Publikationen der Deutschen
Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation e.V.

Band 30 **2022**

Beiträge

42. Wissenschaftlich-Technische Jahrestagung der DGPF

5.-6. Oktober 2022
in Dresden

Dreiländertagung DGPF – OVG – SGPF
Photogrammetrie - Fernerkundung - Geoinformation - 2022

ISSN 0942-2870
DOI: 10.24407/KXP:1795622415

Thomas P. Kersten & Nora Tilly, Hrsg.

Publikationen
der Deutschen Gesellschaft für
Photogrammetrie, Fernerkundung
und Geoinformation e.V.

Band 30 **2022**

Beiträge

42. Wissenschaftlich-Technische Jahrestagung der DGPF

5.-6. Oktober 2022
in Dresden

ISSN 0942-2870
DOI: 10.24407/KXP:1795622415

Thomas P. Kersten & Nora Tilly, Hrsg.

ISSN 0942-2870

Publikationen der
Deutschen Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Band 30, 365 S., Hamburg und Köln 2022
Hrsg.: Thomas P. Kersten & Nora Tilly

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Stuttgart 2022

Geschäftsstelle der DGPF
c/o Universität Stuttgart,
Institut für Photogrammetrie (ifp)
Geschwister-Scholl-Str. 24D
70174 Stuttgart
E-Mail: geschaeftsstelle@dgpf.de

Redaktion:

Thomas P. Kersten
HafenCity Universität Hamburg
Labor für Photogrammetrie & Laserscanning
Henning-Voscherau-Platz 11, 20457 Hamburg
E-Mail: Thomas.Kersten@hcu-hamburg.de

Nora Tilly
Stadt Köln - Amt für Stadtentwicklung und Statistik
Willy-Brandt-Platz 2, 50475 Köln
E-Mail: Nora.Tilly@uni-koeln.de

INHALTSVERZEICHNIS

BEITRÄGE DREI-LÄNDER-TAGUNG 2022

AHMED ALAMOURI, ASTRID LAMPERT & MARKUS GERKE:
New UAS Regulations in the EU and their Impact on Effective Usage of UAS. 1

JENS BRAUNECK & RAINER WIRTH:
UAV-gestützte Untersuchung der durch Blattschneiderameisen verursachten Bioturbation im Catimbau-Nationalpark (NE-Brasilien). 9

MORITZ CAMENZIND, MICHAEL THALER, ULF FEUERSTEIN & KANG YU:
Multi-angular Multispectral Sensing of Barley Awn Coloring. 19

KARIN ERBE, MELANIE BRANDMEIER, MICHAEL SCHMITT, ANDREAS DONAUBAUER, JAN-ANDREAS LIEBSCHER & THOMAS H. KOLBE:
Detektion von Fahrradständern in Luftbildern mittels Deep Learning. 27

SELINA GANZ & PETRA ADLER:
Biodiversity monitoring based on remote sensing - Assessment of forest height changes considering inaccuracies in image-based digital elevation models. 40

VERENA HÄNDLER, OLIVER REUß, JONAS ROTH, JÜRGEN SELENSKI & ANSGAR BRUNN:
Punktwolkenbasierte 3D-Vermessung für KMU in Handwerk und Bauwesen. 52

PAUL HERZIG, UWE KNAUER, PETER BORRMANN, DAVID KILIAS, HANS-CHRISTIAN KLÜCK, UDO SEIFFERT, KLAUS PILLEN & ANDREAS MAURER:
Bewertung des Phänotyps in der landwirtschaftlichen Züchtungsforschung anhand der photogrammetrisch bestimmten Wuchshöhe und spektraler Indizes am Beispiel verschiedener Genotypen der Gerste. 65

UWE KNAUER, KATHARINA HOLSTEIN, BONITO THIELERT, DENNY SCHMELZ & INA EHRHARDT:
Eschentriebsterben – Erkennung der Eschenvitalität mit hyperspektralen Fernerkundungsdaten. 84

KATRIN KRZEPEK, JAKOB SCHMIDT & DOROTA IWASZCZUK:
Potenzial der Fernerkundung für den Moorschutz – Bewertung von Methoden zur Erkennung von stark bewachsenen Wasserflächen. 94

FRANK LIEBOLD & HANS-GERD MAAS:
3D-Deformationsanalyse und Rissdetektion in multitemporalen Voxeldaten von Röntgentomographen. 105

FELIX LOBERT, NORBERT RÖDER, ALEXANDER GOCHT, MARCEL SCHWIEDER & STEFAN ERASMI:
Mowing Detection from Combined Sentinel-1, Sentinel-2, and Landsat 8 Time Series on Fallow Cropland with Transfer Learning. 117

DAVID MADER, KATJA RICHTER, PATRICK WESTFELD & HANS-GERD MAAS:
Volumetrisches Full-Waveform Stacking zur Ableitung der Gewässerbodentopographie aus Laserbathymetriedaten – Erste Ergebnisse einer Pilotstudie am Fluss Elbe. 127

QUSAI MARASHDEH, LUKAS DREES & RIBANA ROSCHER:
Semantic UAV Image Segmentation of Mixed Cropping Fields. 140

KARAM MAWAS, MEHDI MABOUDI & MARKUS GERKE:
Direct Co-registration of As-built and As-designed Data in Digital Fabrication. 149

SILKE MECHERNICH & ROBERT WEIß:
Fernerkundungs-Aktivitäten zu großräumigen hydrologischen Extremereignissen im Kontext des Nationalen Hochwasserschutzprogramms (NHWSP). 160

THERESA MEYER, ANSGAR BRUNN & UWE STILLA:
Genauigkeitsbetrachtung voxelbasierter Änderungsdetektion im Gebäudeinnenbereich zur automatisierten Baufortschrittsüberwachung. 173

MARKUS MUERTH, MARCO ZULKOWSKI, NINA KRÜGER, INES KOPPENHAGEN & DANIEL HOLWEG:
Automatisierte Einbindung und Verarbeitung von Copernicus-Daten in ein Digitales Lager geotopografischer Daten. 186

RUWEI NIE, MICHAEL CRAMER & PATRICK LUBIG:
Processing of Allied Air Forces World War II Strategic Air Reconnaissance Imagery. 193

FRANZ OKYERE, THEA MINNICH, MORITZ SPROLL, EDMUND MENSAH, LILIAN AMARTEY, CARLOS OTOO-KWOFIE & ANSGAR BRUNN:
Implementation of a low-cost Ambulance Management System. 208

SOPHIE PROKOPH, JASMIN CHEEMA, ANITA KIRMER, ANGELA LAUSCH & LUTZ BANNEHR:
Monitoring von blütenreichen Flächen mittels Fernerkundung. 220

THUNYATHEP SANTHANAVANICH, PATRICK WUERSTLE, RUSHIKESH PADSALA & VOLKER COORS:
Digital 3D City Models Towards Urban Data Platform using OGC 3D GeoVolumes API. 237

ROBIN GILBERT SCHULZE, DAVID MADER & ANETTE ELTNER:
Genauigkeitspotential direkt georeferenzierter UAS-Messungen am Beispiel eines Bahnhaltepunktes. 243

JANIK STEIER, LINA E. BUDDE & DOROTA IWASZCZUK:
Machbarkeitsstudie zur Prädiktion von Waldbränden
aus Sentinel-2 Zeitreihen. 256

ROBERT WEIß & SILKE MECHERNICH:
Beschreibung der geometrischen Qualität eines
Digitalen Geländemodells des Wasserlaufs. 267

WEI ZHANG, DAVID SKUDDIS, PHILIPP J. SCHNEIDER & NORBERT HAALA:
Comparison of different 2D and 3D Sensors and Algorithms
for Indoor SLAM on a low-cost Robotic Platform. 280

Xinlong ZHANG, RUIHANG XUE, Michael KÖLLE & Uwe SÖRGEL:
Point Surfel Transformer Network for Semantic Segmentation
of Large-Scale ALS Point Clouds. 290

BEITRÄGE KARL-KRAUS-NACHWUCHSPREIS

SYLVIA ACKERMANN & LENA JOACHIM:
Simulation und Auswertung eines photogrammetrischen Bildverbandes
aus Krankamerabildern. 297

MONA GOEBEL, LINA E. BUDDE & DOROTA IWASZCZUK:
Edge-related Deep Learning for Semantic Segmentation. 304

ANTONIA GOLAB:
It's also about timing!
When do Pedestrians Want to Receive Route Instructions. 315

YUCHANG JIANG:
Multi-Target Multi-Camera Drone Tracking. 322

SIMON KÖNIG, FRANK THONFELD, OLENA DUBOVYK & MARCO HEURICH:
Combining Multisensor Satellite Remote Sensing Time Series
to Assess Disturbances in Central European Temperate Forests. 330

LISA MANDL & STEFAN LANG:
Early Detection of Bark Beetle Induced Forest Stress
using Sentinel-2 Data. 342

NANDO METZGER:
Forecasting Urban Development from Satellite Images. 355

AUTORENVERZEICHNIS 362

PROGRAMMKOMITEE 365

New UAS Regulations in the EU and their Impact on Effective Usage of UAS

AHMED ALAMOURI[1], ASTRID LAMPERT[2] & MARKUS GERKE[1]

***Abstract**: A key hurdle in most UAS based applications is to conduct a safe UAS operation. To reach the highest safety level, minimizing risks to other airspace users, people & property is requested. Risk avoidance requests a clear presence of legislation for UAS operation. In this vein, UAS regulators in EU are making efforts towards enabling a reliable legal framework. The recent outcome of these efforts is that new EU drone regulations are into force since 1st Jan. 2021. Although the new regulations are seen as an essential step for harmonizing UAS rules in Europe, they may lead to new challenges that influence UAS operations. In this opinion paper, the authors will proceed with background information about the recent UAS rules in Europe and their impact on UAS operation and use.*

1 Introduction

Unmanned aircraft systems (UAS, also known as drones) are increasingly becoming popular and are used in many application and research fields such as topographic mapping, infrastructure maintenance, inspections, atmospheric research, etc. One key hurdle in most UAS-based applications is to operate flight missions with a high safety level. This is due to the fact that realizing a safe UAS operation is not always achievable or possible due to the urgent need for coordination between many factors such as the knowledge of all barriers and obstacles existing in the operation scenes. However, a safe UAS operation imposes minimizing the risks that may arise during the operation time and influence other airspace users, bystanders and property on the ground. In addition, avoiding or minimizing UAS risk needs a legal framework for UAS planning and operation. Such kind of a legalisation should provide rules with clear technical and operational definitions that can support overcoming common problems arising during flight missions; for instance the time of flight operation - e.g. within/outside the peak-hours (RANGO, ALBERT & LALIBERTE 2010), the complexity in the execution of administrative affairs that can hinder the desired operational flexibility (STÖCKER et al. 2017). Such high demand for a reliable legal framework for UAS operations is an important motivation that many national and international authorities and organizations in Europe started with tremendous efforts towards updating and modernizing the first wave of UAS regulations that have been adopted in 2017. The main aim of the update process and the modernization is to keep up with the fast and enormous developments of UAS technology in a way that respects the needs and laws of EU member states. The mentioned efforts have culminated in the adoption of new EU drone regulations that are into force since the beginning of 2021. In gen-

[1] Technical University of Braunschweig, Institute of Geodesy and Photogrammetry, D-38106 Braunschweig, E-Mail: [a.alamouri, m.gerke]@tu-braunschweig.de
[2] Technical University of Braunschweig, Institute of Flight Guidance, D-38106 Braunschweig, E-Mail: astrid.lampert@tu-braunschweig.de

eral, the current EU drone rules are seen as a positive step towards harmonizing the drone regulations in Europe. From a practical point of view, implementing the new rules into national level may need more time, probably to the mid of 2022. This can be justified by the lack of adequate interpretation of some rules in terms of administrative and technical issues (AYAMGA et al. 2021). To this end, we aim in this opinion paper to provide a look back on the first wave of the regulations in 2017. Next, we address the recent developments of EU drone regulations with a focus on the most important amendments and changes in the regulations since 2017. Finally, we point at the impact of new rules on drone use i.e. how they may help or hinder the deployment of drone technology.

2 Briefly looking back on drone regulations (2017-2020)

In this section, we briefly introduce the status of the drone regulations in Europe within the past years, namely the last four years before adopting the new rules at the beginning of 2021. Basically, EU member states have started with the process for drafting or enacting drone legislation. JONES in 2017 reported about this process, and pointed out that it was somehow cumbersome and largely vague due to a lack of requirements or technical expertise. He provided a list of tables showing the requirements for UAS operations in different regions in the world. To give a close view about the regulation status at that time, we refer - for instance - to the first wide wave of drone regulations in Germany approved in spring 2017. In that year, the rules for UAS operation have been laid down with focus a.o. on air traffic management and navigation services (BGBL 2017). One distinctive issue to be noted is that the mentioned regulation wave considered the Maximum Take-Off Mass (MTOM) - allowed for UAS - as a key factor for UAS classification. For this reason, UAS were categorised in three systems: with MTOM less than 5kg, between 5-25kg and more than 25kg. The most important operational requirements for UAS categories are represented in Tab. 1 (CRAMER & WIELAND 2019).

Tab. 1: Important operational requirements of UAS according to their MTOM (German rules 2017-2020)

MTOM	<5KG	[5-25]kg	>25kg
Flight permission	Not necessary for each flight	requested	a fundamental operation is prohibited, although exceptions are possible
Labeling of UAS	Mandatory	Mandatory	
Maximum flight height	<100m	<100m	
Pilot certificate of knowledge	Requested	Requested	

In the end of 2017, MTOM was not considered anymore in UAS classification, where the European Council, European Commission and the European Parliament have proposed a new concept regulating all UAS regardless of their MTOM (EASA 2017-05). This concept was developed in the following year with focus on realizing a legal framework that enables a high uniform level of safety in the European civil aviation (EPC 2018/1139). For this aim, effective aviation policies in EU were achieved in 2019 (EASA/947 2019), and their main outcome was that new requirements for technology and personnel were included in the new EU rules for UAS operations adopted in the end of 2020 (EASA_Rev. 2021). This will be highlighted in the next section.

3 New EU drone regulations

After a lot of efforts towards UAS rules harmonization, EU Aviation Safety Agency (EASA https://www.easa.europa.eu/) laid down the new UAS regulations in December 2020. Basically, the new rules have been developed based on the core notices of the amendment A-NPA proposed in 2015 (EASA_NPA 2015). The new rules are structured in two main parts: the first one is Implementing Rules (IR) that cover issues related to operations of UAS. While the second part is Delegated Rules (DR), which deals with technical requirements for the design and manufacture of UAS (EASA_Rev. 2021). In the following sections, IR and DR will be highlighted.

3.1 Implementing rules (IR)

IR have been formulated through 23 articles that deliver comprehensive provisions for UAS operations. Key articles are those focusing on a safe UAS operation, personnel and organizations involved in UAS operations. In this context, IR classified UAS operations in three categories based on the operation risk that may be arisen or posted to third parties (EASA_Rev. 2021); these categories are: open, specific and certified (Tab. 2).

Tab. 2: Categories of UAS operations

UAS operational categories		
Open	Specific	Certified
• Low risk • No involvement of aviation authority • VLOS based operation	• Medium risk • SORA based approval • An operating permit is required	• High risk • Certified operator • Certified UAS

One main issue in the open category is a low risk during UAS operations. Low risks are seen as an advantage in UAS operations, because it enables UAS pilots to fly without the need for an operation permission. For further characteristics of the open category, we can refer to the article 4 of IR (EASA_Rev. 2021): *"UAS operations are classified in the open category only where the following requirements are met:*

- *UAS has a class that is set out in DR EU 2019/945 (see 3.2)*
- *UAS MTOM should be less than 25kg*
- *UAS operation is conducted in the Visual Line of Sight VLOS and the UAS is kept at a safe distance of at least 1.5km from inhabited areas, airports and sensitive zones, and at least 100m from infrastructure like highways, hospitals, power plants, etc.*
- *During an operation, UAS do not carry dangerous goods and do not drop any material*
- *Flying height is limited to 120m above the surface of the Earth".*

If one or more of the abovementioned requirements of the open category cannot be complied within UAS operation, the specific category will be in force. The assumption here is that a higher operation risk can be expected. In this vein, UAS pilots need an operational permission issued by the competent authority pursuant to the risk assessment or the robustness of measures that keeps a high safety level during the operation - more details can be found in the article 12 of IR

(EASA_Rev. 2021). Based on this, UAS pilots need to obtain an operation licence connected to a risk assessment for UAS activities in the specific category. According to the article 11 of IR (EASA_Rev. 2021), *"an operational risk assessment shall include, but not limited to:*

- *A solid description of UAS operation*
- *propose comprehensive operational safety objectives*
- *identification of ground and air risks to, for instance uninvolved persons, objects, etc.*
- *measures for risk mitigation*
- *technical characteristics of the UAS and competencies of the personnel"*

IR proposed a methodology that can be applied to conduct the abovementioned risk assessment; this is known as the Specific Operation Risk Assessment (SORA). SORA has been developed by the Joint Authorities for Rulemaking on Unmanned Systems JARUS (http://jarus-rpas.org/). More details about this approach and its concept are available at the article 11 of IR (EASA_Rev. 2021). Within the third operation category, UAS activities are regulated according to the article 6 of IR. Here UAS operations shall be classified in the certified category if only the following requirements are met (EASA_Rev. 2021):

- *"if the UAS is certified pursuant to Article 40 of DR; here it deals with the design, production and maintenance of UAS*
- *if the operation is conducting over assemblies of people, involves the transport of people or involves the carriage of dangerous goods that may result in high risk to other parties*
- *if the competent authority may assess the operational risk such that the operation falls into the certified category"*

Beside the UAS certification according to the article 40 of DR, it is necessary to certify the entire operation, i.e. the company involved in the operation, pilots, the measures of UAS maintenance, the monitoring of the maintenance Continuing Airworthiness Management Organization (CAMO).

3.2 Delegated Rules (DR)

The subject matter of DR is the design and manufacture of UAS to be operated with respect to the IR introduced in (3.1). In this paper, we focus on DR that define the type and class of UAS to be certified; which is - from our point of view - an important topic for UAS operation. In this context, and from operational aspects, the new EU regulations provided a class-identification of UAS that shows on one hand the technical properties of UAS for e.g. MTOM, and on the other hand the operational requirements to be respected during flight missions. As result, UAS are classified into seven classes C0 to C6. Most important characteristics of these classes are presented in Tab. 3.
However, UAS regulators achieved a linkage between UAS classes and the operation distance to persons, which has an impact on the operation categories described in (3.1), namely on the open category. This led to the requirement that the open category got structured into three subcategories: A1, A2 and A3. For instance, when operating UAS with class C0 in A1, pilots do not need an operation license, but they have to be familiar with the UAS user manual guide. In contrast, flying in A1, but with UAS class C1 imposes that pilots should have a sold background about the UAS user manual guide and passed an online training course.

Tab. 3: UAS classes C0–C6, source (EASA_Rev. 2021)

Subcategory	UAS Class	MTOM (kg)	Max. velocity (m/s)	Max. AGL (m)	Qualification
A1 Fly over people	C0	<0.25	19	120	Familiar with operation instructions
	C1	<0.9	19	120	Familiar with operation instructions Online training and test
A2 Fly close to people	C2	<4	-	120	Familiar with operation instructions Online training and test Certificate "proof of knowledge" (according to German rules)
A3 Fly far from people	C3	<25 < Diameter 3m	-	120	Familiar with operation instructions Online training and test
	C4	<25	-	-	
-	C5	Not defined	-	-	
-	C6	Not defined	50	-	

4 Regulation impact on UAS Operation

Since the adoption of the new EU drone rules in early 2021, drone users and manufacturers have been following the repercussions that may result from applying these rules. Although the new regulations are seen as a positive step towards modernizing and harmonizing UAS rules in Europe, they will lead to some challenges that affect UAS operation and uses. More specifically, in the following subsections we discuss two challenges that may be crucial in the early phase of rules implementation, and these are: UAS user registration and visibility restriction.

- **UAS user registration**

In order to realize a better accommodation of UAS operations, registration of UAS user (pilot or operator) within EU in the respective EU country is only required once. The UAS user's ID assigned during the registration must then be visibly fixed to each drone. The registration process itself depends on UAS design, class (see Tab. 3) and on the operation category where the UAS should be operated. In this context and according to the article 14 of IR (EASA_Rev. 2021), "*UAS operators shall register themselves:*

a) when operating within the open category any of the following UAS:
   i. MTOM is 250g or more
   ii. UAS is integrated with a payload like for instance a sensor that could be used for personal data collection
b) when operating within the specific category an unmanned aircraft of any mass"

To this end, we agree that the regulations addressed clearly who should register and when, but at the same time, having a competent registration system in a practical way is still a key challenge. This is because each EU country needs to create specific databases, and to create an online platform for the registration process, which requires skills for data managements and data provision. These requirements do not yet exist in many EU member states, and therefore the interoperability, mutual access and exchange of registration information might be affected. In addition, the need to indicate the operator information at each device physically might bring challenges in practice, e.g. when several people are using the same system.

- **Visibility restriction**

As previously mentioned in section (3.1), the open category is denoted as being low risk during UAS activities and imposes operating the UAS with VLOS conditions, which means that pilots must keep continuous visual contact with UAS. From an economic point of view, VLOS based operations may not be the best option for various UAS based applications; for instance, VLOS operations may not be the best scenario for monitoring issues in dense forest researches, where keeping a direct contact to UAS is not always possible (PANEQUE-GÁLVEZ et al. 2014; YAUN et al 2015; THIEL et al. 2020; DAINELLI et al. 2021). In contrast, Beyond Visual Line of Sight (BVLOS) based flight missions are a significant aspect for the drone operation (POLITI et al. 2021). This is due to the fact that in BVLOS based operations UAS can be operated away from the visual range. Therefore, it allows UAS covering long distances and large areas, which is interesting in terms of time and cost reduction. In this way, BVLOS opens chances for a large variety of services such infrastructure inspection, deliveries, etc. Nevertheless, BVLOS operations are still limited in current regulations, and regulators in EU are making efforts to develop and adopt clear rules for BVLOS operations (WACKWITZ & SCHROTH 2021). In this direction, we refer to the efforts towards developing the Standards and Recommended Practices (SARPs) for BVLOS operations that started in 2018 with adoption foreseen in 2020 – but not realized – and leading to operations from 2023 on (SESAR 2018). The mentioned developments are focusing on BVLOS main challenges such as security, safety and communications (POLITI et al. 2021), which are considered critical issues hindering the wide use of BVLOS operations.

From a practical point of view, and to understand the potential of VLOS or BVLOS based operations, we provide in Fig. 1 a comparison between typical applications. One can see that there is a tendency in the UAS markets to adopt BVLOS techniques for better UAS functionalities and therefore to open new commercial services such as for first responders, package delivery, etc.

However, the decision to conduct a VLOS or BVLOS operation depends on various factors such as safety, security, time, cost, etc.; and this will influence the selection of the operation category where UAS should be operated. In this context, UAS flights in the specific and certified categories are allowed with BVLOS conditions. This might be seen as a practical solution to get rid of VLOS confinements, but at the same time it requires additional processes that lead to more administrative and bureaucratic complexities that caused somehow a reluctance to conduct flights in these categories.

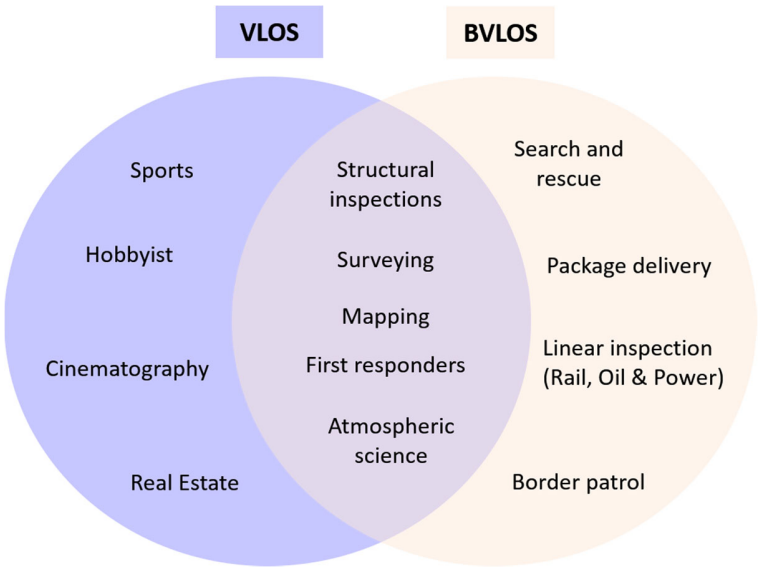


Fig. 1: VLOS- and BVLOS-based typical applications

5 Summary

Due to the rapid development of UAS technology and its usage, drone regulators in EU - and worldwide - are striving to integrate drones into their aviation regulation frameworks. In this paper, after a brief review of drone regulations from 2017 to 2020, we introduced the new EU drone regulations adopted in the beginning of 2021. The new rules are considered a positive step towards rules modernization and harmonization in Europe. They are structured into two parts: implementing and delegated rules, and detailed guidelines of how to define and conduct UAS operations with respect to critical factors such as safety, security, privacy, etc. The implementing rules deal with issues related to realizing safe UAS operations and avoiding risks to persons and organizations involved in UAS operations. While the delegated rules focus on issues related to UAS design and manufacture. Finally, we showed that the new regulations will lead to some challenges that may be faced in the early stage of rules implementation, here we discussed two aspects: UAS user registration and visibility restriction. However, there are other challenges that may be crucial like the risk assessment of UAS operations. In this context, a risk assessment should reflect the safety level expected during UAS operation, and this needs a collaborative effort from different responsibilities and processes that extend beyond a single agency or organization.

6 References

AYAMGA, M., TEKINERDOGAN, B. & KASSAHUN, A., 2021: Exploring the Challenges Posed by Regulations for the Use of Drones in Agriculture in the African Context. Land, **10**(2), 1-13. https://doi:10.3390/land10020164.

BGBL, 2017: Verordnung zur Regelung des Betriebs von unbemannten Fluggeräten. https://www.bgbl.de/xaver/bgbl/start.xav#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgb

l117s0683.pdf%27%5D__1642074397859.

CRAMER, M. & WIELAND, M., 2019: UAS-Regulierung, oder: "Nichts ist beständiger als der Wandel". PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **87**(3), 123-135.

DAINELLI, R., TOSCANO, P., DI GENNARO, S. F. & MATESE, A., 2021: Recent Advances in Unmanned Aerial Vehicles Forest Remote Sensing - a Systematic Review. Part II: Research Applications. Forests, **12**(4). https://doi:10.3390/f12040397.

EASA_NPA, 2015: Advance Notice of Proposed Amendment 2015-10 Introduction of a Regulatory Framework for the Operation of Drones. 1-41.

EASA_Rev., 2021: Easy Access Rules for Unmanned Aircraft Systems Sept 2021. 1-292.

EASA/947, 2019: Implementing Regulation 2019/947.

EASA, 2017: Notice of Proposed Amendment 2017-05 (A) - Introduction of a Regulatory Framework for the Operation of Drones. 05, 1-128.

EPC, 2018: Regulation (EU) 2018/1139 of the European Parliament and of the Council. *Official Journal of the European Union* 2018(July), 1-122.

JONES, T., 2017: International Commercial Drone Regulation and Drone Delivery Services. International Commercial Drone Regulation and Drone Delivery Services. https://doi:10.7249/rr1718.3.

PANEQUE-GÁLVEZ, J., MCCALL, M. K., NAPOLETANO, B. M., WICH, S. A. & KOH, L. P., 2014: Small Drones for Community-Based Forest Monitoring: An Assessment of Their Feasibility and Potential in Tropical Areas. Forests, **5**(6), 1481-1507. https://doi: 10.3390/f5061481.

POLITI, E., PANAGIOTOPOULOS, I., VARLAMIS, I. & DIMITRAKOPOULOS, G., 2021: A Survey of UAS Technologies to Enable Beyond Visual Line Of Sight (BVLOS) Operations. June, 505-512. https://doi:10.5220/0010446905050512.

RANGO, A. & LALIBERTE, A., 2010: Impact of Flight Regulations on Effective Use of Unmanned Aircraft Systems for Natural Resources Applications. Journal of Applied Remote Sensing**,** **4**(1), 043539. https://doi:10.1117/1.3474649.

SESAR, 2018: European ATM Master Plan : Roadmap for the Safe Integration of Drones into All Classes of Airspace. 1-33.

STÖCKER, C., BENNETT, R., NEX, F., GERKE, M. & ZEVENBERGEN, J., 2017: Review of the Current State of UAV Regulations. Remote Sensing**,** **9**(5), 33-35. https://doi:10.3390/rs9050459.

THIEL, C., MÜLLER, M., BERGER, C., CREMER, F., DUBOIS, C., HESE, S., BAADE, J., KLAN, F. & PATHE, C., 2020: Monitoring Selective Logging in a Pine-Dominated Forest in Central Germany with Repeated Drone Flights Utilizing a Low Cost Rtk Quadcopter. Drones, **4**(2), 1-26. https://doi:10.3390/drones4020011.

WACKWITZ, K. & SCHROTH, L., 2021: BVLOS Operations Report 2021. Drone Idustry Insights.

YAUN, C., LIU, Z. & ZAHNG, Y., 2015: UAV-Based Forest Fire Detection and Tracking Using Image Processing Techniques. IEEE International Conference on Unmanned Aircraft Systems, 639-643.

UAV-gestützte Untersuchung der durch Blattschneiderameisen verursachten Bioturbation im Catimbau-Nationalpark (NE-Brasilien)

JENS BRAUNECK[1] & RAINER WIRTH[2]

Zusammenfassung: Gegenstand dieses interdisziplinären Forschungsprojekts sind die ökologischen Auswirkungen anthropogener Störungen auf das Caatinga-Trockenwaldökosystem, eines der am stärksten bedrohten semiariden Naturgebiete Südamerikas. Vegetationssukzession und Ökosystemfunktionen werden im Untersuchungsgebiet nachhaltig von Blattschneiderameisen beeinflusst, da sich diese Ameisen aufgrund der anthropogenen Landnutzungseffekten drastisch vermehren und zu dominanten Schädlingen werden können. Mit Hilfe UAV-basierter Vermessung soll die Bioturbationsleistung der Ameisen auf Landschaftsebene quantifiziert werden. Dazu werden sowohl volumetrische Berechnungen auf Nestebene durchgeführt als auch die Verbreitung der Nester in der Landschaft per UAV-Kartierung erfasst.

1 Einleitung

Im Rahmen eines binationalen Forschungsprogramms wird aktuell in einer Kooperation der TU Kaiserslautern und der Universidade Federal de Pernambuco, Brasilien (DAAD, PROBRAL-ID: 57576555) analysiert, wie sich chronisch anthropogene Störungen auf Ökosystemfunktionen Desertifizierung und Regenerationsdynamik des nordostbrasilianischen Caatinga-Trockenwaldes auswirken. Blattschneiderameisen der Gattung *Atta* spielen in diesem Ökosystem eine Schlüsselrolle als dominante Pflanzenfresser und Ökosystem-Ingenieure, die sich infolge menschlicher Landnutzung drastisch vermehren und zu dominanten Schädlingen werden können. Im Fokus der Untersuchungen stehen die langlebigen, bis zu 200 m^2 großen Ameisennester und deren Effekte auf die chemisch-physikalischen Eigenschaften des Oberbodens (MEYER et al. 2013; BIEBER et al. 2011).

Die Erkennung und Bestandsaufnahme der Blattschneiderameisennester ist für eine Abschätzung des Einflusses auf die Landschaftsentwicklung notwendig, stellt in dem je nach Vegetationsdichte unzugänglichen Gelände eine besondere Herausforderung dar, da sie kaum über herkömmliche freie Satellitendaten erkannt werden können. Der Einsatz von UAV soll helfen, hier die Grundlage für topografische und ökologische Erhebungen auch in abgelegenen Gebieten zu schaffen. Die UAV-Technologie schließt hier eine methodische Lücke zwischen satelliten- und bodengestützter Datenerfassung. In der einfachsten Konfiguration können mit Digitalkameras ausgerüstete UAVs hochauflösende digitale Luftbilder des Geländes erstellen, die dann mittels einer geeigneten Photogrammetrie-Software analysiert und in ein dreidimensionales digitales Oberflächenmodell

[1] Frankfurt University of Applied Sciences, FB 1, Labor für Laserscanning, Photogrammetrie und Fernerkundung, Nibelungenplatz 1, D - 60318 Frankfurt am Main, E-Mail: jens.brauneck@fb1.fra-uas.de
[2] TU Kaiserslautern, Molekulare Botanik (Sektion Pflanzenökologie), Erwin-Schrödinger-Str. Geb. 70/219, D - 67653 Kaiserslautern, E-Mail: wirth@rhrk.uni-kl.de

umgerechnet werden können. Mit dieser Methode soll sowohl eine Kartierung der Nestvorkommen als auch eine direkte Ableitung der Bioturbationsleistung der Ameisen vorgenommen werden.

2 Daten und Methoden

Die Untersuchungen wurden im Catimbau-Nationalpark (NPC) durchgeführt, einem 607 km² großen Gebiet im nordostbrasilianischen Caatinga-Trockenwald im Zentrum des Bundesstaates Pernambuco. Der Nationalpark ist durch ein stark heterogenes Landnutzungsmosaik mit sehr unterschiedlichem Ausmaß an anthropogenen Störungen und einem ausgeprägten W-O verlaufenden Niederschlagsgradienten (650 bis 1100 mm mittlerer Jahresniederschlag) gekennzeichnet. In der Caatinga-Kulturlandschaft wurde bereits etwa 60 % der ursprünglichen Waldbedeckung durch Sekundärvegetation ersetzt oder weitgehend degradiert (SILVA et al. 2017).

2.1 Problemstellung

Pilzzüchtende Ameisenarten der Gattung *Atta* sammeln Biomasse in Form frischer Blattfragmente, um darauf einen Pilz in ihrem unterirdischen Nest zu kultivieren (WEBER 1966), der ihnen als Nahrungsgrundlage dient. *Atta*-Nester sind allgegenwärtig in tropischen und subtropischen Lebensräumen Amerikas von semiariden Grasländern, Savannen, Trocken- und Regenwäldern bis hin zu Weideland und landwirtschaftlich genutzten Flächen (FOWLER & CLAVER 1991). Die kolossalen Nester der Gattung bestehen aus bis zu 8000 unterirdischen miteinander verbundenen Pilz-Kammern, die bis zu 7–8 m unter der Erde reichen (MOREIRA et al. 2004) sowie einem umfangreichen, komplex verzweigten Tunnelsystem zum Transport von Blättern und zur Belüftung und Drainage des Nestes (WIRTH et al. 2003).

Die Nester zeigen oberflächlich zahlreiche Auswurf- und Belüftungshügel (s. Abbildung 1), unterirdisch zeigen sie sehr komplexe Strukturen und sind stark verzweigt. Die sichtbare Nestoberfläche nimmt mit dem Alter zu, da die Neststruktur ein direktes Produkt der wachsenden Ameisenkolonie ist (FOWLER et al. 1986). Dementsprechend nehmen auch die unterirdischen Wühltätigkeiten in ausgewachsenen Nestern zu, um den ständig wachsenden Pilz zu versorgen (RÖMER & ROCES 2015), da bei einigen Arten die entstehenden Abfälle, wie auch der gesamte ausgehobene Mineralboden außerhalb des Nests abgelagert werden (LEAL et al. 2014).

Diese kontinuierlichen Bioturbationsprozesse während des Nestbaus verändern beständig die chemisch-physikalische Struktur und die Nährstoffverfügbarkeit der Böden in unmittelbarer Nähe der Nester. Die durch die Grabungsaktivitäten entstehenden Nesthügel werden durch flächenhafte Wassererosion immer wieder eingeebnet und das Material in die unmittelbare Nestumgebung verlagert. Durch den in aller Regel geringeren Gehalt an Nährstoffen und insb. organischem Kohlenstoff in den unteren Bodenhorizonten (MEYER et al. 2013) erscheint die Nestoberfläche in Luftbildern erkennbar heller als die Umgebung (Abb. 1). Niedrige Vegetation an der Nestoberfläche stirbt bei größeren, resp. älteren Populationen durch die Grabungsaktivitäten und die damit einhergehende reduzierte Wasserverfügbarkeit ab.

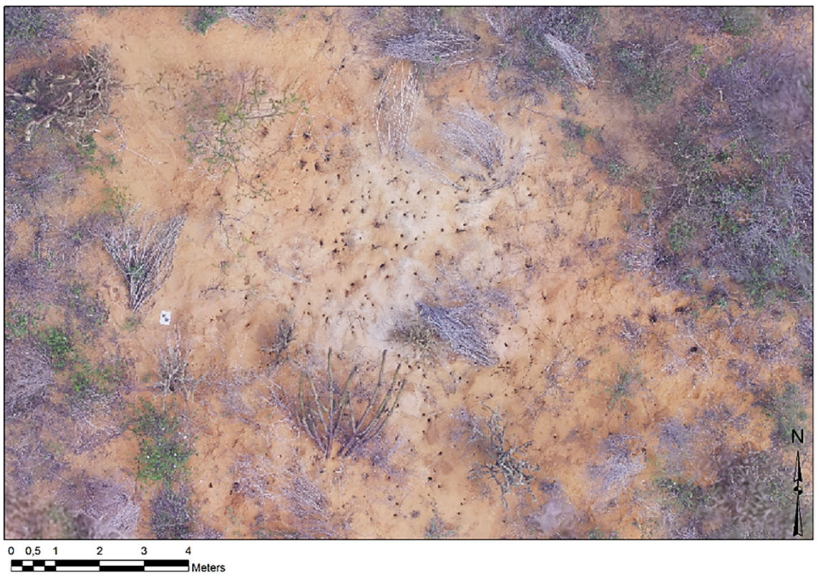


Abb. 1: Orthofoto einer Nestoberfläche der Art Atta opaciceps im Untersuchungsgebiet NPC, NE-Brasilien

2.2 Geländearbeiten

Gegenstand der Untersuchungen war die Art *Atta opaciceps* - die häufigste der Blattschneiderameisenarten im NPC und die einzige endemische Art im Caatinga-Trockenwald (SIQUEIRA et al. 2017). Die Studie beschränkt sich auf drei Teilgebiete des NPC, in denen ausschliesslich *A. opaciceps* vorkam. Sie wurden aufgrund der lokalen Merkmale und bestehender Ortsnamen "Brejo", "Canyon" und "Casa-d'água" genannt (Abb. 2). Auf allen Untersuchungsflächen wurde früher Landwirtschaft betrieben, erst in jüngerer Zeit enwickelte sich die Nutzung als extensive Viehweiden, wobei Viehhaltung in dem Gebiet insgesamt rückläufig ist. In früheren sowie in allen aktuellen Felduntersuchungen wurden in den Referenzgebieten sämtliche Ameisenkolonien erfasst, indem alle sichtbare Neststrukturen lokalisiert und deren Position mit einem GNSS-Handgerät aufgezeichnet wurde. Das Vorhandensein erkennbarer Nesthügel galt als primäres Merkmal für die Identifikation eines Ameisennests. Zusätzlich wurde die Nestaktivität nach den folgenden Anzeichen ermittelt; (1) Vorhandensein aktiv genutzter Ameisenstraßen, (2) Existenz von frisch geschnittenem Blattmaterial auf den Ameisenstraßen oder in unmittelbarer Nähe der Nesteingänge, und (3) das Erscheinen von Arbeitern nach einer Störung am Nesteingang. Für jede Kolonie wurde die Artzugehörigkeit mittels einschlägiger morphologischer Merkmale verifiziert.

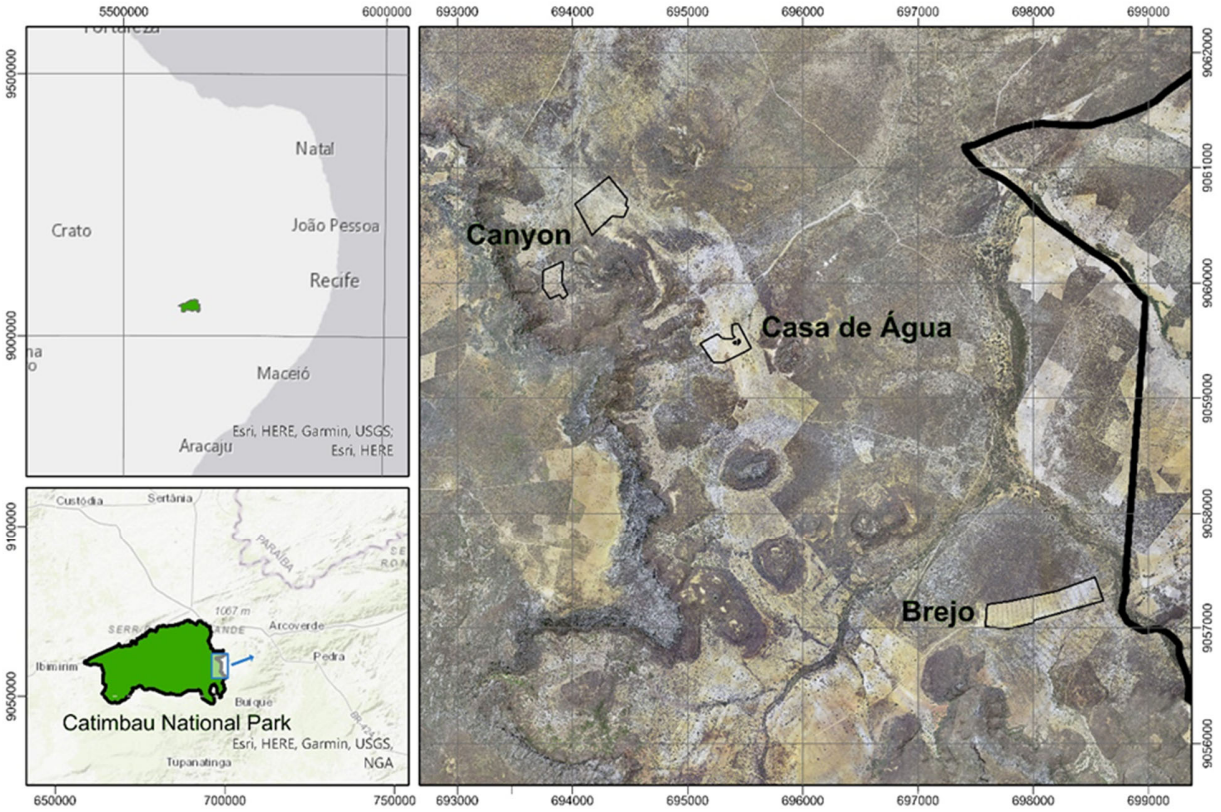


Abb. 2: Lage der Untersuchungsgebiete "Brejo", "Canyon" und "Casa-d'água"

Alle Untersuchungsgebiete wurden 2018 mit Hilfe von UAV-Flügen kartiert, die zu hochauflösenden digitalen Orthofotomosaiken mit einer hohen Auflösung (engl. high resolution, kurz: HR) von 5x5 cm und digitalen Oberflächenmodellen mit einer Auflösung von 10 cm führten. Für die Luftaufnahmen wurde ein handelsüblicher Quadrocopter (DJI Phantom III) verwendet, dessen Flughöhe von 50 m (Kartierung der Untersuchungsgebiete) bis hinunter zu 2 m (Detailaufnahmen einzelner Nester) reichte. Die Onboard-Kamera verwendet einen 1/2,3"-CMOS-Sensor und bietet eine effektive Auflösung von 12,4 Megapixeln. Das Objektiv hat eine Brennweite von 20 mm. Die Luftbilder wurden auf der Grundlage des integrierten GPS vorreferenziert und mit einem bestehenden Datensatz aus dem staatlichen Programm "Pernambuco Tridimensional" korrigiert.

Die erreichbare Genauigkeit in Höhe und Lage hängt von einer Vielzahl an Kriterien ab, darunter sind Flughöhe, Güte und Kalibrierung der Kameras oder die Ausgangspunkte zur Referenzierung (SANZ-ABLANEDO et al. 2018). Die relativen Genauigkeiten für handelsübliche UAV-basierte Oberflächenmodelle aus photogrammetrischen Auswertungen liegen bei etwa dem 1-3-fachen der Bodenauflösung, was für die Untersuchungen im Nahbereich als ausreichend angesehen wird. Insbesondere für niedrige Flughöhen besteht zusätzlich die Möglichkeit, Oberflächen und Kubaturen mit hohen Genauigkeiten auch ohne Bodenkontrollpunkte zu erfassen, wenn zusätzlich Maßstäbe verwendet werden können (MADOUBI et al. 2021). Zusätzlich zu den Erhebungen aus großer Höhe wurden sehr hochauflösende Bilder (engl very high resolution, kurz: VHR) von 12 identifizierten Neststandorten sowohl per Drohne als auch per Bodenfotografie aufgenommen, um

eventuelle Unterschiede in der Genauigkeit ableiten zu können. Da die absolute Genauigkeit für diese Frage weniger wichtig war, wurden als Kontrolle lokale Referenzobjekte auf der Grundlage von digitalen 12-Bit-Markern und Schachbrettzeichen verwendet. Diese Bilder wurden als Grundlage zur volumetrischen Modellierung der Bioturbation einzelner Ameisennester verwendet. Die daraus enstandenen Modelle werden im Folgenden als VHR-Modelle bezeichnet.
Im Dezember 2021 erfolgten erneut Befliegungen in den Untersuchungsgebieten. Dazu wurde ein DJI P4 RTK Quadrokopter eingesetzt. Die Onboard-Kamera verwendet einen 1"-CMOS-Sensor und bietet eine effektive Auflösung von 20 Megapixeln. Das Objektiv hat eine Brennweite von 24 mm. Mit Flughöhen von 90 m konnten digitale Orthofotos mit ca. 3×3 cm Auflösung erzeugt werden. Analog zu 2018 wurden auch Detailaufnahmen einzelner Nester mit Flughöhen unter 5 m erzeugt, um Veränderungen in der Oberflächenstruktur erfassen zu können.

2.3 Nestkartierung

Zur Erfassung der Verbeitung sowie zur Kalkulation der durchschnittlichen Nestoberflächen wurden die bekannten, per GNSS markierten Nester innerhalb einer GIS-Umgebung (ArcGIS Pro) in den Orthofoto-Mosaiken (HR) identifiziert und anschließend auf dieser Grundlage manuell kartiert. Für jedes detektierte Nest in den Orthofotomosaiken beider räumlichen Auflösungen (HR & VHR) wurde die jeweilige Nestfläche durch Erfassung der Form entlang der Nestkontur bestimmt.
Eine methodische Unterscheidung musste zwischen einem ökologischen und dem technischen Kartierungsansatz getroffen werden. Der methodische Unterschied bestand darin, dass beim sog. technischen Ansatz konsequent die gesamte Vegetation um das Nest herum ausgeschlossen und nur der unbedeckte Boden erfasst wurde, um eine Referenz für spätere spektrale Untersuchungen der helleren Nestflächen zu schaffen. Bei dem sog. ökologischen Ansatz wurde der gesamte Nestbereich auf der Grundlage des biologischen Fachwissens über die Struktur von Ameisennestern erfasst, einschließlich abgelegener Eingänge ohne sichtbare Ablagerungen. Die verschiedenen Ansätze sollen es später ermöglichen, aus der automatischen Erfassung der hellen Nestoberfläche auf die Gesamtgröße eines Nests schließen zu können (Abb. 4).
Alle bekannten Ameisennester wurden anhand vorgegebener Kriterien als Polygone in einer GIS-Umgebung kartiert (FINKLER 2020). Dabei konnte während der Kartierung bereits auf den Zustand des Nests (aktiv/inaktiv/zerstört) geschlossen werden.
Im Allgemeinen heben sich die Nester in ihrem Kerngebiet aufgrund des helleren Farbtons des Mineralbodens am deutlichsten von ihrer Umgebung ab. Aufgrund von stark variierenden Bodenfarben in den Untersuchungsgebieten und den zum Rand hin graduellen Farbverläufen erwiesen sich die Standardmethoden zur Kantenerkennung als nicht geeignet. Höhere Vegetation oberhalb eines identifizierten Ameisennests wurde bei der Kartierung nicht ausgelassen aber in den der Bioturbationsberechnung zugrundelegenden Punktwolkenmodellen entfernt. Problematisch für eine Detektion über Luftbilder erscheinen Nester unterhalb von Bäumen und hohen Sträuchern mit hohen Anteilen an Bedeckung.
Betrachtet man nur diejenigen Nester, die einwandfrei erkennbar und nicht weitgehend zerstört sind, konnten 66 Nester einwandfrei kartiert werden. In den Ergebnissen zeigte sich zusätzlich, dass die durchschnittliche Größe der Nestoberflächen zwischen den Untersuchungsgebieten

variiert. So liegen die durchschnittlichen Nestflächen in den Gebieten Canyon und Casa ca. 30 bzw. 20 m² über denen im Gebiet Brejo. Inwiefern hier die Bodenbedingungen oder das Nahrungsangebot eine Rolle spielen, muss in weiteren Untersuchungen geklärt werden.

Tab. 1: Ergebnisse der Nest-Kartierung in den drei Untersuchungsflächen mit 5×5 cm Auflösung HR

	Anzahl Untersuchungsnester	Mittlere Nestgröße (m²)	Minimale Nestgröße (m²)	Maximale Nestgröße (m²)
Brejo	38	72,74	36,15	215,12
Canyon	21	100,89	37,63	180,43
Casa	7	92,92	48,21	157,46
gesamt	***66***	***83,84***		

2.4 Berechnung der Bioturbation

Zwölf aktive, bereits mehrjährig bestehende Ameisennester wurden 2018 per UAV mit einer resultierenden Bodenauflösung im Orthofoto von 5x5 mm (VHR) erfasst. Diese Bilder bildeten die Grundlage für die volumetrische Modellierung der Bioturbation einzelner Ameisennester. In einem weiteren Verarbeitungsschritt wurde das Nesthügelvolumen dieser detaillierten Nester auf der Grundlage digitaler Oberflächenmodelle (DOM) berechnet. Für die Berechnung wurden die durch ökologische Konturierung ermittelten mittleren Nestflächen berücksichtigt.

Um der komplexen Topographie Rechnung tragen zu können und den Effekt von geneigten Oberflächen zu kompensieren, wurde die volumetrische Berechnung direkt an die Nestmorphologie angepasst. Bei dieser Vorgehensweise wird das Volumen jedes Nests in Bezug auf eine vorgegebene virtuelle Oberfläche als normalisiertes digitales Oberflächenmodell (nDOM) gebildet. Diese Bezugshöhe wird direkt durch die Lage und Höhe der Nestkontur vorgegeben.

Um dies zu erreichen werden zunächst sämtliche Polygonstützpunkte (engl. vertices) der kartierten Flächen mit den an dieser Stelle ermittelten Höhen aus den Flächenmodellen versehen. Mit Hilfe dieser Punkte wird eine interpolierte Fläche konstruiert, die im weiteren Verlauf als Referenzhorizont oder Urgelände dient. Dieses Urgelände schneidet die eigentliche Neststruktur und stellt eine vereinfachte Geländeoberfläche ohne Nest dar.

Subtrahiert man diese Fläche von dem digitalen Oberflächenmodell mit Nest, erhält man ein Differenzmodell in Form eines normalisierten digitalen Oberflächenmodells (nDSM), bei dem die ursprüngliche Geländeoberfläche eine Höhe von 0 hat (Abb. 3). Diese Methode funktioniert auch bei komplexem Relief, jedoch ist es hier wichtig zu berücksichtigen, wie hoch die Kubaturen im Vergleich zum gesamten Höhenunterschied sind. Je größer der Unterschied zwischen Nesthöhe und Relief, und je geringer die allgemeine Rauheit der Oberfläche, desto geringer wird der Fehler bei der Berechnung der Kubatur sein. Die hier kartierten Nester sind grundsätzlich auf schwach geneigten Stufenflächen zu finden. Die Stufenhänge sind frei von Nestern. Größere Fehler aufgrund starker Reliefunterschiede sind daher nicht zu erwarten.

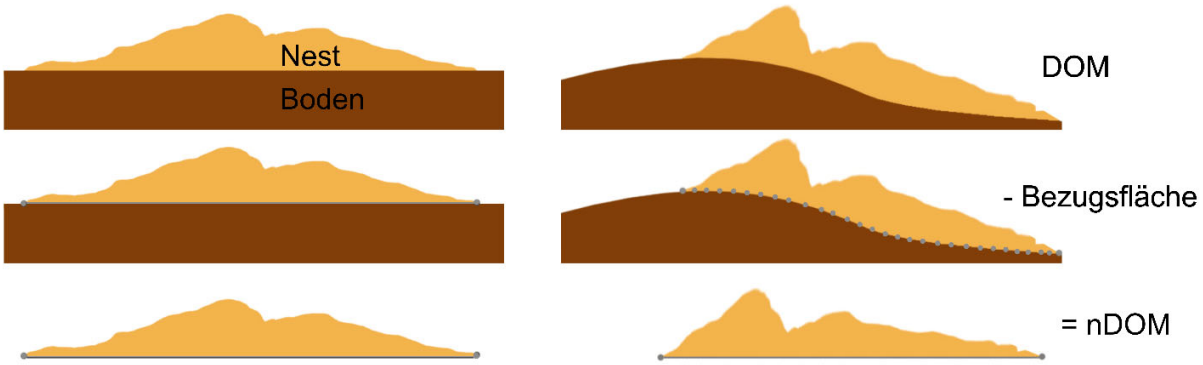


Abb. 3: Schematische Ableitung eines nDOM

Die Berechnung von oberflächlichen Erdvolumina kann grundsätzlich mit unterschiedlichen Methoden, wie Dreiecksvermaschung oder Trapezberechnung erfolgen. In aller Regel sind dabei zunächst Flächen zu berechnen, die im Anschluss eine Höhenkomponente und eventuell eine Bezugshöhe erhalten. Bei digitalen Geländemodellen bietet sich die Berechnung über die Rastermethode an. Die Fläche liegt im digitalen 2,5 D-Modell in Form gleichmäßiger quadratischer Zellen (Fläche eines Pixels= A) mit jeweils einem Höhenwert (h) vor. Bei der Bezugshöhe Null ergibt sich für das Volumen einer Zelle die Rechnung V = h * A. Innerhalb von ArcGIS Pro kann mit Hilfe der Funktion „Oberflächenvolumen“ (engl. surface volume, 3D-Analyst) dann das Volumen eines Körpers zwischen einer digitalen Oberfläche und einem Urgelände (Höhe = 0) oder einer beliebigen Bezugsebene berechnet werden. Berechnet man das Volumen oberhalb (aufgeworfene Fläche) und unterhalb der Bezugshöhe (ausgehobene Eingänge) erhält man das gesamte Volumen des oberflächlichen Nestanteils. Zur Kalkulation der Bioturbation wurde jedoch nur das Volumen oberhalb der Bezugsfläche herangezogen (Tab. 2).
Bei der photogrammetrischen Erfassung dieser Daten wird auch die Höhe der Vegetation erfasst und fließt so in das digitale Oberflächenmodell (DOM) ein. Mit Hilfe der technischen Konturierung der Nestoberfläche kann die umgebende Vegetation, d. h. die Vegetation im Randbereich des Ameisennests, aus der Modellierung ausgeschlossen werden, nicht aber die Vegetation, die sich möglicherweise direkt auf der Oberfläche des Nesthügels befindet.

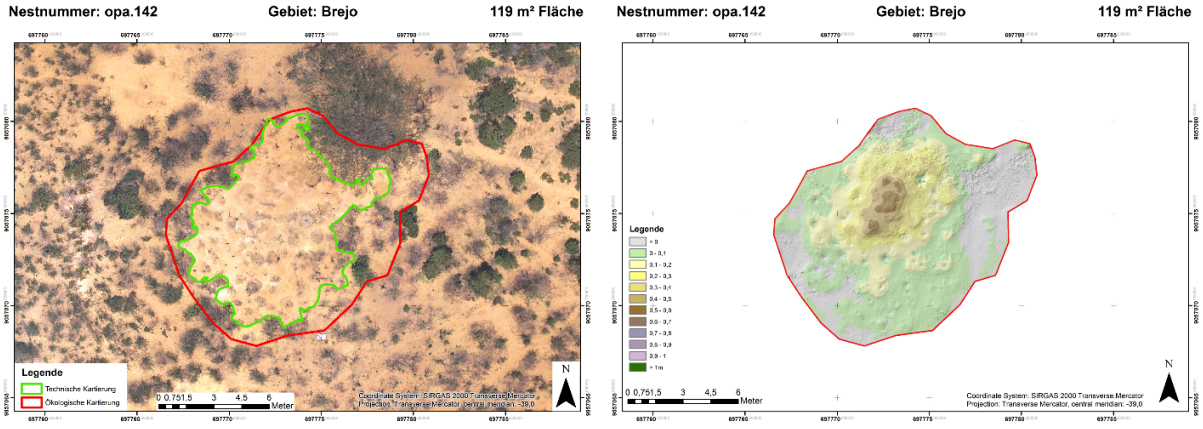


Abb. 4: Kartierung (links) und abgeleitetes nDOM (rechts) eines Nests im Untersuchungsgebiet Brejo

Daher muss die Vegetation entweder manuell oder durch Klassifizierung von Bodenoberfläche und Vegetation entfernt werden, was nicht immer ohne Rückstände möglich ist. Um diese Ungenauigkeiten bei der Darstellung der Nestoberfläche und bei der Bestimmung des Nesthügelvolumens zu reduzieren, wurden die Modelle zusätzlich durch Tiefpassfilterung geglättet.
Von den vorhandenen 12 Nestern musste bei zwei nahe beieinander liegenden Nestern davon ausgegangen werden, dass diese zusammenhängen und nur von einer Kolonie bewohnt werden. Daher reduziert sich die Auswertung auf insg. 11 Nester. Zusammenfassend lässt sich auf dieser Basis ableiten, dass es zum einen auffällige Unterschiede in der maximalen Fläche der Nester gegenüber der Kartierung auf Grundlage der 5 cmm HR-Daten gibt (Tab. 1), da auf der hochaufgelösten Grundlage (VHR) kleine Eingänge, die ebenfalls zur Struktur des Nests gehören besser erkannt werden können und somit weitere Flächen zur Nestfläche einbezogen werden. Farbunterschiede treten hierbei stärker in den Hintergrund. Zum anderen lässt sich erkennen, dass das Verhältnis von ausgehobenem Material zur Nestfläche trotz der Variation in der durchschnittlichen Größe stabil bei ca. 0,1 (m^3/m^2) bleibt. Durch die Berechnung weiterer Nestvolumina im Untersuchungsgebiet die auf Grundlage der Geländekampagne im Dezember 2021 durchgeführt werden können, sind weiterführende Ergebnisse und insbesondere eine Erhöhung der Nestanzahl erwartet.

Tab. 2: Zusammengefasste Ergebnisse der Volumenberechnung von 11 Nesthügeln auf Basis der hochaufgelösten Grundlage (VHR). Angegeben ist das mittlere Bodenvolumen pro m2 Nestoberfläche

	Ökologische Kartierung (VHR)			
	MittlereNestgröße (m^2)	Bodenvolumen (m^3 pro m^2)	Anzahl Nester	max m^2
Brejo	**151,70**	**0,10**	4	224,42
Canyon	**104,21**	**0,12**	5	210,06
Casa	**146,78**	**0,10**	2	169,94

2.5 Detailbetrachtung Nestentwicklung 2018 bis 2021

Die Kartierungen von 2018 und 2021 ermöglichen einen direkten Vergleich der Nestentwicklung und werden in den kommenden Wochen ausgewertet. Im aufgeführten Beispiel (Abb. 5) zeigte sich anhand der Orthofotos, dass die Nestgröße bei einem älteren Nest sich zwischen 2018 und 2021 kaum weiterentwickelt hatte. Dafür hatte sich die Zahl kleinerer Öffnungen deutlich erhöht. Ebenso zeigten sich einige neue größere Öffnungen, die mutmaßlich von Räubern stammen können. Die Vegetationsbedeckung im südlichen Randbereich hatte deutlich zugenommen. Trotz dieser Anzeichen zeigten frische Auswürfe aus dem Jahr 2021, dass dieses Nest eine aktive Kolonie enthielt. Außerdem zeigten auffällig helle Stellen im Nordosten Hinweise auf neu aufgeworfenes Material. Durch die Vegetationsbedeckung könnte aber bereits mit den Orthofotos aus 2018 nicht eindeutig eine Aktivität in diesem Bereich detektiert werden. Diese Problematik wird sich nur durch einen höheren Anteil an Schrägaufnahmen oder den Einsatz von UAV-gestütztem LiDAR beheben lassen

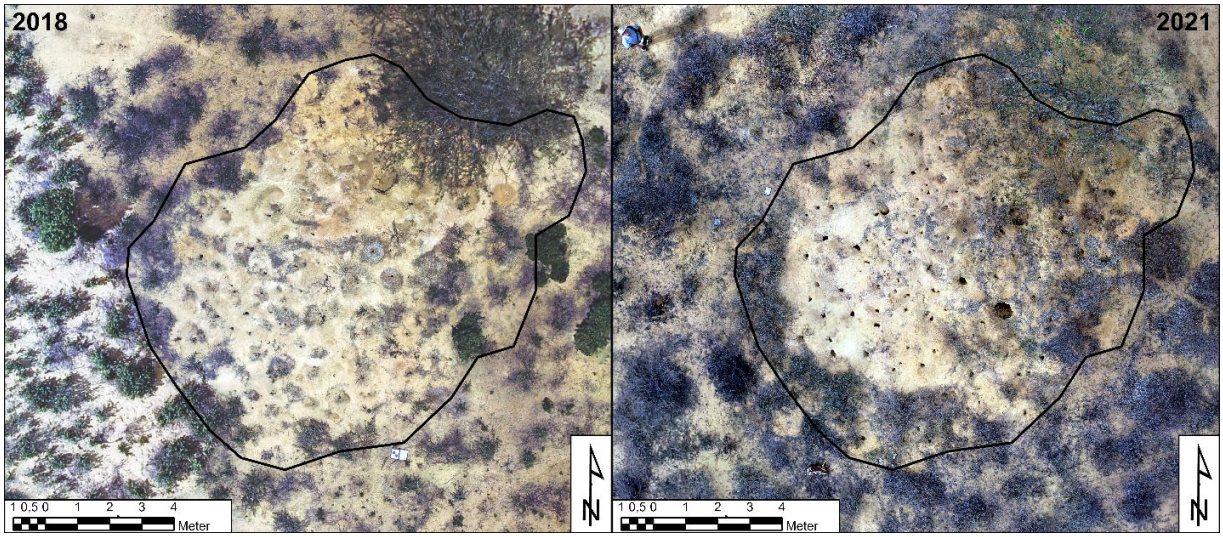


Abb. 5: Beispielhafte Darstellung eines Nests mit VHR Orthofotos aus den Jahren 2018 (links) und 2021 (rechts), kartierte Nestfläche (schwarz) von 2018

3 Fazit & Ausblick

UAV-gestützte Feldarbeit ermöglicht grundsätzlich eine einfache Quantifizierung der durch Blattschneiderameisen verursachten Bodenverlagerung an den Nestoberflächen. Außerdem wird eine Bestandsaufnahme der Nester, sowohl hinsichtlich ihrer Verbreitung und ihrer Größe als auch Ihrer Aktivität erreicht. Dies ist besonders relevant, da es kaum andere Möglichkeiten gibt, die Verteilung von Ameisennestern auf Landschaftsebene zu kartieren. Neben den methodischen Aspekten müssen bei der Kartierung von Ameisennestern auch biologische Faktoren berücksichtigt werden, insbesondere, wenn ökologische Kriterien zur Bestimmung der Grenzen dieser Strukturen herangezogen werden. Dabei muss natürlich in erster Linie berücksichtigt werden, dass es sich um Momentaufnahmen der Nester handelt und somit die ökologisch wichtige zeitliche Dynamik außer Acht gelassen werden muss.

Eine Ableitung der gesamten Bioturbationsleistung der Blattschneiderameisen ist grundsätzlich durch die Verbindung von Expertenkartierungen und detaillierten Volumenberechnungen auf der Landschaftsebene möglich. Für eine exakte Abschätzung der Bioturbationsprozesse über die gesamte Lebensdauer eines Ameisennestes müssen zukünftig wiederholte Messungen der Oberflächenstrukturen, der Erosionsrate und auch genauere Untersuchungen des tatsächlichen Nestvolumens unterhalb der Bodenoberfläche durchgeführt werden. Erst wird es möglich sein, die Auswirkungen dieser Organismen auf der Landschaftsebene zu beurteilen.

Ein Upscaling des durch Blattschneiderameisen verursachten edaphischen Ökosystem-Engineerings soll mit Hilfe fortgeschrittener Fernerkundungs-Klassifikationsmethoden auf hochaufgelösten Satellitendaten (WorldView 2) durchgeführt werden. Bei erfolgreicher Durchführung der Untersuchungen wäre es erstmals möglich, das Vorkommen aktiver Nester großflächig zu erfassen und deren Aktivitäten und Einfluss auf Ökosysteme zu quantifizieren.

4 Literaturverzeichnis

BIEBER, A. G. D., OLIVEIRA, M. A., WIRTH, R., TABARELLI, M. & LEAL, I. R., 2011: Do abandoned nests of leaf-cutting ants enhance plant recruitment in the Atlantic Forest? Austral Ecol., **36**, 220-232.

FINKLER, D., 2020: Ökosystem-Engineering von Blattschneiderameisen durch Bioturbation im Caatinga-Trockenwald (Brasilien). Diplomarbeit, Technische Universität Kaiserslautern, FB Biologie.

FOWLER, H.G. & CLAVER, S., 1991: Leaf- cutter ant assemblies: effects of latitude, vegetation, and behavior. Ant- Plant Interactions, Huxley, C. R., Cutler, D. F. (Hrsg.), Oxford University Press, 51-59.

FOWLER, H.G., PEREIRA DA SILVA, V., FORTI, L.C. & SAES, N.B., 1986: Population dynamics of leaf-cutting ants: a brief review. Fire ants and Leaf-cutting ants, Lofgren, C.S., Vander Meer, R.K. (Hrsg.), Boulder and London, Westview Press, 123-145.

LEAL, I. R., WIRTH, R. & TABARELLI, M., 2014: The multiple impacts of leaf- cutting ants and their novel ecological role in human- modified neotropical forests. Biotropica, **46**, 516-528.

MABOUDI, M., ELBILLEHY, A., GHASSOUN, Y. & GERKE, M., 2021: Improving the Internal Accuracy of UAV-Image Blocks Using Local Low-Altitude Flights and Scale-Bars, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B1-2021, 183-188, https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-183-2021.

MEYER, S. T., NEUBAUER, M., SAYER, E. J., LEAL, I. R., TABARELLI, M. & WIRTH, R., 2013: Leaf-cutting ants as ecosystem engineers: topsoil and litter perturbations around Atta cephalotes nests reduce nutrient availability. Ecological Entomology, **38**, 497-504.

MOREIRA, A., FORTI, L. C., ANADRADE, A. P., BOARETTO, M. A. & LOPES, J. F. S., 2004: Nest architecture of Atta laevigata (F. Smith 1858) (Hymenoptera: Formicidae). Stud. Neotrop. Fauna E., **39**, 109-116.

RÖMER, D. & ROCES, F., 2015: Available space, symbiotic fungus and colony brood influence excavation and lead to the adjustment of nest enlargement in leaf-cutting ants. Insectes Sociaux, **62**, 401-413.

SANZ-ABLANEDO, E., CHANDLER, J.H., RODRÍGUEZ-PÉREZ, J.R. & ORDÓÑEZ, C., 2018: Accuracy of Unmanned Aerial Vehicle (UAV) and SfM Photogrammetry Survey as a Function of the Number and Location of Ground Control Points Used. Remote Sens., **10**(10), 1606. https://doi.org/10.3390/rs10101606.

SILVA, J.M.C., LEAL, I.R. & TABARELLI, M. (Hrsg.), 2017: Caatinga: The largest tropical dry forest region in South America. Springer International Publishing.

SIQUEIRA, F. F., RIBEIRO-NETO, J. D., TABARELLI, M., Andersen, A. N., Wirth, R. & Leal, I. R., 2017: Leaf-cutting ant populations profit from human disturbances in tropical dry forest in Brazil. Journal of Tropical Ecology, **5**, 337-344.

WEBER, N. A., 1966: Fungus growing ants. Science, **153**, 587-604.

WIRTH, R., W. BEYSCHLAG, H. HERZ, R. J. RYEL, & HOELLDOBLER. B., 2003: Herbivory of leaf-cutter ants: a case study of Atta colombica in the tropical rainforest of Panama. Ecological Studies, **164**, Springer-Verlag, Berlin, Germany.

Multi-angular Multispectral Sensing of Barley Awn Coloring

MORITZ CAMENZIND[1], MICHAEL THALER[1], ULF FEUERSTEIN[2] & KANG YU[1]

Abstract: Barley awn coloring is a descriptive trait for barley (Hordeum vulgare) testing by the European Union. Rating the coloring visually is difficult and measuring canopy reflectance by a multispectral canopy might be more accurate. Reflectance of barley canopies with different awn coloring levels was measured from different viewing angles using a multispectral camera mounted to an unmanned aerial vehicle. The bi-directional reflectance distribution of the canopies was determined and used for deriving the anthocyanin reflection index, which was then used to identify the optimal viewing angles for the detection of Barley awn coloring. The index differed between the canopies without awn coloring and the canopies with awn coloring. The proposed method failed to detect subtle changes of awn coloring but it might be further improved and used to characterize canopy traits from multiple viewing angles.

1 Introduction

Barley awn coloring is a result of the accumulation of color pigments such as anthocyanins in the barley awn tips during flowering. It is a descriptive trait for barley (*Hordeum vulgare*) plants during the testing process for distinctness, uniformity and stability by the community plant variety office (CPVO) of the European Union. The coloring is usually rated visually, which is difficult because barley awn coloring is very subtle, is only present during a short time period and its perception depends on the viewing geometry of the examiner. Barley awn coloring might be measured with a higher consistency and accuracy by multispectral reflectance measurements of the canopy. Experience from visually rating barley awn coloring suggests that a viewing zenith angle that is higher than nadir might be beneficial for the detection of barley awn coloring with multispectral cameras as well.

Multispectral reflectance of canopies is widely used in agriculture remote sensing to estimate e.g. nitrogen use efficiency (ARGENTO et al. 2021) or green fraction (JAY et al. 2017). Often, canopy reflectance is measured from nadir by using a Structure from Motion (SfM) algorithm, which stiches pixels of single images into an orthomosaic. However, canopy reflectance is highly dependent on the geometry of observation (SANDMEIER & ITTEN 1999) and can be described by a bi-directional reflectance distribution function (NICODEMUS et al. 1977). Models accounting for the effects of bi-directional reflectance distribution (BRD) have been used to observe the leaf area index (LAI) and chlorophyll content (ROOSJEN et al. 2018) as well as reproductive organs (LI et al. 2021).

The data collection as well as the data processing have to be adjusted from measurement procedures used for the collection of nadir observations when measuring BRD using a multispectral camera mounted to an unmanned aerial system (UAS). LI et al. (2021) adjusted the flight path, which limited the area that can be covered by a single flight to few hundred square meters. ROOSJEN et al. (2018) down sampled the image resolution to a ground sampling distance of 5 m to calculate viewing angles for single pixels. Both approaches are limited at measuring BRD

[1] Technical University of Munich, Chair of Precision Agriculture, Dürnast 9, D-85354 Freising
E-Mail: [moritz.camenzind, michael.david.thaler, kang.yu]@tum.de
[2] Deutsche Saatveredelung AG, Steimker Weg 7, D-27330 Asendorf
E-Mail: ulf.feuerstein@dsv-saaten.de

for trials with several hundred plots of a size of few square meters as can be found in breeding trials.
A workflow to extract LAI from multi-view RGB images developed by ROTH et al. (2018) was adjusted to cope with multispectral images in order to measure canopy BRD of barley canopies during the appearance of awn coloring. The BRD was then used to identify the optimal viewing configuration to detect barley awn coloring by canopy reflectance measurements.

2 Materials and Methods

All measurements were taken at the breeding station of the Deutsche Saatveredelung AG (DSV) in Asendorf, Lower Saxony, Germany (52.777° N, 8.682° E, 50 m a.s.l.). A total number of 150 Barley (*Hordeum vulgare*) genotypes were drill sown in plots of 1.5 m × 5 m size, which were placed randomly in a grid on the trial site. Fertilization and plant protection were carried out according to local practice. The anthocyanin coloring of the awns was rated visually according to a protocol used by the European Union for variety testing (COMMUNITY PLANT VARIETY OFFICE 2019; Tab. 1). The BBCH stage of the canopies was 60 to 65 during measurements and the expression of the barley awn coloring was maximal.

Tab. 1: Definitions of the rating scores.

Score	Definition
1	no awn coloring
3	single awns were colored but no homogenous coloring
5	light awn coloring
7	intermediate awn coloring
9	strong awn coloring

The multispectral images were taken using a MicaSense Altum camera (MicaSense Inc., Seattle, USA), which has five bands (475 nm, 560 nm, 668 nm, 717 nm and 842 nm) and a field of view of 48° × 37°. The camera was mounted on a Matrice 600 Pro rotary wing unmanned aerial system (Dà-Jiāng Innovations Science and Technology Company, Shenzhen, China). During flight, the camera was oriented in nadir and the width of the sensor was aligned to flight direction, which was east-east-south to west-west-north and the opposite direction. Flight height was 37 m above ground level, resulting in a ground sampling distance (GSD) of 16 mm. Flight speed was 1.6 m s^{-1}, front and side overlap were 90%. The flight took place under stable and sunny weather conditions on May 31, 2021 at 12:48 local time (UTC + 1) and lasted 11 min.
Goal of the image processing was to obtain reflectance values from single plots defined by ROIs on single images (Fig. 1). As a first step, the images were calibrated geometrically as well as radiometrically. The geometric calibration included distortion correction, the radiometric calibration included corrections for vignette, black level, exposure and gain. Radiance was converted into reflectance by measuring the radiance of a panel with a known reflectance before and after the flight. Calibration values during the flight were linearly interpolated. Finally, the images of single bands were co-registered. The mentioned steps were all carried out using python code (Python Software Foundation, https://www.python.org/), which was provided by the camera manufacturer (MICASENSE, 2021).

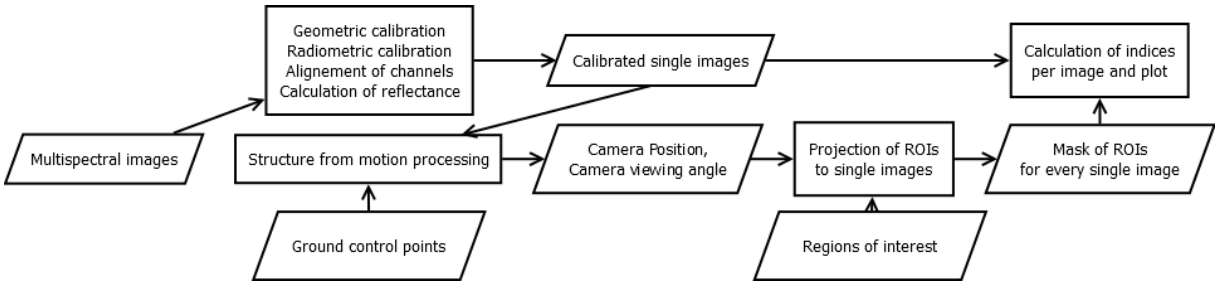


Fig 1: Diagram of the applied Workflow

After calibration, the images were imported into Agisoft Metashape Professional 1.7.2 (Agisoft LLC, St. Petersburg, Russia) and processed as described by ROTH et al. (2018). The coordination system used was ESPG: 32632 (WGS 84 / UTM zone 32N), size of the ROIs, containing information about the location of the plots, was 1.1 m × 4.4 m. After processing, a geojson file was exported containing information about the position of the plots in each single image, as well as on the viewing geometry, such as the viewing azimuth and the viewing zenith angle.
In order to detect barley awn coloring the anthocyanin reflection index (ARI) was calculated on a pixel basis according to GITELSON et al. (2001):

$$ARI = \frac{1}{R_{560}} - \frac{1}{R_{717}}$$

where R is the reflectance of the respective wavelength. A median was calculated per plot, image, band or index, including all the pixels within the ROI.
Further data analysis and plotting was done in R 4.0.5 (R CORE TEAM, 2021). To produce a visual representation of the BRD, the reflection measurements of single plots were interpolated using inverse distance weighing. Only measurements that were taken from azimuth angles relative to the sun of – 10° to 10° and 170° to 190° and zenith angles of more than 11° were considered for comparison of the rating with the reflectance measurement. Within these regions, the ARI was not dependent on the viewing angles. Analysis of variance and the Tukey post hoc test were carried out on a significance level of 5 %.

3 Results

The visual rating of the barley awn coloring found 77 plots, which showed no awn coloring, 15 plots where single awns were colored, 21 plots with light awn coloring, 31 plots with intermediate awn coloring and 5 plots with high awn coloring. Imaging yielded 304 image stacks, on each stack we located 30 plots on average (sd = 15). We therefore obtained 9084 single measurements per band or index, resulting in an average of 60 measurements per plot (sd = 9). Plots located in the middle of the trial were measured more often than those at the border. Viewing azimuth angles ranged from 0° to 360°, most measurements were made perpendicular to flight direction. Viewing zenith angles range from 0° to 28°, most measurements were taken at a viewing zenith angle of 12°.
A lower reflectance in the green band was measured for plots with a rated awn coloring of 1 (0.032, cv = 27.6) compared to plots with a rated awn coloring of 9 (0.029, cv = 26.1). Reflectance in the rededge band was similar for both awn color ratings (0.089, cv = 18.7 and 0.090, cv = 17.0). The ARI was lower for plots with an awn color rating of 1 (21.4, cv = 29.9) compared to a rating of 9 (24.8, cv = 28.8).

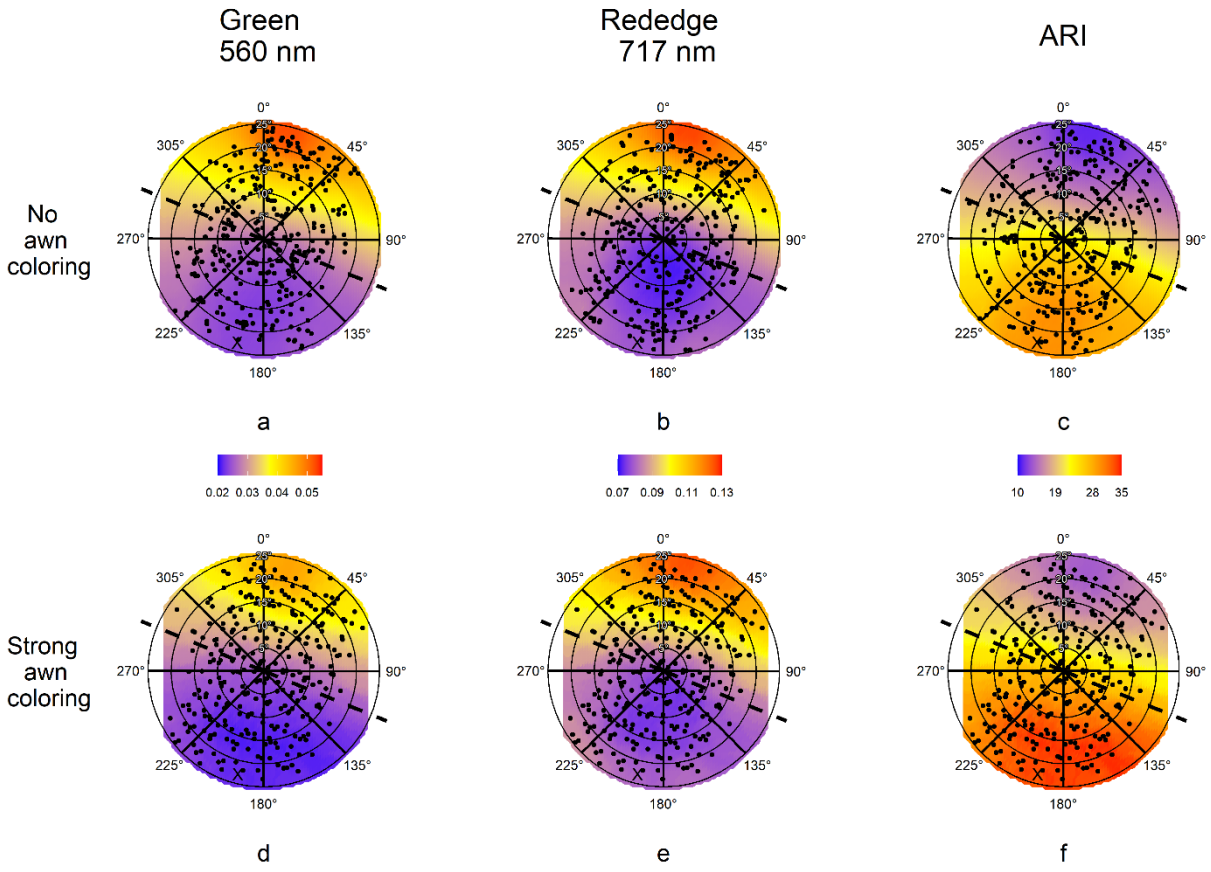


Fig. 2: Representation of the measured BRD from all plots with no (Rating 1, top) and high awn coloring (Rating 9, bottom). (a, c) represent the green, (b, c) the rededge and (c, f) the ARI BRD. The x-axis displays the viewing azimuth angle (0 ° corresponds to north, 90 ° to east), the y-axis corresponds to the viewing zenith angle. The black cross represents the sun azimuth position, the black dashed line the direction of flight. The black dot represent locations of measurements of which only 300 measurements are shown per subfigure to make the figure more readable. Images were taken at 12:48 local time (UTC + 1).

The reflectance in both bands decreased with the viewing azimuth angle relative to the sun (Fig. 2). The reflection in the rededge band was further negatively correlated with the zenith viewing angle (Fig. 2b). The same effect was observed in the near-infrared band, where the effect was more distinct (data not shown). The ARI increased with a lower viewing azimuth angle relative to the sun (Fig. 2cd) and was maximal when the angle is 0 ° and minimal for a viewing azimuth angle relative to the sun of 180 °. A positive relationship between the viewing zenith angle and the ARI was found when filtering for viewing azimuth angles relative to the sun of 350 ° to 10 ° and 170 ° to 190 °. The variability in the ARI however increased with the viewing zenith angle (Fig. 3c).

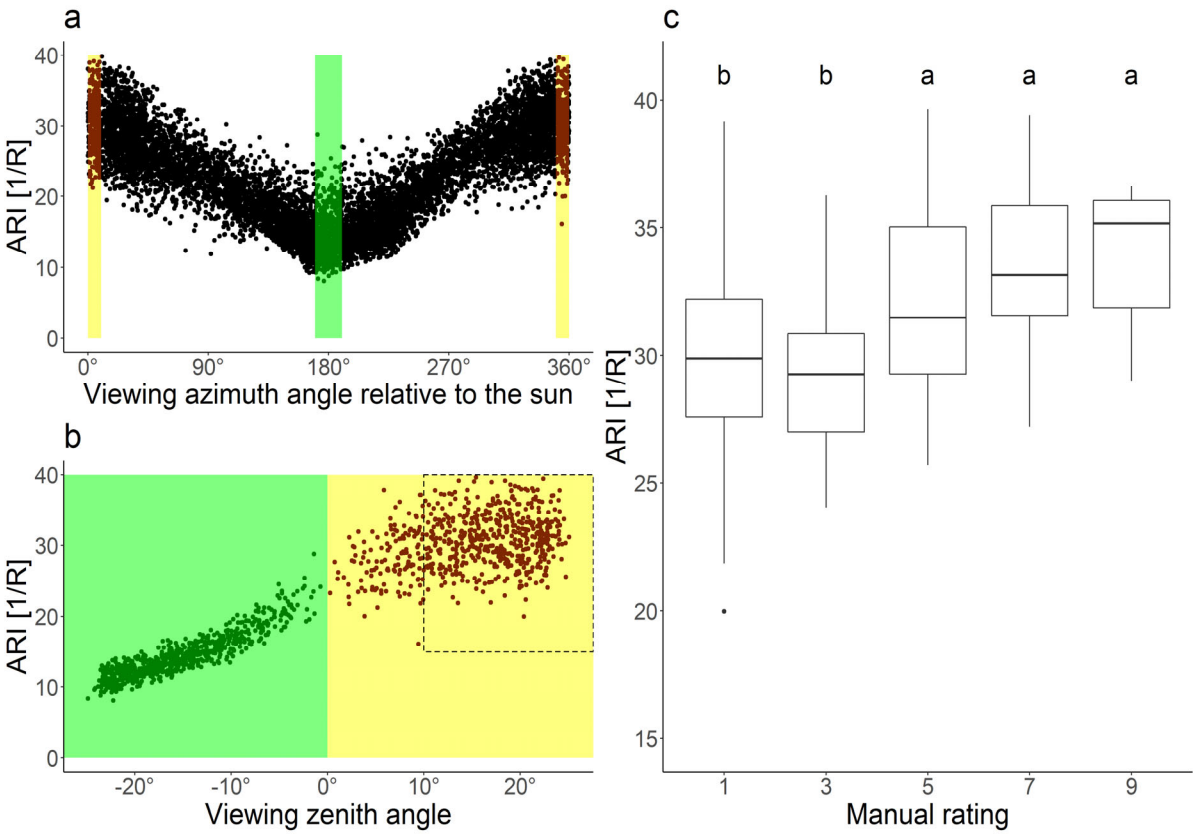


Fig. 3: Influence of viewing angles on the ARI (a, b) as well as the ARI of different rating groups (c). (a) shows the influence of the viewing azimuth angle relative to the sun, where 0 ° is the forward and 180 ° is the backward scattering direction. The colored areas display the region from which the values were filtered for further analysis. (b) shows the influence of the viewing zenith angle at the mentioned viewing azimuth angles. Negative angles correspond to zenith angles in the backward scattering direction, positive angles to the forward scattering direction. Colored regions correspond to (a), the dashed line box includes values with a viewing zenith angle of 11 ° and more. (c) displays the ARI for measurements filtered by the mentioned viewing azimuth and viewing zenith angles grouped by their rating score. The letters correspond to the statistical group of the Tukey comparisons.

After filtering for the aforementioned sun azimuth angles and viewing zenith angles of more than 11 °, significant differences were found in the ARI between plots where no or only single awns were colored and plots where the awns were colored (Fig. 3c). For viewing zenith angles of more than 0 ° and less than 11 ° no significant differences among the rating groups in the ARI were found (Data not shown).

4 Discussion and Outlook

Adapting the workflow developed by ROTH et al. (2018) to multispectral imaging was successful. The importance of choosing the same viewing geometry when comparing reflectance measurements is underlined by the relatively big coefficient of variation (cv) for measurements of a single rating group including all viewing angles. Differences in the mean of the ARI between rating groups are much smaller than within the group.

Tab. 2: Statistical parameters of ARI measurements per rating group

Rating	1	3	5	7	9
n	220	62	65	119	12
ARI mean	29.7	29.0	32.2	33.4	33.9
ARI cv	11.2	9.3	311.6	8.4	7.7

Generally, the reflectance decreased with a lower viewing angle relative to the sun and increased with increasing viewing angle (Fig. 2). The highest reflectance was measured with a viewing angle relative to the sun of around 180°. The measured BRD therefore stands in opposition to other studies, which found an increasing reflectance in the single bands with a lower azimuth angle relative to the sun (BURKART et al. 2015; LI et al. 2021). One reason for the unexpected BRD might be due to the image reflectance calibration. We initially planned to calibrate the images using an incident light sensor (ILS) mounted to the UAS. However, the irradiation measured by the ILS was highly affected by the yaw angle of the UAV and therefore not suitable for reflectance calibration on single images. Using reflectance panels on the ground is a well-known approach for the radiometric calibration of multispectral images but was limited to one calibration image before and one calibration image after the flight. A calibration with the method used, requires perfectly stable light conditions, which are rare in Lower Saxony. For future flight campaigns, reflectance panels should be put within the trial site to calibrate each image individually. These panels will further allow to calculate a bi-directional reflectance distribution function (BRDF) of the canopies.
Further, the BRD of canopies with colored barley awns might be different than the BRD of canopies measured in other studies. LI et al. (2021) measured the BRD of wheat at a similar BBCH stage, when awns were present in the canopy as well but the awns were not colored. We might therefore review our findings in a next study, measuring the BRD using established techniques for BRD measurements such as a hand held hyperspectrometer.
The optimal viewing angle range should be as narrow as possible to exclude viewing angle effects within the selected range. However, the range needs to be big enough in order to obtain enough reflectance measurements. We found that the ARI measurements stay stable in relation to the viewing azimuth angle in ranges with a size 20° centered on a sun azimuth angle of 0° and on a sun azimuth angle of 180° (Fig. 3). The ARI in these viewing azimuth angle ranges is highly correlated to the viewing zenith angle. Therefore, as for the viewing zenith angle, a range has to be chosen to obtain comparable measurements. We filtered the dataset for a viewing zenith angle of more than 11°, since the values in this range do not change significantly with the viewing zenith angle. Using this viewing configuration yielded significant differences in the ARI between the plots with non-colored and the plots with colored awns. Still, the variability in the ARI within one level of rated awn coloring is very high and an ARI measurement cannot clearly be assigned to a rating group, making it difficult to use the procedure in practice (Tab. 2). The detection might be improved by a higher viewing zenith angle. Currently, the viewing zenith angles are limited by the field of view of the camera and have not been high enough to measure the full BRD and observing phenomena such as the hot spot effect (KUUSK 1991). This shortcoming might be solved by tilting the camera on the UAV or choosing a flight time when the sun zenith angle is higher. Another improvement to the workflow would be the use of smaller ROIs. The used ROIs have almost the same size as the plot and the viewing angles of single pixels within one ROI vary by several degrees. Using smaller ROIs would

allow for smaller variation in the viewing angles within the ROI and therefore allow choosing the viewing angles more precisely. Smaller ROIs would however be less representative for the whole plot.
Using the proposed method for multi-angular multispectral measurements for canopy reflectance measurements shows promising first results. The method might be further improved and used to characterize canopy traits from multiple viewing angles. However, compared to classical reflectance measurements, the proposed method is more complex and requires optimal measuring conditions as well as further improvements.

5 Acknowledgements

We thank the Deutsche Saatveredelung AG (DSV) for providing access to their breeding trial as well as to their infrastructure for data collection.

6 Literature

ARGENTO, F., ANKEN, T., ABT, F., VOGELSANGER, E., WALTER, A. & LIEBISCH, F., 2021: Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. Precision Agriculture, **22**(2), 364-386, https://doi.org/10.1007/s11119-020-09733-3.

BURKART, A., AASEN, H., ALONSO, L., MENZ, G., BARETH, G. & RASCHER, U., 2015: Angular Dependency of Hyperspectral Measurements over Wheat Characterized by a Novel UAV Based Goniometer. Remote Sensing, 7(1), 725-746, https://doi.org/10.3390/rs70100725.

COMMUNITY PLANT VARIETY OFFICE, 2019: Protocol for distinctness, uniformity and stability tests. Angers, France, https://cpvo.europa.eu/sites/default/files/documents/hordeum_5_0.pdf.

GITELSON, A. A., MERZLYAK, M. N. & CHIVKUNOVA, O. B., 2001: Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochem Photobiol, **74**(1), 38-45, https://www.doi.org/10.1562/0031-8655(2001)074%3C0038:OPANEO%3E2.0.CO;2.

JAY, S., MAUPAS, F., BENDOULA, R. & GORRETTA, N., 2017: Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. Field Crops Research, **210**, 33-46, https://doi.org/10.1016/j.fcr.2017.05.005.

KUUSK, A., 1991: The Hot Spot Effect in Plant Canopy Reflectance. In R. B. Myneni & J. Ross (Eds.), Photon-Vegetation Interactions: Applications in Optical Remote Sensing and Plant Ecology, 139-159, Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-642-75389-3_5.

LI, W., JIANG, J., WEISS, M., MADEC, S., TISON, F., PHILIPPE, B., COMAR, A. & BARET, F., 2021: Impact of the reproductive organs on crop BRDF as observed from a UAV. Remote Sensing of Environment, **259**, 112433, https://doi.org/10.1016/j.rse.2021.112433.

MICASENSE, 2021: MicaSense RedEdge and Altum Image Processing Tutorials. Micasense Inc., https://github.com/micasense/imageprocessing.

NICODEMUS, F.E., KICLUNOND, J.C., HSIA, J.J., GINSBERG, I.W., LIMPERIS, T., 1977: Geometrical Considerations and Nomenclature for Reflectance (no. NBS MONO 160). National Bureau of Standards, Gaithersburg, MD, https://doi.org/10.6028/NBS.MON0.160.

R CORE TEAM, 2021: R: A Language and Environment for Statistical Computing. R foundation for statistical computing, Vienna, Austria, https://www.R-project.org/.

ROOSJEN, P. P. J., BREDE, B., SUOMALAINEN, J. M., BARTHOLOMEUS, H. M., KOOISTRA, L. & CLEVERS, J. G. P. W., 2018: Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data – potential of unmanned aerial vehicle imagery. International Journal of Applied Earth Observation and Geoinformation, **66**, 14-26, https://doi.org/10.1016/j.jag.2017.10.012.

ROTH, L., AASEN, H., WALTER, A. & LIEBISCH, F., 2018: Extracting leaf area index using viewing geometry effects—A new perspective on high-resolution unmanned aerial system photography. ISPRS Journal of Photogrammetry and Remote Sensing, **141**, 161-175, https://doi.org/10.1016/j.isprsjprs.2018.04.012.

SANDMEIER, S. R. & ITTEN, K. I., 1999: A field goniometer system (FIGOS) for acquisition of hyperspectral BRDF data. IEEE Transactions on Geoscience and Remote Sensing, **37**(2), 978-986, https://doi.org/10.1109/36.752216.

Detektion von Fahrradständern in Luftbildern mittels Deep Learning

KARIN ERBE[1], MELANIE BRANDMEIER[2], MICHAEL SCHMITT[3], ANDREAS DONAUBAUER[1], JAN-ANDREAS LIEBSCHER[4] & THOMAS H. KOLBE[1]

Zusammenfassung: Dieser Beitrag erörtert anhand der Detektion und Klassifikation von unüberdachten Fahrradständern aus Luftbildern und Oberflächenmodellen, inwiefern Deep Learning eingesetzt werden kann, um Stadtmöbel automatisch zu erfassen und für 3D-Stadtmodelle nutzbar zu machen. Dafür wird am Beispiel der Fahrradständer der Landeshauptstadt München ein Trainingsdatensatz erzeugt und ein Mask R-CNN Modell trainiert. Um die Ergebnisse zu evaluieren, wird eine umfangreiche quantitative und qualitative Fehleranalyse durchgeführt. Dabei zeigte sich, dass bei der Verwendung der RGB-Bilddaten in Verbindung mit dem nDOM-Raster die besten Detektionsergebnisse erzielt werden. Die Detektion erreicht eine Genauigkeit (Precision) von 83,2% und eine Trefferquote (Recall) von 77,3%. Als Einschränkungen erweisen sich der Schattenwurf durch Gebäude und Bäume. Zudem wird deutlich, dass die Zahl der abgestellten Fahrräder einen signifikanten Einfluss auf den Erfolg der Detektion hat.

1 Motivation

3D-Stadtmodelle spielen eine wichtige Rolle bei Planungs- und Verwaltungsaufgaben. Durch fortwährende Veränderungen im Stadtbild ergibt sich der Bedarf nach einer regelmäßigen Aktualisierung der 3D-Stadtmodelle. Insbesondere Digitale Zwillinge von Städten, die versuchen die physikalische Umwelt möglichst exakt abzubilden, benötigen eine zeitnahe Synchronisation der abgebildeten Objekte. Während für Gebäude, beispielsweise mit Hilfe von Laserscanning-Befliegungen, dem kontinuierlich aktualisierten Liegenschaftskataster und darauf aufsetzender Gebäuderekonstruktion bewährte Fortführungsprozesse genutzt werden, hat sich bei der automatisierten Erfassung von Stadtmöblierung noch keine geeignete Methode durchgesetzt. Deep Learning bietet sich in diesem Feld an, da es eine effiziente Möglichkeit bietet, aus Bilddaten automatisiert Objekte zu erfassen. Deep Learning ist ein Teilbereich des maschinellen Lernens, welcher die Lehre von tiefen künstlichen neuronalen Netzen umfasst. Einmal trainiert kann ein Deep Learning Netzwerk auf ähnliche, aber neue Bilder angewendet werden. Inwiefern Deep Learning eingesetzt werden kann, um Stadtmöbel automatisch zu erfassen und für 3D-Stadtmodelle nutzbar zu machen, wird in diesem Beitrag anhand der Detektion und Klassifikation von unüberdachten Fahrradständern aus Luftbildern und Oberflächenmodellen

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: [karin.erbe, andreas.donaubauer, thomas.kolbe]@tum.de
[2] Hochschule Würzburg-Schweinfurt, Röntgenring 8, D-97070 Würzburg, E-Mail: melanie.brandmeier@fhws.de
[3] Universität der Bundeswehr München, Institut für Raumfahrttechnik und Weltraumnutzung, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, E-Mail: michael.schmitt@unibw.de
[4] Landeshauptstadt München - Kommunalreferat - Geodatenservice, Denisstraße 2, D-80335 München, E-Mail: jan.liebscher@muenchen.de

erörtert. Für den Aufbau des Deep Learning Frameworks und das Training werden dabei die Deep Learning Funktionalitäten von ArcGIS Pro (ESRI 2022A) und der ArcGIS API für Python (ESRI 2022B) verwendet.

2 Verwandte Arbeiten

In der Literatur finden sich bisher keine einschlägigen Vorarbeiten zur automatischen Erfassung von Fahrradständern und Straßenmöbeln ähnlicher Art, wie Bänke und Mülleimer, mittels Deep Learning aus Luftbildern. Veröffentlichungen über die Erkennung von Fahrzeugen oder Flugzeugen (z.B. CHEN 2018) sind dagegen verbreiteter. Die Fernerkundung stellt besondere Anforderungen an die Objekterkennung und die Instanzsegmentierung. Dies liegt an der besonderen Beschaffenheit der Daten, die aus der Vogelperspektive und nicht wie die meisten Bilddaten im Profil erfasst werden. Daraus ergeben sich 360°-Objektdrehungen, komplexe Hintergründe (LI et al. 2018), unterschiedliche Aufnahmebedingungen wie Wetter, Jahreszeit und Auflösung (LI et al. 2020), aber auch Verdeckungen, Beleuchtungsänderungen und Schattenwurf (SU et al. 2019). Darüber hinaus können Fernerkundungsdaten aus mehr als den drei üblichen Farbbändern Rot, Grün und Blau bestehen.
Ein Beispiel für ein 3D-Stadtmodell, in dem Stadtmobiliar enthalten ist, wurde für die Stadt Rotterdam implementiert (GEMEENTE ROTTERDAM 2022). Es enthält beispielsweise Laternen, Mülleimer, Parkautomaten und auch Fahrradständer. Ob es möglich ist Objekte dieser Art automatisch in Luftbildern mit Hilfe von Deep Learning zu detektieren, wird im Folgenden anhand von Fahrradständern in München erörtert.

3 Deep Learning-basierte Erkennung von Fahrradständern

3.1 Anwendungsgebiet & Daten

Als Untersuchungsgebiet wird die Landeshauptstadt München gewählt, die an etwa 350 Orten mehr als 34.000 öffentliche Fahrradstellplätze bereitstellt (Stand: 2019). Eine Reihe von Fahrradstellplätzen, die baulich als zusammenhängend erkennbar sind, kann als ein Fahrradständer-Objekt bezeichnet werden. Die Landeshauptstadt München führt diese Fahrradständer-Objekte als linienhafte Geoobjekte in einem Geodatensatz, welcher als Grundlage für die Erstellung der notwendigen Trainingsdaten für den Deep Learning Ansatz dient. Jedem der Fahrradständer ist dabei ein Fahrradständer-Typ zugewiesen. Dabei handelt es sich beispielsweise um Anlehnbügel, Klemmbügel oder verschiedene Typen von Vorderradhaltern, die ein Anschließen des Rahmens ermöglichen.
Als Bildgrundlage dienen RGB-Orthophotos aus einer Winterbefliegung im Jahr 2019 mit einer Bodenauflösung von 8 cm. Zusätzliche Rasterdaten werden hinzugezogen, um zu testen, inwiefern CIR (color-infrared), nDOM (normalisiertes digitales Oberflächenmodell) und NDVI (Normalized Difference Vegetation Index) die Ergebnisse verbessern können. Dabei werden folgende fünf Datenkombinationen getestet: RGB, CIR, RGB+nDOM, RGB+NDVI bzw. RGB+nDOM+NDVI.

3.2 Erstellung der Trainingsdaten

Die Erstellung der Trainingsdaten erfolgt mit dem ArcGIS-Pro-Werkzeug *Beschriften von Objekten für Deep Learning*. Dafür werden alle Linienobjekte des Fahrradständer-Datensatzes einzeln betrachtet. Jeder im Luftbild erkennbare Fahrradständer wird entsprechend seines Typs manuell mit einem Rechteck eingerahmt. Von 1.999 Fahrradständer-Objekten im Geodatensatz sind 1.463 sichtbar und als Trainingsobjekte nutzbar. Die übrigen 536 Objekte können aufgrund der Überdachung, der Verdeckung durch Bäume oder aufgrund von sehr dunklem Schatten nicht verwendet werden. Darüber hinaus enthält der Datensatz Fahrradständer, die zum Zeitpunkt der Luftbildaufnahme nicht mehr oder noch nicht aufgestellt waren. Einige der Objekte im Datensatz werden durch mehrere Rechtecke repräsentiert, sodass ein Trainingsdatensatz entsteht, der mit ca. 1.800 markierten Fahrradständer-Objekten als Grundlage für das Deep Learning dient.
Bei der Erstellung der Trainingsdatensätzen werden bereits Herausforderungen dieses Deep-Learning-Ansatzes deutlich (siehe Abb. 1): Schattenwurf durch Gebäude und Bäume erschweren das Erkennen von Fahrradständern ebenso wie Äste und leere Stellplätze. Hinzu kommt, dass einige der Linienobjekte verschoben oder mit falscher Länge hinterlegt sind und somit den Markierungsprozess (*Labeling*) erschweren.

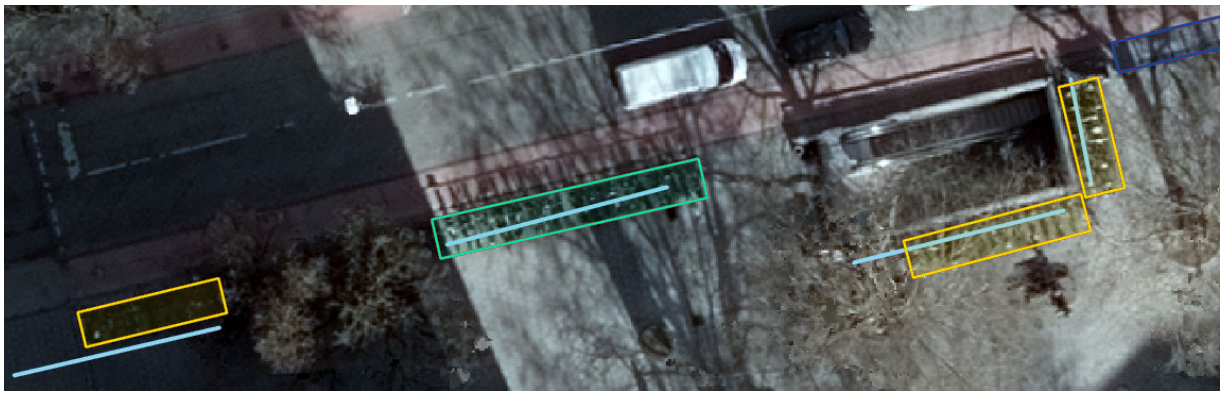


Abb. 1: Ausschnitt aus RGB-Orthophoto (Rotkreuzplatz) mit markierten Fahrradständern unterschiedlichen Typs; links: verschobene Liniengeometrie (Hellblau) und dunkler Schatten durch Gebäude, der ein Erkennen des Objekts erschwert; Mitte: Fahrradständer im Schatten eines Baumes; rechts: Äste verdecken Teile des Fahrradständers (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

Anschließend werden die Bildkacheln mit den zugehörigen Labeln für das Training exportiert. Dazu werden im selben ArcGIS-Pro-Werkzeug TIFF-Dateien der Größe 256×256 Pixel exportiert. Zu jedem Bild wird eine binäre Maske der gleichen Größe ausgegeben, welche die Position der Fahrradständer einer Klasse ausgibt. Für die gesamte Fläche Münchens ergeben sich damit ca. 7500 Bildkacheln.

3.3 Training

ArcGIS Pro stellt für die Detektion von Objekten verschiedene Deep-Learning-Frameworks zur Verfügung. Einer der Netzentwürfe, Mask R-CNN (HE et al. 2017), bietet über die Objekterkennung hinaus zusätzlich eine Instanzsegmentierung, sodass nicht nur das Objekt durch eine rechteckige Maske markiert, sondern pixelgenau lokalisiert wird. Dabei handelt es sich um die Weiterentwicklung der R-CNN Familie: Convolutional Neural Networks, die zusätzlich Region Proposal Networks implementieren, d.h. Layer, die in einem zusätzlichen Schritt dafür zuständig sind, Regionen von Interesse zu identifizieren. Mask R-CNN implementiert zwei Head Netzwerke, die parallel arbeiten: eines für die Klassifikation und die Bounding Box Regression und das andere für die exakte Berechnung der Segmentierungsmaske. Die Verlustfunktion ergibt sich dabei aus der Summe der Verlustfunktionen für Klassifikation (log loss), Bounding Box Regression (L_1 loss) und Maskenberechnung (average binary cross-entropy loss) (HE et al. 2017 & GIRSHICK 2015).

Für die Detektion der Fahrradständer-Objekte wurde ein Mask R-CNN Deep-Learning-Modell mit vortrainiertem ResNet-50 Backbone mit Hilfe der ArcGIS API für Python über 100 Epochen trainiert.

Als ein wichtiger Teil der Preprocessing-Aufgaben stellt sich die Data Augmentation, also die Erweiterung der Daten, heraus. Die Bilder im Trainingsdatensatz werden durch Drehungen, Helligkeits- und Kontraständerungen, sowie Zoomen verändert und verwendet, um den Trainingsdatensatz künstlich zu vergrößern (FAST.AI 2021). In diesem Projekt wurden 50% der Bilder um einen Winkel zwischen -30° und 30° gedreht. Diese Werte werden gewählt, um den Schattenwurf durch Bäume und Gebäude weiterhin realistisch zwischen West und Ost zu erhalten. Zudem werden 50% der Bilder horizontal an der vertikalen Bildmittelachse gespiegelt. Der Kontrast wird zufällig mit Werten zwischen 1 (gleichbleibend) und 1,5 erhöht und die Helligkeit leicht verändert (Werte zwischen 0,4 und 0,6 (0,5 ist gleichbleibend)). Zudem wird jedes Bild zufällig auf Zoomwerte zwischen 100% und 120% vergrößert. Mit Hilfe dieses Vorgangs wird das Training verbessert.

3.4 Nachbereitung

Die Anwendung der trainierten Modelle erfolgt anschließend in ArcGIS Pro mit dem Werkzeug *Objekte mit Deep Learning erkennen*. Detektierte Objekte werden durch die Segmentierungsmaske eingerahmt und durch diese in Position und Ausdehnung markiert. Im Rahmen des Projekts wurde zudem eine entsprechende Toolbox realisiert, die es Sachbearbeitern des 3D-Stadtmodells ermöglicht, auf Basis der trainierten Netze die Erkennung auf einfache und wiederholte Weise durchzuführen. Die Toolbox berechnet aus den Segmentierungsmasken (Abb. 2: Lila) Rechtecke (Abb. 2: Orange) und Linien (Abb. 2: Grün), welche nun für den gewünschten Zweck weiterverarbeitet werden können.

Abb. 2: Ausschnitt aus RGB-Orthophoto (Weißenburger Straße) mit gelabelten Fahrradständern in Rot, detektierten Fahrradständern in Lila, daraus berechneten Rechtecken in Orange und Linien in Grün (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

4 Evaluation der Ergebnisse

4.1 Vorgehen

Für die Evaluation des Trainingserfolgs müssen Bereiche verwendet werden, die einerseits nicht in den Trainingsdaten enthalten sind, andererseits aber Ground-Truth-Daten aufweisen. Dafür werden vier Gebiete aus dem Trainingsdatensatz entfernt (Abb. 3): die Gegenden um die S- und U-Bahn-Station Harras, um den Rotkreuzplatz, die U-Bahn-Station Messestadt West und das Gebiet zwischen Rosenheimer Platz und Ostbahnhof. Diese werden gewählt, da diese unterschiedliche Stadtstrukturen repräsentieren. Während Harras und Rotkreuzplatz belebte Plätze mit unterschiedlicher Baumdichte in der Stadt vertreten, liegt die Messestadt weiter außerhalb und ist mit vergleichsweise neuen Fahrradständern ausgestattet. Die Gegend nahe dem Ostbahnhof ist hingegen durch mehrstöckige Gebäude und damit viel Schatten und Hinterhöfe gekennzeichnet. Insgesamt liegen in diesen Stadtgebieten 128 gelabelte Fahrradständer.

Für jede der fünf Datenkombinationen werden alle detektierten Objekte einzeln betrachtet und in richtig positiv (RP) bzw. falsch positiv (FP) eingeordnet. Bei richtig positiven Detektionen wird dabei zusätzlich unterschieden, ob es sich um Fahrradständer handelt, welche auf Privatgrund stehen und daher nicht in den Ground-Truth-Daten enthalten sind ($RP_{sonstige}$), oder ob es Fahrradständer sind, die im kommunalen Datensatz enthalten sind (RP). Zweitere werden ferner darin unterschieden, ob sie korrekt oder falsch klassifiziert werden. Darüber hinaus fallen in die Gruppe der $RP_{sonstige}$ einzeln detektierte Fahrräder und öffentliche Fahrradständer, die fälschlicherweise nicht im Geodatensatz der Stadt München enthalten sind, aber korrekt detektiert werden. Falsch positive Detektionen werden anhand des fälschlich markierten Objektes gruppiert. Beispiele hierfür sind Autos, Dächer oder Vegetation. Umgekehrt werden auch die Ground-Truth-Daten betrachtet und alle nicht detektierten, also falsch negativen (FN), Objekte gezählt. Daraus ergibt sich nun folgende Tabelle (Tab. 1), die einen Vergleich der Datenkombinationen für die gewählten Stadtgebiete ermöglicht.

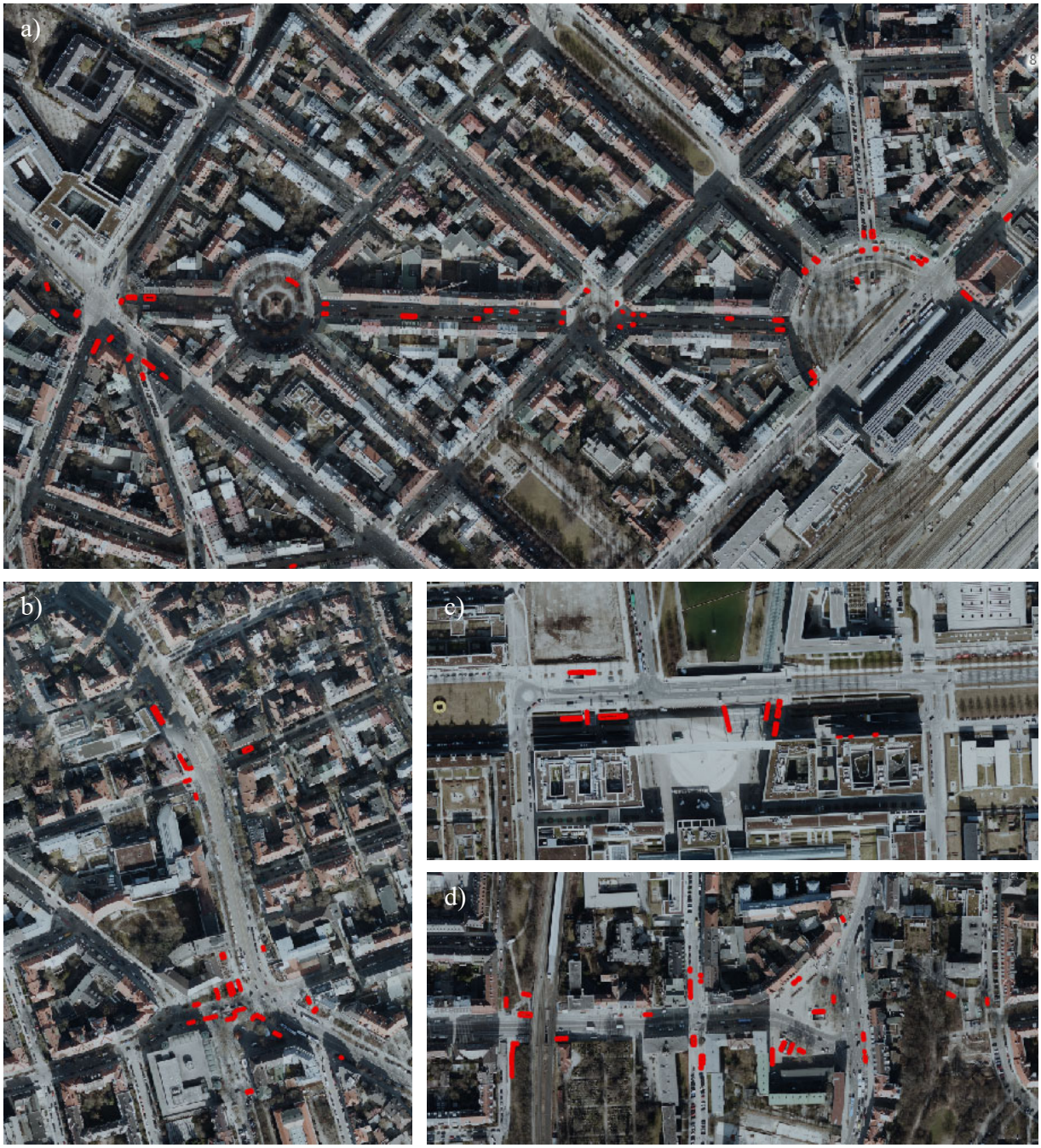


Abb. 3: Aus Trainingsdaten herausgenommene Gebiete: gelabelte Fahrradständer-Objekte in Rot; a) Rosenheimer Platz bis Orleansplatz/Ostbahnhof (50 Objekte); b) Rotkreuzplatz (31 Objekte); c) Harras (30 Objekte); d) Messestadt West (17 Objekte) (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

Tab. 1: Detaillierte Auflistung aller RP, FP und FN Detektionen in allen Datenkombinationen zusammengefasst über alle Gebiete

		RGB	Anteil an detektierten Objekten	CIR	Anteil an detektierten Objekten	+nDOM	Anteil an detektierten Objekten	+NDVI	Anteil an detektierten Objekten	+nDOM +NDVI	Anteil an detektierten Objekten
$RP + RP_{sonstige}$	**korrekt detektiert**	**170**	**86,3%**	**187**	**82,4%**	**188**	**90,4%**	**174**	**89,7%**	**166**	**85,1%**
	privat	71	36,0%	78	34,4%	69	33,2%	70	36,1%	71	36,4%
	korrekt klassifiziert	67	34,0%	79	34,8%	80	38,5%	77	39,7%	75	38,5%
	falsch klassifiziert	20	10,2%	17	7,5%	19	9,1%	16	8,2%	17	8,7%
	einzelne Fahrräder	11	5,6%	10	4,4%	17	8,2%	6	3,1%	1	0,5%
	ungelabelt & öfftl.	1	0,5%	3	1,3%	3	1,4%	5	2,6%	2	1,0%
FP	**falsch detektiert**	**27**	**13,7%**	**40**	**17,6%**	**20**	**9,6%**	**20**	**10,3%**	**29**	**14,9%**
	Dach	6	3,0%	5	2,2%	0	0,0%	1	0,5%	2	1,0%
	Bodenmuster	8	4,1%	7	3,1%	4	2,4%	3	1,5%	8	4,1%
	Friedhof	4	2,0%	3	1,3%	0	0,0%	2	1,0%	3	1,5%
	Auto	2	1,0%	2	0,9%	0	0,0%	0	0,0%	4	2,1%
	Busch, Baum, Gras	6	3,0%	9	4,0%	8	3,8%	4	2,1%	8	4,1%
	Schatten	0	0,0%	4	1,8%	0	0,0%	2	1,0%	1	0,5%
	rechteckiges Objekt	0	0,0%	3	1,3%	2	1,0%	4	2,1%	1	0,5%
	Motorrad, o.ä.	1	0,5%	7	3,1%	5	2,4%	4	2,1%	2	1,0%
Σ	**detektiert**	**197**	**100,0%**	**227**	**100,0%**	**208**	**100,0%**	**194**	**100,0%**	**195**	**100,0%**

			Anteil an zu detektieren-den Objekten		Anteil an zu detektieren-den Objekten		Anteil an zu detektieren-den Objekten		Anteil an zu detektieren-den Objekten		Anteil an zu detektieren-den Objekten
	zu detektieren	128		128		128		128		128	
RP	detektiert & gelabelt	**87**	**68,0%**	**96**	**75,0%**	**99**	**77,3%**	**93**	**72,7%**	**92**	**71,9%**
	korrekt klassifiziert	67	52,3%	79	61,7%	80	62,5%	77	60,2%	75	58,6%
	falsch klassifiziert	20	15,6%	17	13,3%	19	14,8%	16	12,5%	17	13,3%
FN	gelabelt & undetektiert	**41**	**32,0%**	**32**	**25,0%**	**29**	**22,7%**	**35**	**27,3%**	**36**	**28,1%**

Aus diesen Werten können nun Genauigkeit (engl.: precision), Falscherkennungsrate (engl.: false discovery rate, FDR), Trefferquote (engl.: recall) und Falsch-negativ-Rate (engl.: false negative rate, FNR) berechnet werden (Tab. 2). Die Unterscheidung zwischen zwei RP-Werten spiegelt sich dabei in der Berechnung der Werte wider. Bei $Precision_2$ und FDR_2 werden die $RP_{sonstige}$ miteinbezogen und RP_2 stellt die Summe aus RP und $RP_{sonstige}$ dar. Da für private Fahrradständer keine Ground-Truth-Daten vorhanden sind, können hier keine FN-Werte in die Berechnung miteinfließen.

Tab. 2: Precision, FDR, Recall und FNR für alle Datenkombinationen

		RGB	CIR	+nDOM	+NDVI	+nDOM +NDVI
$Precision_2$	$\frac{RP_2}{RP_2 + FP}$	86,3%	82,4%	90,4%	89,7%	85,1%
FDR_2	$\frac{FP}{RP_2 + FP}$	13,7%	17,6%	9,6%	10,3%	14,9%
Recall	$\frac{RP}{RP + FN}$	68,0%	75,0%	77,3%	72,7%	71,9%
FNR	$\frac{FN}{FN + RP}$	32,0%	25,0%	22,7%	27,3%	28,1%
Precision	$\frac{RP}{RP + FP}$	76,3%	70,6%	83,2%	82,3%	76,0%
FDR	$\frac{FP}{RP + FP}$	23,7%	29,4%	16,8%	17,7%	24,0%

Wenn $RP_{sonstige}$ einbezogen wird, erzielt RGB+nDOM mit einer $Precision_2$ von 90,4 % die besten Ergebnisse. CIR schneidet mit 82,4 % am schlechtesten ab. Wenn $RP_{sonstige}$ außer Acht gelassen wird, liegen die Recall-Werte zwischen 68,0 % (RGB) und 77,3 % (RGB+nDOM). Die Precision erreicht Werte zwischen 70,6% (CIR) und 83,2% (RGB+nDOM). In allen Fällen erreicht RGB+nDOM die besten Werte für die Erkennung von Fahrradständern (Recall: 77,3%, Precision: 83,2%). Wenn es um den Anteil der RP im Vergleich zu allen erkannten Objekten (Precision) geht, schneidet CIR (70,6%) weiterhin am schlechtesten ab. Beim Recall schneidet RGB (68,0 %) am schlechtesten ab. RGB+NDVI kommt RGB+nDOM in Bezug auf die Precision am nächsten. Beim Recall ist es CIR.

4.2 Richtig positive Detektionen

In der Analyse zeigt sich, dass die Segmentierungsmasken die Fahrradständer sehr genau erfassen und einrahmen können. Abb. 4a) zeigt korrekt detektierte Fahrradständer mit den roten Rechtecken als Ground-Truth. Zu richtig positiven Detektionen zählen auch Fahrradständer, die nicht von der Stadt München verwaltet werden, da sie beispielsweise auf Privatgrund stehen (Abb. 4b).

Gelegentlich werden auch einzelne Fahrräder als Fahrradständer-Objekt erkannt und markiert (Abb. 4c).

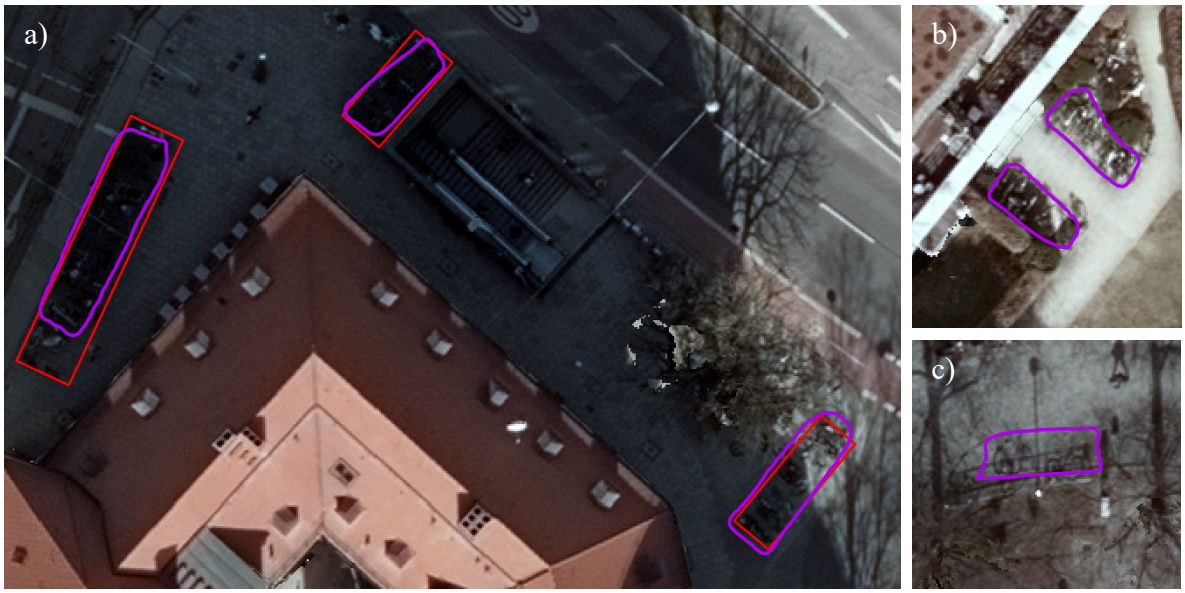


Abb. 4: Beispiele richtig positiver Detektionen: a) korrekt markierte Fahrradständer-Objekte in Lila mit gelabelten Fahrradständern in Rot (Rosenheimer Platz); b) private Fahrradstellplätze (Spicherenstraße); c) einzelne Fahrräder angelehnt an ein Geländer (Weißenburger Platz) (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

4.3 Falsch positive Detektionen

Für fälschlicherweise detektierte Objekte sind Bildstrukturen verantwortlich, die denen eines Fahrradständer-Objektes gleichen. Abb. 5 zeigt Beispiele für die in Tab. 1 aufgeführten Objekte. Bodennahe Strukturen beispielsweise wie Büsche (Abb. 5a), Gräber auf Friedhöfen (Abb. 5b) und Schatten (Abb. 5c) sorgen für falsch positive Detektionen. Auch Gleisbetten (Abb. 5d) ähneln in ihrem Muster denen eines Fahrradständers und können vom Modell falsch erkannt werden. Diese lassen sich aber beispielsweise durch die Verwendung einer Maske eliminieren, die nur Flächen enthält, an denen Fahrradständer aufgestellt werden. Bahnstrecken können dafür ausgeschlossen werden. Falsch detektierte Dachstrukturen (Abb. 5e) können genauso wie fälschlich detektierte Autos (Abb. 5f) durch die Verwendung des nDOM-Rasters vermieden werden.

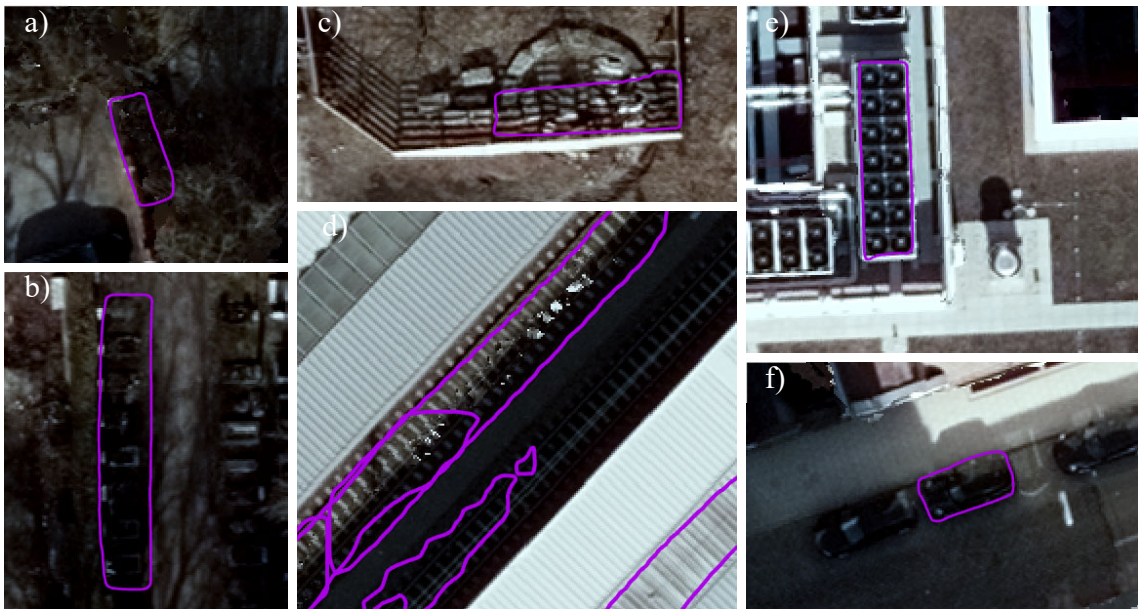


Abb. 5: Beispiele falsch positiver Detektionen: a) Busch (Kidlerplatz); b) Friedhof (Friedhof Sendling); c) Schatten (Kidlerplatz); d) Gleisbett und Dach (Ostbahnhof); e) Dach (Willy-Brandt-Platz); f) Auto (Volkartstraße) (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

4.4 Falsch negative Detektionen

Werden bei der Anwendung von Deep Learning Objekte nicht erkannt, liegt dies zum einen daran, dass nicht genug Beispiele in den Trainingsdaten für ein spezielles Pixelmuster vorliegen. Darüber hinaus bestätigt die Analyse der falsch negativ detektierten Fahrradständer die bereits bei der Trainingsdatensatzerstellung beobachteten Herausforderungen: Verdeckungen durch dichte Äste und sehr dunkle Schatten können eine Detektion verhindern (Abb. 6a, 6b und 6d). Des Weiteren zeigt sich, dass die Anzahl abgestellter Fahrräder von Bedeutung ist: sind diese chaotisch geparkt, beispielsweise weil zu viele Räder auf zu wenig Stellplätze kommen (Abb. 6c) oder ist der Fahrradständer leer (Abb. 6e), werden die Fahrradständer-Objekte unter Umständen ebenfalls nicht erkannt.

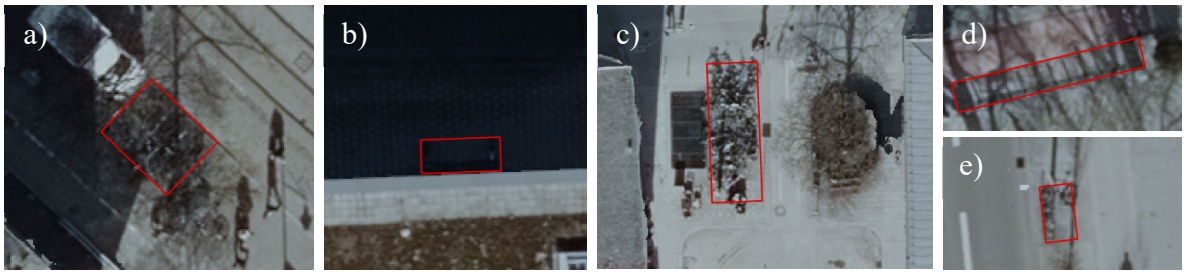


Abb. 6: Beispiele falsch negativer Detektionen (Label in Rot): a) dichte Äste (Orleansplatz); b) dunkler Schatten (Willy-Brandt-Allee); c) große Anzahl chaotisch abgestellter Fahrräder (Albert-Roßhaupter-Straße); d) Baumschatten und leerer Fahrradständer (Rotkreuzplatz); e) (fast) leerer Fahrradständer (Plinganserstraße) (Bildquelle: LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT – GEODATENSERVICE 2020)

4.5 Klassifizierung

Um neben der reinen Detektion auch eine Klassifizierung der Fahrradständertypen zu untersuchen, wird ein Trainingsdatensatz mit allen 17 enthaltenen Fahrradständertypen erstellt. Da zehn der Fahrradständertypen aber durch weniger als 50 manuell markierte Objekte repräsentiert sind, werden anschließend ähnlich aussehende Typen in einer Oberklasse gruppiert. Dadurch entstehen sieben Klassen, die für einen zweiten Trainingsdatensatz verwendet werden. Ebenso wird ein Trainingsdatensatz erstellt, bei dem keine Unterscheidung der Klassen gemacht wird.
Auf der einen Seite kommt es beim Training mit allen Klassen durch die geringe Anzahl an Trainingsbeispielen für einzelne Klassen zu Unter- und Überanpassungen. Auf der anderen Seite zeigt sich keine Verbesserung mehr, wenn keine Gruppierung vorgenommen wird, sodass sich ein Gruppieren als sinnvoll erweist.
Dabei können die Objekte mit einer Rate von 62,5% korrekt klassifiziert werden. Es zeigt sich, dass eine Unterscheidung zwischen Anlehnbügeln und Fahrradständern, bei denen das Vorderrad gehalten wird, möglich ist. Durch die begrenzte Größe des Datensatzes gibt es allerdings nicht für alle Fahrradständertypen genügend Trainingsbeispiele.

5 Diskussion

Insgesamt zeigen alle Datenkombinationen gute Ergebnisse, RGB+nDOM kann aber durch das Hinzuziehen der Objekthöhen zum Beispiel Gebäude und Autos ausschließen und die besten Ergebnisse erzielen. Vegetationssensitive Raster wie Nahinfrarot und NDVI können wenig dazu beitragen, dass Vegetationsobjekte wie Büsche und Bäume nicht fälschlicherweise detektiert werden, da es sich um unbelaubte Winterbilder handelt.
Bei der Betrachtung der Ergebnisse zeigt sich, dass der Schattenwurf insgesamt eine große Rolle spielt. Die Lichtverhältnisse zur Zeit der Bildaufnahme beeinflussen maßgeblich die Trainingsdaten. Zum einen kommt es durch die Aufnahme in Wintermonaten zu ausgeprägtem Schattenwurf durch Gebäude und Bäume, sodass 74% der Fahrradständer in den Trainingsdaten im Schatten stehen. Durchgängige Schatten verändern den Kontrast zwischen Fahrradständer-Objekten und Hintergrund, während unbelaubte Laubbäume ein Schattenmuster werfen, welches die Detektion erschweren kann. Zudem werden Fahrräder detektiert, die einen Schatten werfen, der vom Deep Learning Netzwerk als Fahrradständer interpretiert wird, da in den Trainingsdaten einige Fahrräder in Fahrradständern ebenso diese Fahrradschatten werfen.
Diese Beobachtungen zeigen, dass die Größe und der Inhalt der Trainingsdaten für diese Anwendung ausreichend sind, aber eine Erweiterung des Datensatzes von Vorteil wäre.
Auch die geringe Zahl leerer Fahrradständer in den Trainingsdaten führt dazu, dass diese seltener erkannt werden. Das Vorhandensein abgestellter Fahrräder und deren regelmäßige Anordnung in Fahrradständern sind somit für die Detektion förderlich.
Auch die Genauigkeit der Klassifizierung ist von der Menge und den Inhalten der Trainingsdaten abhängig. Insgesamt lässt sich aber die Aussage treffen, dass anhand der Anordnung der Fahrräder in den Fahrradständern (sehr geordnet oder chaotisch und mit welchem Abstand) eine erste Klassifizierung möglich ist.

6 Fazit & Ausblick

Es zeigt sich, dass bei der Verwendung der RGB-Bilddaten in Verbindung mit dem nDOM Raster die besten Detektionsergebnisse erzielt werden. Die Detektion erreicht eine Genauigkeit von 83,2% und eine Trefferquote von 77,3%. Eine Einschränkung stellt hierbei insbesondere der Schattenwurf durch Gebäude und Bäume dar. Da 74% der Fahrradständer in den Trainingsdaten im Schatten liegen, werden zum einen Fahrradständer, die in der Sonne stehen, schlechter detektiert. Zum anderen verhindern sehr dunkle Schatten ebenfalls eine Detektion. Die Analyse der detektierten Objekte zeigt zudem, dass Fahrradständer besser erkannt werden, wenn dort viele Fahrräder abgestellt sind. Der Grund ist auch hier in den Trainingsdaten zu finden, da diese nur wenige leere Fahrradständer enthalten. Eine Klassifizierung der Fahrradständertypen ist nur bedingt möglich: Aus den 17 unterschiedlichen Typen wurden zunächst sieben Kategorien eingeteilt, welche dann mit einer Rate von 62,5% korrekt klassifiziert werden können. Für eine erste grobe Einteilung der Objekte ist die Klassifizierung allerdings ausreichend.
Aus den detektierten und klassifizierten polygonförmigen Objekten können nun Informationen zur Größe und Ausrichtung der Fahrradständer abgeleitet werden, um diese im 3D-Stadtmodell durch standardisierte 3D-Modelle aus einer Modellbibliothek ersetzen zu können.
Um die Detektion und insbesondere die Klassifizierung noch weiter zu verbessern, ist eine größere Zahl von Trainingsdaten notwendig. Auch um die Wahrscheinlichkeit zu erhöhen, dass das Modell auf andere Städte angewendet werden kann, ist die Erweiterung des Trainingsdatensatzes wünschenswert. Dies adressiert die Problematik, dass im Bereich der Fernerkundung für viele Objekte bisher nicht ausreichend große Datensätze und Normen vorhanden sind, um auf passende Trainingsdaten zugreifen zu können. Dies muss durch Bereitstellung und Katalogisierung der vorhandenen Datensätze vorangetrieben werden.

7 Literaturverzeichnis

CHEN, Z., ZHANG, T. & OUYANG, C., 2018: End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images. Remote Sensing, 10:139. https://doi.org/10.3390/rs10010139.

ESRI, 2022A: Überblick über das Toolset „Deep Learning“. ArcGIS Pro Dokumentation. https://pro.arcgis.com/de/pro-app/latest/tool-reference/image-analyst/an-overview-of-the-deep-learning-toolset-in-image-analyst.htm, letzter Zugriff am 07.01.2022.

ESRI, 2022B: ArcGIS API for Python. ArcGIS Developer. https://developers.arcgis.com/python/, letzter Zugriff am 07.01.2022.

FAST.AI., 2021: vision.transform. List of transforms for data augmentation in CV. https://fastai1.fast.ai/vision.transform.html, letzter Zugriff am 07.01.2022.

GEMEENTE ROTTERDAM, 2022: Rotterdam 3D. https://www.3drotterdam.nl, letzter Zugriff am 07.01.2022.

GIRSHICK, R., 2015: Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV), 1440-1448, https://doi.org/10.1109/iccv.2015.169.

HE, K., GKIOXARI, G., DOLLÁR, P. & GIRSHICK, R., 2017: Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV), 2980-2988, https://doi.org/10.1109/iccv.2017.322.

LANDESHAUPTSTADT MÜNCHEN – KOMMUNALREFERAT - GEODATENSERVICE, 2020: RGB, CIR, nDOM, DOM, DGM, NDVI aus Winterbefliegung 2019.

LI, K., CHENG, G., BU, S. & YOU, X., 2018: Rotation-Insensitive and Context-Augmented Object Detection in Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing, **56**(4), 2337-2348, https://doi.org/10.1109/TGRS.2017.2778300.

LI, K., WAN, G., CHENG, G., MENG, L. & HAN, J., 2020: Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, **159**, 296-307, https://doi.org/10.1016/j.isprsjprs.2019.11.023.

SU, H., WEI, S., YAN, M., WANG, C., SHI, J. & ZHANG, X., 2019: Object Detection and Instance Segmentation in Remote Sensing Imagery Based on Precise Mask R-CNN. In IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 1454-1457, https://doi.org/10.1109/IGARSS.2019.8898573.

Biodiversity Monitoring based on Remote Sensing - Assessment of Forest Height Changes Considering Inaccuracies in Image-based Digital Elevation Models

SELINA GANZ[1] & PETRA ADLER[1]

Abstract: Multitemporal monitoring of vegetation heights holds great potential for forest management and protection. For this purpose, time series of image-based digital elevation models can be used. However, the determination of forest height changes is challenging due to missing reference data and varying quality of the input data. In this study, a method is developed for automatic classification of forest height changes considering inaccuracies in elevation information. To analyse these inaccuracies, heights of buildings scattered throughout the study area were analysed. Based on this analysis, height changes were classified into i) stable ii) stable/growth iii) growth and iv) decline. The developed methods create a change map for each user-defined forest area within Baden-Württemberg, showing forest height changes over six years and provide the potential for a long-term, remote-sensing based biodiversity monitoring.

1 Introduction

Forest structures are essential elements of forests, which are linked with habitat requirements of many protected species (FRANK et al. 2009; KYWE 2012). Old forest stands in particular are rich in structure and essential for carbon storage and –sequestration, as well as for water provisioning and buffering of the microclimate (SPRACKLEN & SPRACKLEN 2019; DE ASSIS BARROS & ELKIN 2021). Old trees play a vital role as habitats for plant, fungus, lichen and animal species that depend on these conserved structures and cannot compete in younger stands, which are often characterized by denser and uniform planting and have unstructured, linear forest edges (BOLLMANN et al. 2009). Unfortunately, despite the high awareness in the European Union about these forests, their abundance is still declining (KNORN et al. 2013; SABATINI et al. 2018). In addition, forest age and vegetation height are important information for management and for the designation of protected areas (VIHERVAARA et al. 2017). Therefore, it is important to monitor forest development in a goal-oriented manner, as discussed in LINDENMAYER & LIKENS (2009) and REYNOLDS et al. (2016).
However, identification of old stands and measurement of tree growth in the field is difficult, destructive and time-consuming (SPRACKLEN & SPRACKLEN 2019). Remote sensing techniques such as airborne laser scanning (ALS), aerial images or satellite data are a cost efficient and promising approach for biodiversity monitoring (JIANYA et al. 2008; KUENZER et al. 2014; VIHERVAARA et al. 2017). Time series of vegetation height changes can help to identify undisturbed, well-growing forest stands for entire landscapes (SPRACKLEN & SPRACKLEN 2019) or countries (GINZLER et al. 2021) and may enable the distinction of old and younger forest stands.

[1] Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg (FVA), Wonnhaldestraße 4, D-79100 Freiburg, E-Mail: [Selina.Ganz, Petra.Adler]@forst.bwl.de

So far, there are no established remote sensing-based monitoring tools for Germany that analyse forest height changes over time.
As a contribution to the project "monitoring of biodiversity with remote sensing tools" (MoBiTools), this study developed remote sensing-based methods to record biodiversity-relevant forest surface structures. For this purpose, image-based digital elevation models were used which cover about one third of the state area of Baden-Württemberg every year and thus provide new data every three years (LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG (LGL) 2021b). A key analysis was the change of vegetation heights, which can be derived from normalized digital surface models (nDSMs). However, the spatially explicit assessment of height changes is challenging: Reference data are rarely available. If reference data is available, the accurate location of reference points and tree canopies, as well as the measurement of tree heights, are subject to uncertainties (ACKERMANN et al. 2020). Moreover, the aerial images from different flight campaigns can differ in quality (ZIELEWSKA-BÜTTNER et al. 2016), and nDSMs can also suffer from inaccuracies for various reasons (WANG et al. 2015; ACKERMANN et al. 2020; JAVADI et al. 2020), which may affect classification. Therefore, a marked change in forest height can only be measured if height change is greater than any biases in the remote sensing measurement (WULDER et al. 2008). This implies that growth increment must exceed the assumed measurement bias (WULDER et al. 2008). Hence, to detect changes, the vertical agreement of the nDSM time series must be evaluated and considered when classifying forest height changes. Consequently, the following research question arises: How can height changes in forests be assessed taking into account inaccuracies in image-based digital elevation models?

2 Methodology

After a short introduction to the study site (2.1), this section describes the nDSM time series as input data for the assessment of forest height changes (2.2), the accuracy assessment of these datasets (2.3) and the change detection of forest heights considering inaccuracies in the input data (2.4). The developed workflow was accomplished using the software R Version 4.0.5 (R CORE TEAM 2021).

2.1 Study site

The study site covers the forest area of Baden-Württemberg, a federal state located in the south-west of Germany. Baden-Württemberg has a total area of about 35,751 km², of which around 40 % (13,718.5 km²) are covered with forest (KÄNDLER & CULLMANN 2014). Most common tree species are Norway spruce (*Picea abies*) with 34 %, European beech (*Fagus sylvatica*) with 22 %, Silver fir (*Abies alba*) with 8 % and oaks (*Quercus sp.*) with 7 %.

2.2 Creation of nDSM time series

The assessment of forest height changes was carried out using aerial images acquired by 82 airborne image flight missions spanning the period from 2011 to 2019. The aerial images were provided by the state agency of spatial information and rural development of Baden-Württemberg (LGL) (LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG (LGL) 2021b) as part of

regular aerial surveys. The surveys are conducted during the vegetation period on a three-year cycle, covering about one-third of the state each year. Using the software SURE of nFrames (ROTHERMEL et al. 2012), photogrammetric point clouds were processed with a ground sampling distance (GSD) of 40 or 50 cm. Flight conditions such as season, time of day and weather conditions, flight settings such as front/side overlap and camera type, as well as image-matching parameters such as SURE version and GSD, varied between the flight missions.
Normalized digital surface models (nDSMs) with a GSD of 1 m generated from aerial images served as the basis for deriving forest heights. A detailed description of the applied methods for deriving aerial image-based nDSMs can be found in SCHUMACHER et al. (2019) and GANZ et al. (2020). The three-year interval of the image flight missions resulted in the time series 2011-2014-2017, 2012-2015-2018, and 2013-2016-2019. The position of the flight missions per year and the composition of the time series are shown in Figure 1. The allocation of relevant flight missions is defined by a regularly updated polygon feature that indicates, in 1x1km squares, the combination of flight missions to be selected for each single year and three-year interval.

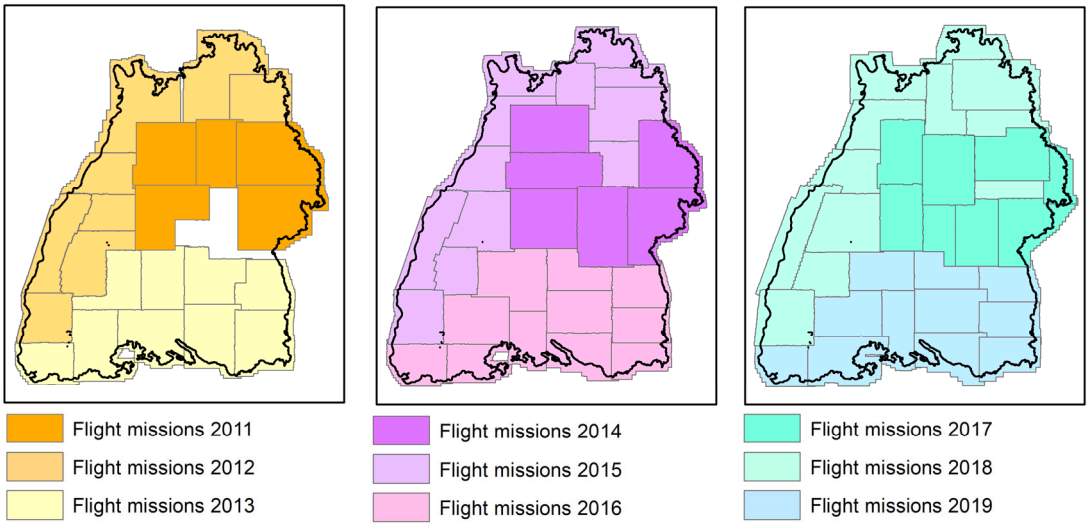


Fig. 1: Flight missions in Baden-Württemberg between 2011 and 2019 are combined into the three time series 2011-2014-2017, 2012-2015-2018 and 2013-2016-2019

2.3 Accuracy assessment of nDSM time series

Within the last decades, various remote-sensing-based methods for the assessment of tree height at stand or single tree level have been developed and studied. The comparison of remote sensing methods for measuring tree height involves a high degree of uncertainty because data acquisition parameters and forest characteristics vary in these studies. As such, it is difficult to transfer measurement errors between the applied methods (GANZ et al. 2019). As a consequence, the evaluation of the accuracy of forest heights cannot be measured against the results of other studies. Depending on the remote sensing system, forest heights can be systematically under- or overestimated. Therefore, in time series analyses, besides the absolute accuracy of the nDSMs, especially the precision of repeated recordings is crucial (WULDER et al. 2008). To evaluate the vertical agreement of the nDSMs, the variation of objects at constant heights was quantified. For

this purpose, the building dataset of the land register of Baden-Württemberg (LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG (LGL) 2021a) was used. A maximum of five buildings per 1x1km quadrant was randomly sampled across Baden-Württemberg so that the buildings were distributed as evenly as possible. For each building, the maximum value of the respective nDSM was extracted for each year of the time series. As the accuracy assessment aimed to determine the nDSM precision rather than changes in buildings, according to the study of DINI et al. (2012) a threshold was set. Only buildings with heights > 2.5 m were evaluated. Furthermore, only nDSM height differences < 2.5 m were considered. Height changes > 2.5 m were considered as architectural changes and were not analysed. Only buildings for which data were available for all three time periods were used.

After applying the thresholds, 87225 building heights could be extracted from three different nDSMs, respectively. The nDSM inaccuracies were evaluated on the basis of the maximum absolute differences of each time series. The distribution of the buildings considered for accuracy assessment, as well as a histogram of the maximum absolute differences, is shown in Figure 2. The maximum absolute differences were ranging between 0 and 2.50 m with a mean value of 0.88 m (standard deviation = 0.46 m). A value of 2 m corresponded to the 97% percentile.

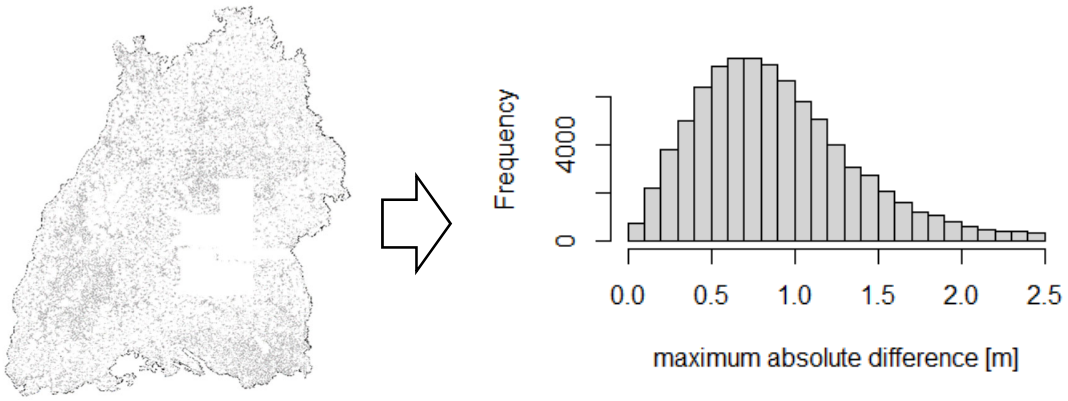


Fig. 2: To evaluate the accuracy of height changes, nDSM values were extracted at 87225 buildings. The maximum absolute differences between nDSMs were calculated to evaluate the vertical agreement of the nDSM time series

2.4 Classification and plausibility check of forest height changes

To minimize inaccuracies at the level of individual pixels, the resolutions of the nDSMs were reduced from 1 m to 10 m, keeping the highest values. Subsequently, the height changes were calculated by subtracting the rasters from each other resulting in nDSM difference rasters (D-nDSMs). Based on the nDSM time series 2011-2014-2017, 2012-2015-2018 and 2013-2016-2019, changes could be analyzed over 6 years as well as over two three-year periods.

Height changes were classified as follows: i) stable, ii) stable/growth, iii) growth, iv) decline. The division of the classes ‘stable’, ‘stable/growth’ and ‘growth’ is based in the height change within six years: The class ‘stable’ covers height differences of < 1 m to restrict height changes to a minimum, while the class ‘growth’ comprises height differences of > 3 m. Based on the estimated inaccuracy of the nDSM time series, a range of 2 m (1 – 3 m height difference) was assigned to the class ‘stable/growth’ to distinguish between the classes ‘stable’ and ‘growth’. The allocation

to the class ‘decline’ was made on the basis of the three-year periods. A significant height reduction of more than 50% of the maximum measured tree height had to occur in one of the three-year periods.
In order to verify the plausibility of the height changes, all available time periods were considered, including two three-year periods and one six-year period. Pixels with non-plausible height changes were not assigned to any category. For example, an increase in height (> 1 m) followed by a decrease in height of the same magnitude was considered to be implausible, since it was presumably due to inaccuracies in the nDSMs. The permitted height changes within the three- and six-year periods are illustrated in Figure 3 and summarized in Tab. 1.

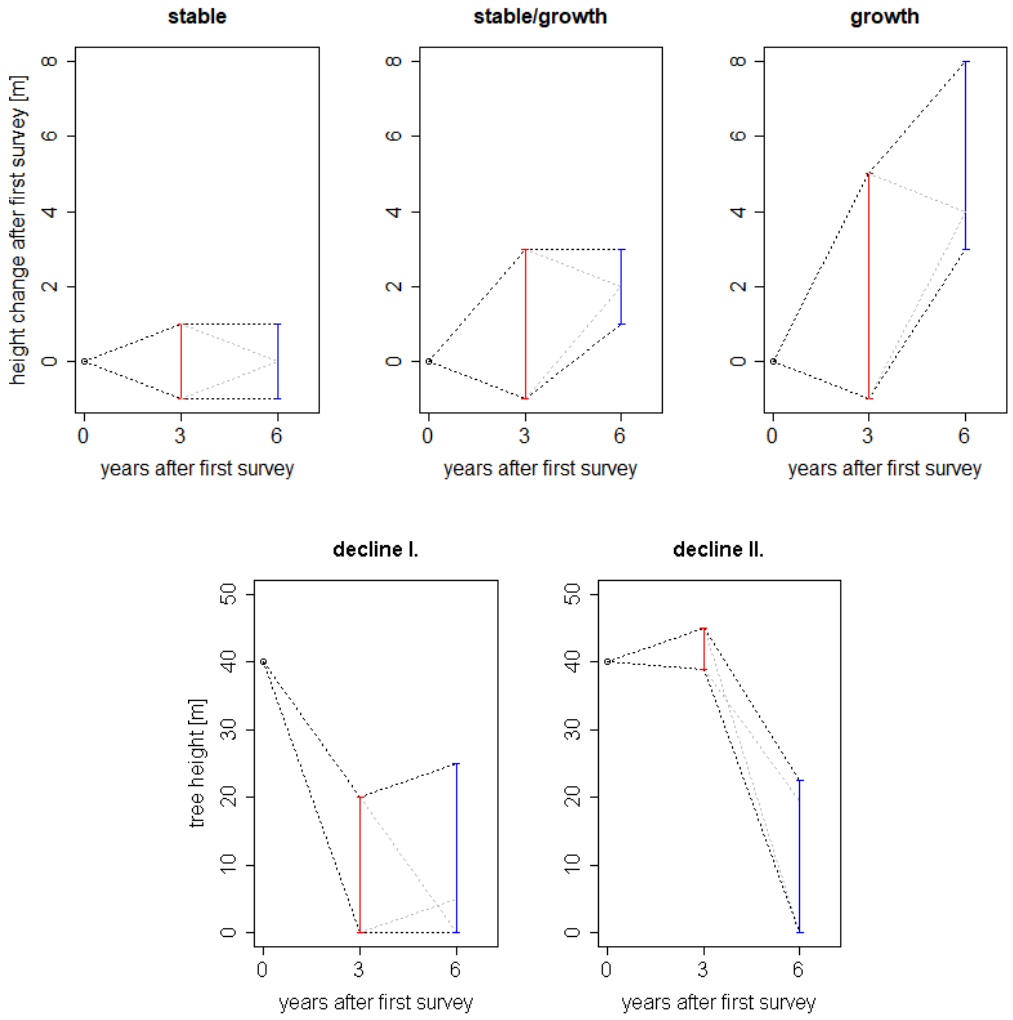


Fig. 3: Categorization of nDSM height changes as ‘stable’, ‘stable/growth’, ‘growth’ and ‘decline’. Height changes are only allowed in the ranges shown: Both after 3 years (red line) and after 6 years (blue line) a plausibility check was conducted. The grey lines correspond to the allowed ranges within the class definition assuming the height development reaches a maximum or minimum value after three years.

Tab. 1: Classification and plausibility check with d as difference between the nDSMs of the time period. The time period defines the analysed height difference after the first survey. Bold values are for classification, the others for plausibility check

Time period	stable	stable/growth	growth	decline
0 to 6 years	**-1 m < d < 1 m**	**1 m < d < 3 m**	**3m < d< 8 m**	**-**
0 to 3 years	-1 m < d < 1 m	- 1 m < d < 3 m	- 1 m < d < 5 m	**i) d < 50% of tree height** OR ii) d < 5 m
3 to 6 years	-1 m < d < 1 m	- 1 m < d < 3 m	- 1 m < d < 5 m	i): d < 5 m OR **ii): d < 50% of tree height**

3 Results

By combining the aforementioned approaches, a tool could be created that automatically evaluates forest height changes for any given area in Baden-Württemberg. While a polygon dataset with the requested areas serves as input dataset, the outputs of the tool are a table with summarized results (*.txt), which are illustrated by graphs and maps for each area (*.png), and optionally raster datasets (*.tif). Thus, spatial and temporal analyses are possible, which can be integrated in research questions regarding biodiversity. Figure 4 shows the developed workflow and illustrative results for four stands of Norway spruce at different ages. The stands are located in the southern Black Forest on comparable site conditions. The examples show that, despite the short time period covered by the time series so far, the proportions of the classes can be very different across stands at different ages. Stand 1 with an age of 31 – 40 years shows with 79% a very high percentage of the change class ‘growth’. The stand 2 and 3 with the ages 61 – 70 years and 81 – 90 years have a mixed proportion of the classes ‘growth’ and ‘stable/growth’, with 41% and 24% ‘growth’ and 27% and 47% ‘stable/growth’, respectively. 28% and 26% each are assigned to ‘NoData’. In stand 2, height reduction was detected on 3% of the area. Stand 4 represents with an age of 121 – 130 years a late-successional forest and shows with 35% ‘stable’, 2% ‘growth’ and 30% ‘NoData’ a high percentage of constant heights and ‘NoData’. In these examples, the class ‘stable’ increases with age while class ‘growth’ decreases. This is consistent with the general finding that height growth of young stands is greatest and decreases with age (PRETZSCH 2019). Although the examples are generally not representative for stands with different tree species, ages, and site conditions, they show that a differentiation of height changes may be possible i) according to the set of categorizations made and ii) across the available time series.

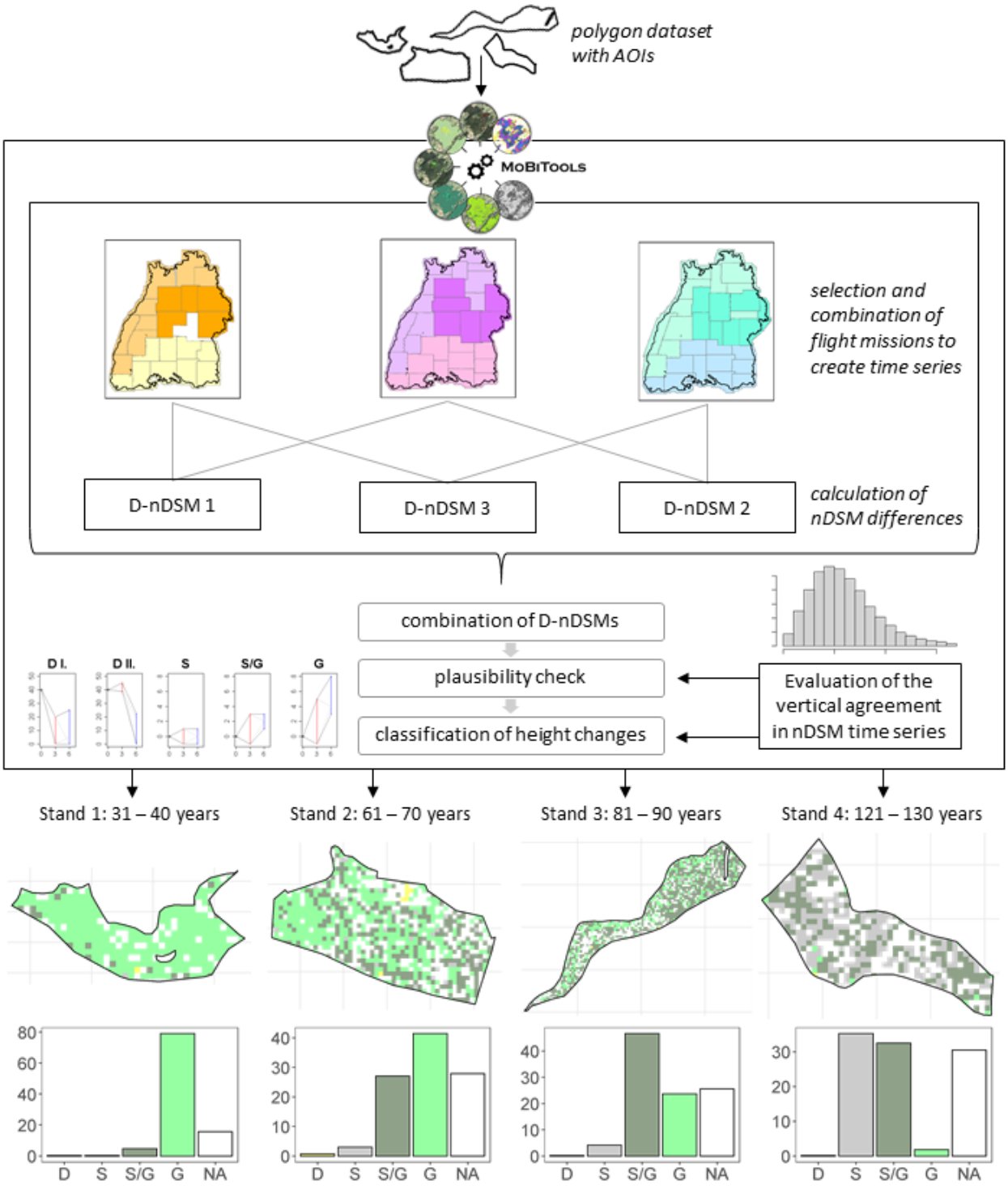


Fig. 4: Developed workflow leading to the assessment of forest height changes for user-defined forest stands in Baden-Württemberg. D-nDSMs are nDSM difference rasters. Graphs with D = decline, S = stable, S/G = stable/growth, G = growth, NA = NoData. The y-axis depicts the percentage of the area

4 Discussion

Continuous change detection based on multitemporal aerial images is a challenge due to variant spectral characteristics, shadows and surface conditions (JAVADI et al. 2020). Changes in the acquisition geometry, the time of acquisition, the weather at the time of flight (ACKERMANN et al. 2020) as well as different view and illumination directions (DINI et al. 2012) influence image-matching for deriving elevation models. Especially in rugged terrain and forests of complex structure, image-based elevation models might be erroneous (HOBI et al. 2015; ZIELEWSKA-BÜTTNER et al. 2016). A direct comparison of height models within a time series can therefore lead to uncertainties and artefacts (ACKERMANN et al. 2020). As a consequence, it is very important to note that changes in tree height can only be measured if the height increase is greater than any biases in remote sensing measurements (WULDER et al. 2008). For that reason, we evaluated the vertical agreement arising by subtracting image-based nDSMs through the analysis of height changes on roofs. Despite filtering by 2.5 m, we cannot exclude the possibility of architectural changes on the analysed roofs. However, due to the high number of samples, the evaluation can be considered reliable. According to the accuracy assessment of 2.3, overlaps between the classes ‘stable’ and ‘growth’ can occur in approximately 3% of the samples.

The accuracies of elevation models vary regarding different land cover categories and provide lower accuracies in forests (ALGANCI et al. 2018). Furthermore, artefacts in elevation models are more frequent in forest areas than in non-forest areas (WANG et al. 2015). Consequently, the threshold for forest areas must be higher than indicated by the analysis of the roofs. A study of GANZ et al. (2019) evaluated the accuracy of tree height measurements based on aerial images from LGL. The accuracy for 30 individual tree heights on a 50-year-old Douglas fir (*Pseudotsuga menziesii*) stand in the Black Forest in Baden-Württemberg was estimated to range between 1 – 2 m. Finally, the accuracy of the measurement depends on the shape of the tree or on treetop visibility (GANZ et al. 2019). According to these findings, the range of 2 m of the class 'stable/growth' can be considered sufficient for differentiation between the classes ‘growth’ and 'stable. According to these findings, a threshold value of 2 m for the determination of height changes can be considered to be sufficient. In order to reduce misclassifications, the forest height changes must lie between certain thresholds to be assigned to one of the change classes. The class ‘stable/growth’ with its range of 2 m acts as a buffer between the classes ‘stable’ and ‘growth’ and thus contains both: areas where tree growth takes place and areas with stable forest heights. To further minimize misclassifications, we decreased the resolution of the nDSMs from 1 m to 10 m keeping the highest value within 10x10 m. As recommended by ACKERMANN et al. (2020), we filter out erroneous data through plausibility check during categorization. However, the relevance and correct categorization of the classes ‘decline’, ‘stable’, ‘stable/growth’ and ‘growth’ cannot be verified due to the absence of reference data.

As the examples show, the results of the time series analysis have the potential to correlate with forest age. Therefore, the nDSM time series can potentially enable the distinction of old forests from younger forest stands. For which tree species, tree age and site conditions this is practicable needs to be further investigated. As height growth is greatest in young trees under good site conditions, the developed methods are only appropriate until a specific forest age and site index.

For older stands with a lower site index, it takes considerably longer for the height increment to exceed the bias of the D-nDSMs. We agree with WULDER et al. (2008) that there is likely to be some error and that this error has to be contrasted with expected growth. For example, tree growth can vary with age and site conditions for Douglas-fir from 0.35 to 1.12 m per year for a 10-year-old stand, from 0.10 - 0.36 m per year for an 80-year-old stand, and from only 0.05 - 0.24 m per year for a 120-year-old stand (WULDER et al. 2008). This illustrates that the present 'stable' and 'growth' classes might be applicable for analysing the age structure of forests, though presumably only under good site conditions. As the examples show, the amount of 'NoData' increases with increasing forest age or stand complexity. The 'NoData' pixels are either related to inaccuracies in the nDSM time series or to heights < 5 m. Old-growth and late-successional forests possess complex structures (FRANK et al. 2009), which complicates the analysis of time series. To be able to evaluate change detection in forests with complex structures or poor site conditions, longer time series are needed.

The required time interval of the images depends not only on the research question but also on the availability of data, e.g. the frequency of aerial image flight missions (ACKERMANN et al. 2020). For our studies, the aerial images were available from 2011 to 2019, resulting in time series of six years. Compared to other studies, the provided time series is very short. VASTARANTA et al. (2015) classified forest stand age using time series of image-based DSMs over a 68-year period. GINZLER et al. (2021) analysed forest height changes with historical vegetation height models based on analogue aerial photographs from the time periods 1980 and 1990 and the current vegetation height model from 2010. The greater the interval between the recording dates and the longer the time series, the more detailed the information on forest height change. With longer time series, the class 'growth' could be further distinguished to differentiate between slow and fast growing forests. High frequency of data recording enables very precise identification of forest dynamics. The three-years interval is sufficient to detect changes in a regular high frequency. At 10×10m, it is possible to detect even small changes in very small groups of trees. When time series grow longer, the current analysis tool, especially the plausibility and categorisation, must be adjusted.

5 Conclusion

Remote sensing enables the observation of natural dynamics that make it a powerful tool for biodiversity-related studies (KUENZER et al. 2014). The derivation of forest stand age (SCHUMACHER et al. 2020) or the identification of the amount and location of old forests (DE ASSIS BARROS & ELKIN 2021) can aid both forest management and forest conservation strategies. Regularly updated image-based digital elevation models from public flight mapping campaigns bear the potential to assess forest height changes in a cost-efficient way, especially when aiming at long-term monitoring. With the methods developed, it is possible to monitor forest height changes over the last six years for a given forest area in Baden-Württemberg. The results can potentially give an estimate of forest age or the proportion of old forests. The calculations are fully automated for a user-defined number of areas and thus offer the possibility to get an overview of entire protected areas as well as individual habitats. Using an appropriately designed workflow, repeated aerial images enable measurement of forest height growth and the differentiation between

tree loss, growth and stable heights over time, giving insights about changes in forest structures and habitats. Although the time series of only six years are still very short for forestry purposes, they already have the potential to contribute to an increase in information. Furthermore, the analyses show the potential of longer time series analyses. If new data sets are added to the time series analyses in the future, the developed analyses can be improved and extended. Therefore, this analysis can be considered as the beginning of a long-term, remote sensing-based biodiversity monitoring program to assess changes of ecologically important forest structures and habitats.

Acknowledgments: We are grateful to Eva Kranefeld for her contribution to literature management as well as for writing and proofreading the manuscript. We acknowledge the provision of the remote sensing data by the land surveying authority of Baden-Württemberg (LGL). This research was funded by "Sonderprogramm zur Stärkung der biologischen Vielfalt" (Land Baden-Württemberg).

6 References

ACKERMANN, J., ADLER, P., AUFREITER, C., BAUERHANSL, C., BUCHER, T., FRANZ, S., ENGELS, F., GINZLER, C., HOFFMANN, K., JÜTTE, K., KENNEWEG, H., KOUKAL, T., MARTIN, K., OEHMICHEN, K., RÜFFER, O., SAGISCHEWSKI, H., SEITZ, R., STRAUB, C., TINTRUP, G., WASER, L. & ZIELEWSKA-BÜTTNER, K., 2020: Oberflächenmodelle aus Luftbildern für forstliche Anwendungen. Leitfaden AFL 2020. WSL-Berichte 87.

ALGANCI, U., BESOL, B. & SERTEL, E., 2018: Accuracy Assessment of Different Digital Surface Models. ISPRS International Journal of Geo-Information, **7**, 114.

BOLLMANN, K., BERGAMINI, A., SENN-IRLET, B., NOBIS, M., DUELLI, P. & SCHEIDEGGER C., 2009: Konzepte, Instrumente und Herausforderungen bei der Förderung der Biodiversität im Wald. Schweizerische Zeitschrift für Forstwesen, **160**, 53-67.

DE ASSIS BARROS, L. & ELKIN, C., 2021: An index for tracking old-growth value in disturbance-prone forest landscapes. Ecol Indicators 121.

DINI, R., JACOBSEN, K., ROTTENSTEINER, F., RAJHI, M. & HEIPKE, C., 2012: 3D building change detection using high resolution stereo images and a GIS database. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIX-B7.

FRANK, D., FINCKH, M. & WIRTH, C., 2009: Impacts of Land Use on Habitat Functions of Old-Growth Forests and their Biodiversity. 429-450.

GANZ, S., ADLER, P. & KÄNDLER, G., 2020: Forest Cover Mapping Based on a Combination of Aerial Images and Sentinel-2 Satellite Data Compared to National Forest Inventory Data. Forests, **11**, 1322.

GANZ, S., KÄBER, Y. & ADLER, P., 2019: Measuring tree height with remote sensing—A comparison of photogrammetric and LiDAR data with different field measurements. Forests, **10**, 694.

GINZLER, C., PRICE, B., WEBER, D., HOBI, M. & MARTY, M., 2021: Dynamik der Vegetationshöhen im Schweizer Wald. Schweizerische Zeitschrift fur Forstwesen 172, 310-317.

HOBI, M., GINZLER, C., COMMARMOT, B. & BUGMANN, H., 2015: Gap pattern of the largest primeval beech forest of Europe revealed by remote sensing. Ecosphere, **6**, art76.

JAVADI, S., DAHL, M. & PETTERSSON M., 2020: Change Detection in Aerial Images Using Three-Dimensional Feature Maps. Remote Sens, **12**, 1404.

JIANYA G., GUORUI M. & ZHOU, Q., 2008: A review of multi-temporal remote sensing data change detection algorithms. Remote Sensing and Spatial Information Sciences, Beijing, China, 3-11 Jul 37.

KÄNDLER, G. & CULLMANN, D., 2014: Der Wald in Baden-Württemberg: ausgewählte Ergebnisse der dritten Bundeswaldinventur. Forstliche Versuchs-und Forschungsanstalt Baden-Württemberg.

KNORN, J., KUEMMERLE, T., RADELOFF, V.C., KEETON, W.S., GANCZ, V., BIRIŞ, I.-A., SVOBODA, M., GRIFFITHS, P., HAGATIS, A. & HOSTERT, P., 2013: Continued loss of temperate old-growth forests in the Romanian Carpathians despite an increasing protected area network. Environ Conserv, **40**, 182-193.

KUENZER, C., OTTINGER, M., LIU, G., SUN, B., BAUMHAUER, R. & DECH, S., 2014: Earth observation-based coastal zone monitoring of the Yellow River Delta: Dynamics in China's second largest oil producing region over four decades. Applied Geography, **55**, 92-107.

KYWE, T.Z., 2012: Habitat Suitability Modeling for Tiger (Panthera tigris) in the Hukaung Valley Tiger Reserve, Northern Myanmar. PhD Dissertation, Georg-August-Universität Göttingen, Göttingen.

LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG (LGL), 2021a: ALKIS. https://www.lgl-bw.de/unsere-themen/Produkte/Geodaten/Daten-des-Liegenschaftskatasters/ALKIS/ 'Accessed' 13.01.2022.

LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG (LGL), 2021b: Digitale Orthophotos. https://www.lgl-bw.de/unsere-themen/Produkte/Geodaten/Digitale-Orthophotos/ 'Accessed' 13.01.2022.

LINDENMAYER, D.B. & LIKENS, G.E., 2009: Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol Evol, **24**, 482-486.

PRETZSCH, H., 2019: Grundlagen der Waldwachstumsforschung.

R CORE TEAM, 2021: R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria.

REYNOLDS, J.H., KNUTSON, M.G., NEWMAN, K.B., SILVERMAN, E.D. & THOMPSON, W.L., 2016: A road map for designing and implementing a biological monitoring program. Environ Monit Assess, **188**, 399.

ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: SURE: Photogrammetric Surface Reconstruction From Imagery.

SABATINI, F., BURRASCANO, S., KEETON, W., LEVERS, C., LINDNER, M., PÖTZSCHNER, F., VERKERK, H., BAUHUS, J., BUCHWALD, E., CHASKOVSKY, O., DEBAIVE, N., HORVATH, F., GARBARINO, M., GRIGORIADIS, N., LOMBARDI, F., DUARTE, I., MEYER, P., MIDTENG, R., MIKAC, S. & KUEMMERLE, T., 2018: Where are Europe's last primary forests? Divers Distrib 24.

SCHUMACHER, J., HAUGLIN, M., ASTRUP, R. & BREIDENBACH, J., 2020: Mapping forest age using National Forest Inventory, airborne laser scanning, and Sentinel-2 data. Forest Ecosystems, **7**, 60.

SCHUMACHER, J., RATTAY, M., KIRCHHÖFER, M., ADLER, P. & KÄNDLER, G., 2019: Combination of multi-temporal sentinel 2 images and aerial image based canopy height models for timber volume modelling. Forests, **10**, 746.

SPRACKLEN, B.D. & SPRACKLEN, D.V., 2019: Identifying European Old-Growth Forests using Remote Sensing: A Study in the Ukrainian Carpathians. Forests, **10**, 127.

VASTARANTA, M., NIEMI, M., WULDER, M., WHITE, J., NURMINEN, K., LITKEY, P., HONKAVAARA, E., HOLOPAINEN, M. & HYYPPÄ, J., 2015: Forest stand age classification using time series of photogrammetrically derived digital surface models. Scand J For Res.

VIHERVAARA, P., AUVINEN, A.-P., POIKOLAINEN, L., TÖRMÄ, M., AHLROTH, P., ANTTILA, S., BÖTTCHER, K., FORSIUS, M., HEINO, J., HELIÖLÄ, J., KOSKELAINEN, M., KUUSSAARI, M., MEISSNER, K., OJALA, O., TUOMINEN, S., VIITASALO, M. & VIRKKALA, R., 2017: How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. Global Ecology and Conservation, **10**, 43-59.

WANG, Z., GINZLER, C. & WASER, L.T., 2015: A novel method to assess short-term forest cover changes based on digital surface models from image-based point clouds. For Int J For Res, **88**, 429-440.

WULDER, M., WHITE, J., GOWARD, S., MASEK, J., IRONS, J., HEROLD, M., COHEN, W., LOVELAND, T. & WOODCOCK, C., 2008: Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sens Environ, **112**, 955-969.

ZIELEWSKA-BÜTTNER, K., ADLER, P., EHMANN, M. & BRAUNISCH, V., 2016: Automated Detection of Forest Gaps in Spruce Dominated Stands Using Canopy Height Models Derived from Stereo Aerial Imagery. Remote Sens, **8**, 175.

Punktwolkenbasierte 3D-Vermessung für KMU in Handwerk und Bauwesen

VERENA HÄNDLER[1], OLIVER REUß[1], JONAS ROTH[1], JÜRGEN SELENSKI[1] & ANSGAR BRUNN[1]

Zusammenfassung: Hochschulen für angewandte Wissenschaften entwickeln beständig die Lehre nicht nur für die eigenen Studierenden, sondern auch die Methoden für den Wissenstransfer in die Unternehmen weiter. Neue didaktische Konzepte werden für anstehende Fachthemen ausprobiert. Im Rahmen des Projekts „Frühe Digitalisierung technischer Planungsprozesse im Bauwesen und Handwerk für kleine und mittlere Unternehmen“ (DiTeP4KMU) sollen daher aktuelle Methoden der Wissensvermittlung (Online-Lehrangebote, Blended Learning sowie hybride Lehrveranstaltungsformen) in die berufsbegleitende Weiterbildung übertragen und realisiert werden. Das Projekt ist eine Fortsetzung und Erweiterung der bereits an der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt umgesetzten Projekte „Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen (DiPhoBi4KMU)“ und „Mobile 3D Vermessung für kleine und mittlere Unternehmen“ (MV-3D4KMU). Ziel des aktuellen Projekts ist die Prozessoptimierung in KMU durch mobile 3D-Vermessung anhand hybrider Wissensvermittlung in Form von E-Learning Angeboten sowie Präsenzveranstaltungen in Form von Netzwerktreffen und Aktionen. Der fachliche Fokus liegt dabei auf der stärkeren Verbreitung der 3D-Daten für alltägliche Aufgaben - insbesondere der schnell zu generierenden 3D-Punktwolken durch moderne Messsysteme - in die Arbeitsprozesse im Bauwesen und Handwerk, in Zusammenarbeit mit den beteiligten Unternehmen. Die Erfahrungen in dem Projekt zeigen bereits nach ca. sechs Monaten ein großes Interesse und bestätigen die Konzeption des Projektentwurfs.

1 Einleitung

Die punktwolkenbasierte 3D-Vermessung hat sich in den letzten Jahren zu einem Standardverfahren der Bauaufnahme und Baustandsdokumentation entwickelt (KICKLER 2021). Dennoch ist die Handhabung und Weiterverarbeitung der entstehenden 3D-Daten für kleinere und mittlere Unternehmen (KMU) im Handwerk und Bauwesen aufgrund der Datenmengen sowie der notwendigen Spezialsoftware immer noch problematisch, was zu einer Begrenzung des Marktwachstums der 3D-Vermessung wie auch der Möglichkeiten der Nutzer (Unternehmen) führt. Gleichzeitig erkennen viele Nutzer die Entwicklungen im Umfeld ihres Unternehmens.
Der technische Fortschritt erfordert immer wieder die Anpassung von Prozessketten im Unternehmen, beginnend beim Auftrag bis zur Umsetzung (Abb. 1). Die Optimierung von Prozessabläufen insbesondere durch frühzeitige Digitalisierung ist dabei eines der Hauptanliegen der Unternehmen, unabhängig von der Branche (LASINGER 2011). Dabei können Großunternehmen meist schnell auf Veränderungen der Umwelt reagieren, da sie die finanziellen Möglichkeiten besitzen, in neue Techniken und Methoden zu investieren, um diese auszuprobieren. Darüber hinaus sind sie in der Lage, Fachkräfte einzustellen, um diese Techniken und Methoden

[1] Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Röntgenring 8, D-97070 Würzburg, E-Mail: [verena.haendler, oliver.reuss, jonas.roth, juergen.selenski, ansgar.brunn]@fhws.de

schnell und gewinnbringend für das Unternehmen einzusetzen. Kleine und mittlere Unternehmen sind hier im Nachteil.

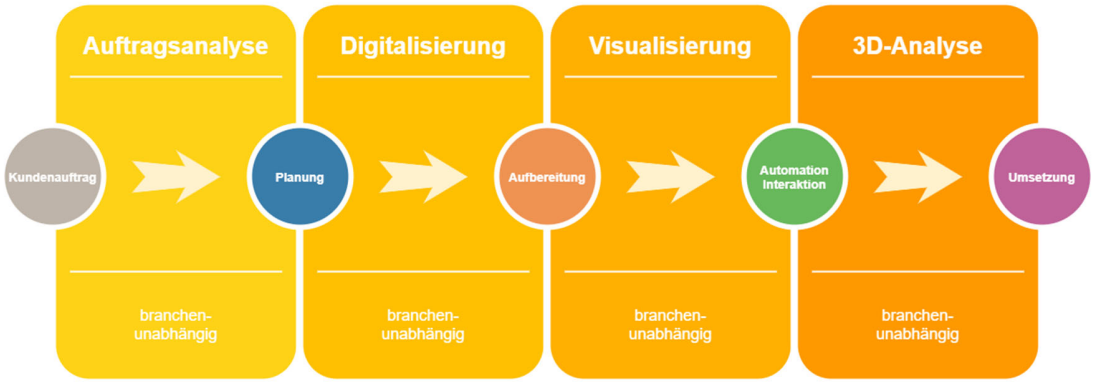


Abb. 1: Prozesskette für 3D-Vermessung in einem Unternehmen

Informations- und Visualisierungstechnologien wie Building Information Modeling (BIM), Augmented- oder Virtual-Reality sind zeitgemäße Mittel, um z.B. für Unternehmen aus dem Bauwesen und Handwerk einen virtuellen Zwilling des Auftragsobjektes zu erstellen, der unabhängig vom Auftragsort bearbeitet werden kann. Die Datengrundlage hierfür liefern 3D-Daten, welche durch modernes Messequipment schnell erfasst werden können. Durch den Einsatz moderner Messverfahren zur Erfassung von 3D-Daten sowie die Verwendung aktueller Software zur Aufbereitung und Weiterverarbeitung der 3D-Daten besteht bei vielen KMU das Potenzial, Arbeitsprozesse effizienter und ökologischer zu gestalten. Dennoch werden diese modernen Möglichkeiten gerade von KMU nur selten genutzt, da der Umgang mit der Erfassung von 3D-Daten sowie deren Verarbeitung nicht vertraut ist.

2 Motivation

Aufgrund der gerade für KMU im Vergleich zu Großunternehmen speziellen Situation bedarf es neuer Methoden, das notwendige Wissen in diese Unternehmen zu bringen. Aus diesem Grund wurde das Projekt „Frühe Digitalisierung technischer Planungsprozesse kleiner und mittlerer Unternehmen im Bauwesen und Handwerk (DiTeP4KMU)“ initiiert.
Ziel des Projektes ist es, neue Technologien zur schnellen 3D-Datenerfassung und –verarbeitung durch moderne didaktische Methoden in die Unternehmen einzubringen und dadurch die Prozesse in den einzelnen Unternehmen zu optimieren. Dabei soll gleichzeitig ein Netzwerk für einen stetigen Wissenstransfer sowohl zwischen den teilnehmenden KMU untereinander als auch zwischen Hochschule und Unternehmen aufgebaut werden, welches über die Projektlaufzeit hinaus besteht. Anhand von E-Learning-Angeboten, Netzwerktreffen und der Analyse einzelner Prozessabläufe soll den KMU Zugang zu den neuen Methoden verschafft und sollen Möglichkeiten aufgezeigt werden, welche Methoden im jeweiligen Unternehmen sinnvoll eingesetzt werden können. Darüber hinaus sollen die KMU auch in die Möglichkeiten der Nutzung von 3D-Daten über das eigene Unternehmen hinaus, also auch in Kooperation mit anderen KMU sowie Vermessern, eingeführt werden.

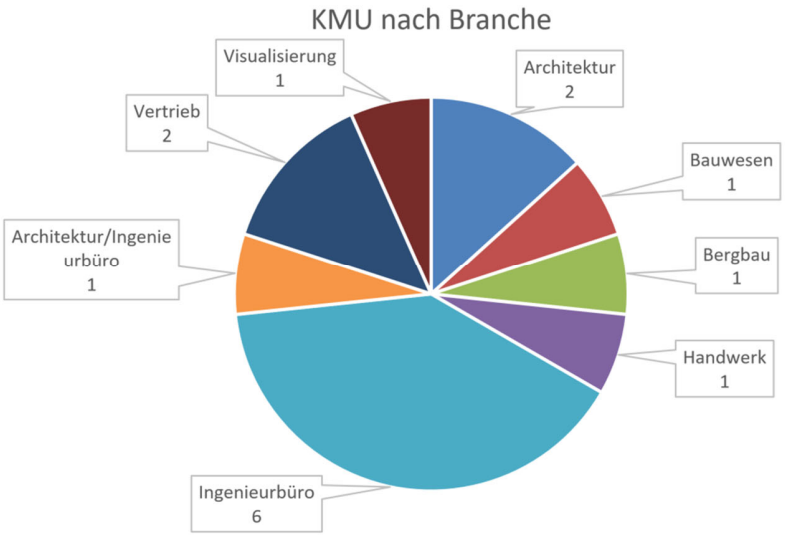


Abb. 2: Teilnehmende KMU nach Branche

Das wissenschaftliche Ziel der Hochschule ist die Erweiterung der modernen Lernvermittlungsmethoden und deren Bewertung durch die praktische Umsetzung. Dazu ist die Interaktion mit und unter den Projektteilnehmenden von großer Bedeutung.
Das Projekt konzentriert sich beispielhaft auf den Umgang mit 3D-Daten innerhalb von KMU. In dem Projekt sind z. Z. 15 Unternehmen aus verschiedenen Branchen (Abb. 2) mit in der Regel weniger als 250 Beschäftigten, vorwiegend aus den strukturschwachen bayerischen Regionen aktiv. Insgesamt nehmen 32 Personen an dem Projekt teil, davon sind 5 Frauen.

3 Handlungskontext

Die fachlich inhaltliche Weiterentwicklung von Unternehmen ist gesellschaftlich als absolut notwendig angesehen und seit den 1970er Jahren mit dem Begriff „Lebenslanges Lernen“ verbunden (BUND-LÄNDER-KOMMISSION FÜR BILDUNGSPLANUNG 1973). Die Unternehmen unterstützen hier vorwiegend die Berufsverbände und die Industrie- und Handelskammern sowie kommerzielle Anbieter, i. d. R. verknüpft mit dem langfristigen Angebot von Hardware, Software und Support. Die Hochschulen für angewandte Wissenschaften widmen sich neben der angewandten Forschung, zukunftsorientierter Lehre dem Wissens- und Technologietransfer in die Praxis, über die eigentliche Lehre der Studierenden hinaus. Unterschiedliche Formate werden hier verfolgt: kooperative Abschlussarbeiten, gemeinsame Forschungs- und Entwicklungsprojekte, Studien mit vertiefter Praxis sowie seminaristische Weiterbildung. Diese Angebote sind i. d. R. von zeitlich kurzer Dauer und behandeln Themen und Fragestellungen sehr fokussiert. Der hier beschriebene neue Ansatz der Unternehmensentwicklung geht darüber hinaus.
Wichtig für diesen Ansatz ist die hybride Informationsvermittlung. Eine anerkannte Methode ist das „Blended Learning“ (z.B. HORN & STAKER 2011, OSGUTHORPE & GRAHAM 2003). An der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (FHWS) wurde das Blended

Learning für KMU bereits in zwei Projekten erfolgreich umgesetzt und kontinuierlich weiterentwickelt. In „Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen - DiPhoBi4KMU“ entstand ein Online-Lernangebot für Photogrammetrie und Bildverarbeitung. Präsenzveranstaltungen wie Projekttreffen oder Workshops ergänzten die Wissensvermittlung zu einem Blended Learning-Modell (MEYER & BRUNN 2016). Gleichzeitig wurde Netzwerkbildung durch Zusammenspiel von realen Treffen und Online-Angeboten im Rahmen eines Blended Networking praktiziert (BRUNN 2018). Das Folgeprojekt „Mobile Vermessung für kleine und mittlere Unternehmen - MV-3D4KMU“ mit dem Fokus auf die mobile 3D Vermessung, erweiterte die Onlinelehre um das Medium Video und das spielerische Lernen (Gamification) (WICH & BRUNN 2019). Die Realisierung einer Prozesskette zur Erstellung von E-Learning Modulen für Blended Learning wird in SELENSKI et al. (2020) beschrieben.
Ein vergleichbares, auch langfristiges Projekt innerhalb der geowissenschaftlichen Hochschullehre ist „Prolearn“ (TEUCHER et al. 2018). Hier wird der Fokus auf das Online-Lernangebot und den Lernfortschritt gelegt. In „Prolearn“ sind jedoch Studierende die Zielgruppe.
In „Fernerkundung in Schulen“ (FIS) wird die digitale Lehre für Schüler langfristig an Schulen etabliert. Dabei werden Werkzeuge für eine bedarfsorientierte Wissensvermittlung, Monitoring des Lernfortschritts und Evaluation der Lernmaterialien (HODAM et al. 2020) angeboten.
Eine Langfristigkeit und hohe Intensität, bei einem gleichzeitig breiten Anwendungsspektrum erreicht keines der genannten Projekte. Hier setzt die Kernidee unseres erweiterten Ansatzes in DiTeP4KMU an.

4 Methode

Die Umsetzung des Projektes erfolgt anhand der fachlichen Impulse aller Beteiligten unter Anwendung moderner didaktischer Methoden wie Blended Learning und Flipped Classroom sowie hybrider Veranstaltungen und praktischen Aktionen. Im Projektdesign sind fünf Komponenten wichtig, die in den folgenden Abschnitten kurz ausgeführt werden.

4.1 Rollenspezifisches Projektkonzept

In dem Projekt werden drei Funktionsgruppen zusammengebracht. Die Leitung sowie die Steuerung des Projekts obliegen der Hochschule. Bei der Wahl der Kooperationspartner wurde darauf geachtet, dass sich verstärkt Unternehmen aus dem Bauwesen und dem Handwerk und aber auch Spezialisten der Vermessung und des Systemvertriebs beteiligen. Dadurch soll sichergestellt werden, dass punktwolkenbasierte Vermessung mit breitem geodätischen Fachwissen und unter Verfügbarkeit modernster Technik für einen möglichst breiten Anwendungsbereich umgesetzt werden kann. Alle Beteiligten können von dem Projekt profitieren. Auf der Anwenderseite können Prozesse optimiert werden. Für die Vermesser bedeutet das Interesse der Anwender eine Erweiterung ihres Tätigkeitsfeldes unter Einsatz modernster Messsysteme. Der Systemvertrieb hat hier die Möglichkeit, neueste Technik für konkrete Anwendungsfälle vorzustellen und somit neue Anwendungsfelder zu bedienen. Die Hochschule wird dabei ihrer Aufgabe gerecht, Wissen und Technologie für ein vielfältiges Anwendungsspektrum in die Praxis zu transferieren. Die Grundlagen werden seitens der Hochschule vermittelt. Konkrete gemeinsam behandelte

Anwendungsfälle stellen die Anwender vor. Die Innovationen werden systemtechnisch durch den Vertrieb sowie methodisch durch die Hochschule eingebracht.

4.2 Blended Learning

Um das Wissen für den Umgang mit 3D-Daten zu den KMU zu bringen, wird mit dem Ansatz des Blended Learning gearbeitet (SEIFERT & PAPE 2014). Das Blended Learning kombiniert Präsenzveranstaltungen mit Online-Angeboten. Es werden sowohl die Vorteile der klassischen Lehre in Präsenz als auch die der virtuellen Lehre genutzt. Durch Online-Kurse können Lernende zeit- und ortsungebunden Wissen aufnehmen. Somit können gerade Berufstätige sich ihre Lernphasen flexibel einteilen und in ihren Arbeitsalltag integrieren. Präsenzveranstaltungen ergänzen diesen Aspekt mit den sozialen Vorzügen der face-to-face-Kommunikation (MEYER & BRUNN 2016). Die Lehrinhalte werden sowohl in Netzwerktreffen in Präsenz vorgetragen als auch auf einer Online-Lehrplattform zur Verfügung gestellt. Präsenz- und Online - Lehrinhalte werden nach einer bestimmten Gliederung abgearbeitet und sind aufeinander abgestimmt.

Blended Learning Präsenzanteil: Durch Präsenzveranstaltungen für die geschlossene Gruppe der KMU findet ein regelmäßiger, persönlicher Austausch und Verbindlichkeit der Teilnehmenden zu dem Projekt statt. Bei den KMU herrscht häufig großer Zeitdruck im Arbeitsalltag und feste Termine neben der Arbeit sind oft nur schwer wahrzunehmen. Daher finden ausführliche Veranstaltungen in Präsenz in größeren Zeitabständen, alle 1-2 Monate statt, was in etwa 10 Präsenzterminen entspricht.

Die Präsenzveranstaltungen in Form von Netzwerktreffen sind ein wesentlicher Bestandteil des Projekts. Inhalte der Netzwerktreffen sind neben Lehreinheiten auch eigene Beiträge der Unternehmen. Die Firmen haben hier die Möglichkeit, sich selber und ihre Projekte allgemein vorzustellen. Außerdem können sie den Ist-Zustand, konkrete Projekte oder Problemfälle bezüglich des Umgangs mit 3D-Daten darlegen. Anschließende Diskussionen der Teilnehmenden sind hier wünschenswert und sollen dazu dienen, verschiedene Fragestellungen und Probleme durch die Expertise und Erfahrungen der anderen Unternehmen beantworten zu können. Die unterschiedlichen Themen und Problematiken werden gesammelt und im Laufe der Projektlaufzeit in engem Austausch mit den KMU bearbeitet. Um die regelmäßige Teilnahme aller kooperierenden Unternehmen zu ermöglichen, finden die Netzwerktreffen in hybrider Form statt.

Blended Learning – Onlineanteil: Der Online-Kurs ist in die E-Learning-Plattform der FHWS eingebunden. Es handelt sich hierbei um die Open-Source-Lehrplattform „Moodle“. Für die Unternehmen werden hier externe Zugänge zur Verfügung gestellt, da im Normalfall nur Studenten und Dozenten Zugriff auf das E-Learning haben. Genauere Informationen zum Aufbau eines solchen Kurses werden in MEYER & BRUNN (2016) gegeben.

Der Kurs baut auf einer einfachen Gliederung der Lehrinhalte auf, die von Anfang an auf der Kursseite abgerufen werden kann. Das gewährleistet eine gewisse Übersichtlichkeit, da der Lernende bei einem unstrukturierten Kurs womöglich schnell das Interesse und die Motivation verliert. Die hierarchisch strukturierten Kurse werden in regelmäßigem Abstand zwischen den Netzwerktreffen hochgeladen.

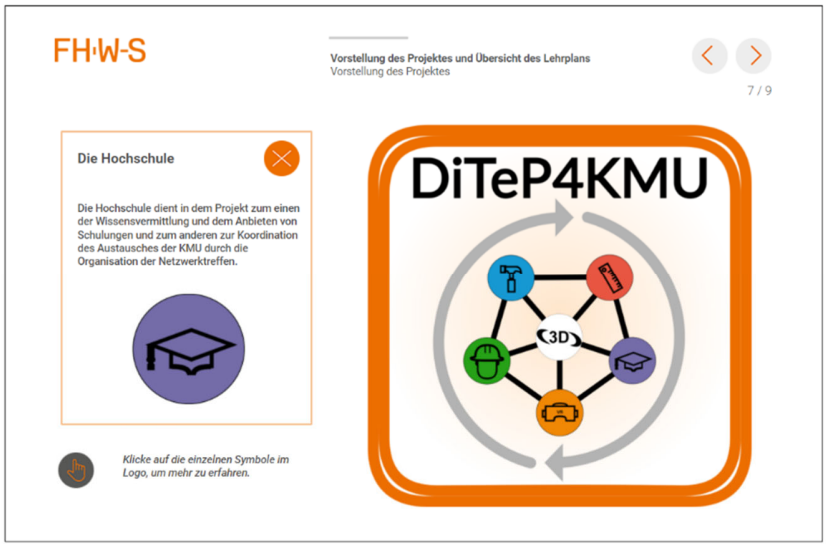


Abb. 3: Die Oberfläche eines Lectora-Kurses

Ein Großteil der theoretischen Lehrinhalte wird durch didaktisch ansprechende, etwa 30minütige E-Learning-Einheiten, die jederzeit unterbrochen werden können, vermittelt. Die Kurse werden mit der Software Lectora erstellt (Abb. 3). Lectora ermöglicht das Erstellen von interaktiven Kursen ohne einen großen Einarbeitungsaufwand. Hier können interaktive, unterhaltsame Elemente wie Drag&Drop-Quizze, Memorys oder Buttons immer wieder problemlos in einen Kurs eingebaut und visualisiert werden. Dazu können selbst erstellte Videos, die ein bestimmtes Themengebiet behandeln, in die Kurseinheit eingebunden werden. Da viel Text oft ungern gelesen wird, dienen solche Videos dazu, die Aufmerksamkeit und Motivation hochzuhalten. Des Weiteren können mit Hilfe von Videos komplizierte Sachverhalte leicht und verständlich beschrieben werden. Evaluationen am Ende der Kurse sollen deren Gestaltung und Inhalt im Laufe des Projektes verbessern. Der exportierte Kurs kann nach einer automatischen Fehlerüberprüfung in E-Learning-Management-Systeme eingebunden werden. Um eine Einheitlichkeit in den Kursen beizubehalten, sind der Aufbau und die Gestaltung der Lectora-Kurse durch hochschuleigene Gestaltungsvorlagen vorgegeben.

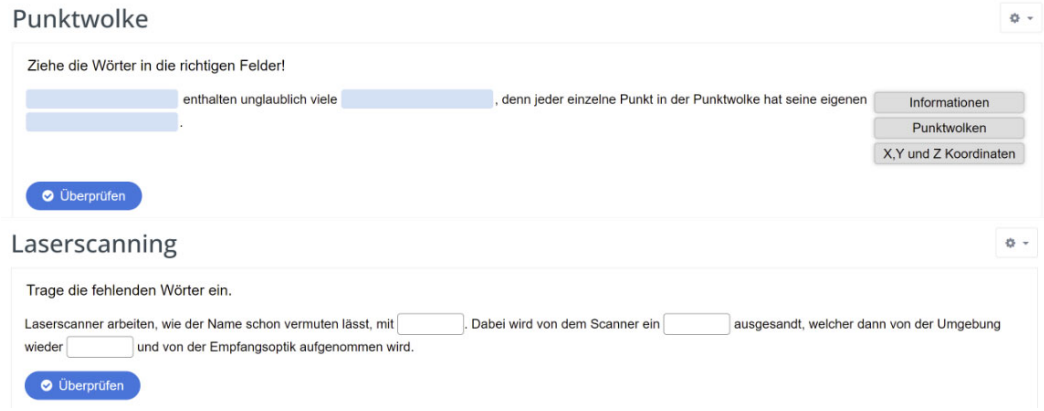


Abb. 4: Oberfläche H5P Software

Auch die in Präsenz vorgestellten Lehrinhalte werden nachträglich im E-Learning als Video zur Verfügung gestellt. Das soll zum einen bewirken, dass Lernende, die nicht an einem Netzwerktreffen teilnehmen konnten, sich auch zeitungebunden die Vorlesung zu den Lehreinheiten ansehen können. Zum anderen können Lernende die Inhalte bei Interesse mehrmals ansehen. Die Videos werden mit der Software H5P eingebunden. Mit H5P können Videos interaktiv gestaltet werden. Das Ziel ist es, durch verschiedene Fragestellungen in den Videos, die Aufmerksamkeit und Motivation des Lernenden hoch zu halten. Die Möglichkeiten dieser Software sollen auch im weiteren Verlauf des Projekts in den Kurs eingebunden werden. Durch Interaktionen wie „Drag the Words“, „Fill in the Blanks“ oder „Drag&Drop“ kann der Kurs interessanter gestaltet werden (Abb. 4).
Die Online-Lehrangebote sollen den Wünschen der KMU angepasst werden. Vor allem in den Netzwerktreffen können unterschiedliche Thematiken, Wünsche und Vorschläge diskutiert werden. Online-Befragungen im weiteren Verlauf des Projekts dienen der Evaluierung des Kurses sowie der Projekttreffen und sollen dazu genutzt werden, den Kurs bzw. die Treffen zu verbessern.

4.3 Flipped Classroom

„Flipped Classroom" bzw. „Inverted Classroom" wird eine Unterrichtsmethode bezeichnet, in der die üblichen Aktivitäten innerhalb und außerhalb des Hörsaals oder Klassenzimmers „umgedreht” werden (E-TEACHING.ORG 2021).
Das bedeutet, die Lehrinhalte eignet sich der Lernende zunächst durch online zur Verfügung gestelltes Lehrmaterial (z.B. in Form von Videos) eigenständig an. In anschließenden Präsenzveranstaltungen wird dieses Wissen dann vertieft und Verständnisfragen geklärt.
In diesem Projekt wird besonders auf die in den Online-Kursen gestellten Quizze eingegangen. Die Fragen bzw. die möglichen Probleme bei der Bearbeitung der Quizze können somit besprochen und diskutiert werden. Da die Netzwerktreffen in dem Projekt nur alle 1-2 Monate stattfinden, muss die Bereitstellung der Online-Kurse zeitlich und thematisch mit den Präsenzveranstaltungen abgestimmt werden.
Zum einen sollten die Kurse nicht zu früh hochgeladen werden, da der Lernende dadurch mögliche Probleme und Fragen bis zum nächsten Treffen vergessen hat. Zum anderen kann ein verspätetes Hochladen des Kurses dazu führen, dass der Lernende zu wenig Zeit zum Vorbereiten hat.

4.4 Hybride Veranstaltungen

Für den Aufbau eines etablierten Netzwerkes zwischen KMU und Hochschule sind Netzwerktreffen in Form von Präsenzveranstaltungen mit einer möglichst vollständigen Teilnehmerzahl unabdingbar. Daher gibt der Projektträger eine Mindestanzahl von Netzwerktreffen sowie eine Mindestanzahl an Teilnehmenden pro Veranstaltung vor. Die ursprünglich ausschließlich in Präsenz angedachten Veranstaltungen werden im Rahmen des Projekts in hybrider Form angeboten. Dies bedeutet, dass die Treffen mit einem Videokonferenzsystem zusätzlich online zugänglich sind. Für dieses Zusatzangebot spricht, dass auch Projektmitglieder, die nicht persönlich anwesend sein können, trotzdem die Möglichkeit haben, aktiv an dem Treffen teilzunehmen und sich in die Diskussionen miteinzubringen. Dagegen, die Veranstaltung ausschließlich online anzubieten spricht, dass der Austausch zwischen den Teilnehmenden ein bedeutendes Element des Projektes ist. Bei einer Präsenzveranstaltung ist

die Hürde niedriger in ein Gespräch zu kommen, da sich die Teilnehmenden vor oder nach der Veranstaltung sowie in den Pausen für gewöhnlich unterhalten.

4.5 Aktionen

Ein wesentliches Mittel für einen praxisorientierten Wissenstransfer ist die Durchführung von durch das Netzwerk abgestimmten gemeinsamen Aktionen. Bei einer Aktion handelt es sich gemäß Projektträger um ein aus den Erkenntnissen der Netzwerktreffen entstandenes gemeinsames Unterprojekt. Zur Bearbeitung des Unterprojekts werden Arbeitskreise aus Hochschule und Unternehmensvertretern aus den KMU gebildet. Die in den Netzwerktreffen angesprochenen Anwendungsfälle können innerhalb der Arbeitskreise detailliert analysiert und Möglichkeiten zur Effizienzsteigerung seitens der Hochschule entwickelt und vorgeschlagen werden. Dadurch kann speziell auf die Bedürfnisse einzelner KMU eingegangen werden. Die Ergebnisse werden im Rahmen von sogenannten Aktionsveranstaltungen allen interessierten Teilnehmenden präsentiert.
Durch die Aktionen kann das vermittelte Wissen direkt für praktische Anwendungsfälle genutzt sowie mögliche Problemstellungen gelöst werden. Dadurch wird die Motivation für die Umsetzung der neuen Methoden gesteigert.

5 Umsetzung und Ergebnisse

5.1 Erstellung Curriculum

Zu Beginn der Projektlaufzeit wurde eingehend der Bedarf an Aus- und Weiterbildung in den beteiligten Unternehmen durch eine Umfrage analysiert. Hierfür wurde mit Hilfe eines Online-Umfragetools ein Katalog angelegt, der Fragen zum Unternehmen und dessen Tätigkeitsschwerpunkt, zum fachlichen Hintergrund der Teilnehmenden und Erfahrungen sowie Wünsche bezüglich der Online-Lehre beinhaltet.

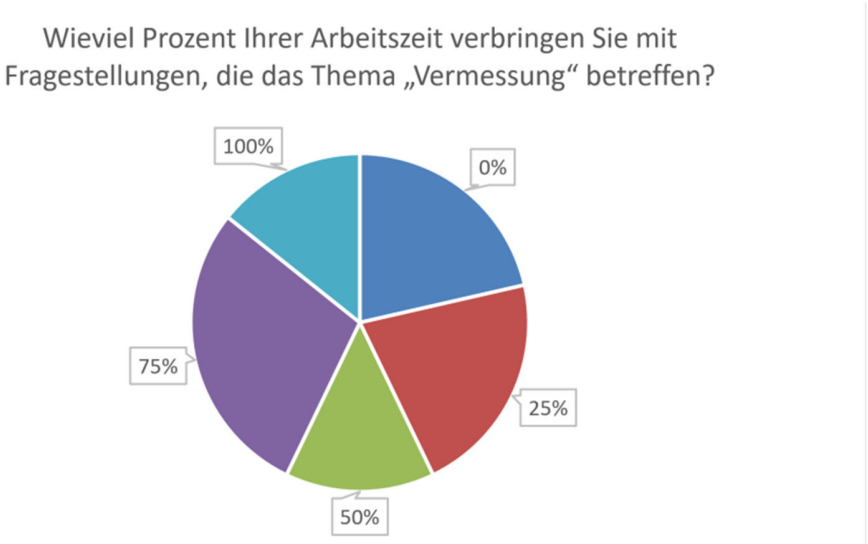


Abb. 5: Teilnehmende beschäftigen sich unterschiedlich stark mit der Vermessung im Arbeitsalltag (in Prozent der täglichen Arbeitszeit)

Die Umfrageergebnisse haben gezeigt, dass die Teilnehmenden verschiedene Wissensstände zur 3D-Vermessung besitzen und sich unterschiedlich stark damit im Berufsalltag beschäftigen (Abb. 5). Außerdem gaben die befragten Personen an, dass theoretische Grundlagen in der 3D-Vermessung für sie wichtig sind. Es war also notwendig, im Curriculum auf die Grundlagen der Verfahren zur Erstellung von 3D-Vermessungsdaten einzugehen, damit alle Beteiligten das gleiche Basiswissen haben.
Die Angabe der im Unternehmen für die Vermessung verwendeten Hard- und Software ist für das Kapitel zur Datenaufbereitung und -weiterverarbeitung hilfreich. Hier bietet es sich an, Theorie und Praxis zu einer Software zu vermitteln, die mit dem bereits in den Unternehmen vorhandenen Equipment und genutzten Dateiformaten kompatibel ist.
Auf die Gestaltung des zuvor genannten Kapitels hatten auch die Antworten auf die Frage “Für welche Aufgaben könnte 3D-Vermessung in ihrem Unternehmen möglicherweise verwendet werden? “ einen Einfluss. Entsprechend der Angaben der Teilnehmenden wurden die Inhalte zur Datenverarbeitung gewählt. Weiterhin zeigten die Befragten Interesse an der Kombination von verschiedenen Vermessungsverfahren. Insbesondere die Antworten zu der Frage nach den Erwartungen an das Projekt haben zur Erstellung des Curriculums beigetragen.
Die Erkenntnisse aus der Umfrage hatten den größten Einfluss auf die Gestaltung des Lehrplans. Ein weiterer Einfluss auf den Aufbau des Curriculums hatten die Präsentationen während dem zweiten Netzwerktreffen. Hier hatten die Unternehmen die Möglichkeit, beispielhaft Anwendungsfälle vorzustellen, für die 3D-Daten genutzt werden sollen. Hierbei wurde unter anderem auch auf Schwierigkeiten beim Workflow im jeweiligen Projekt verwiesen.
Insgesamt besteht das Curriculum aus vier Kapiteln, welche die grobe Struktur des Lehrplans vorgeben und nicht mehr verändert werden. Sie sollen chronologisch möglichst nah am Workflow der Arbeit mit 3D-Daten stehen. Für die Unterkapitel wird zunächst ein Entwurf erstellt, der über die Projektlaufzeit entsprechend den Wünschen und Bedürfnissen der Projektteilnehmenden flexibel angepasst werden kann. Außerdem wurde für die Unterkapitel geplant, welche nur online oder auch in Präsenz (bei einem Netzwerktreffen) behandelt werden (Abb. 6).

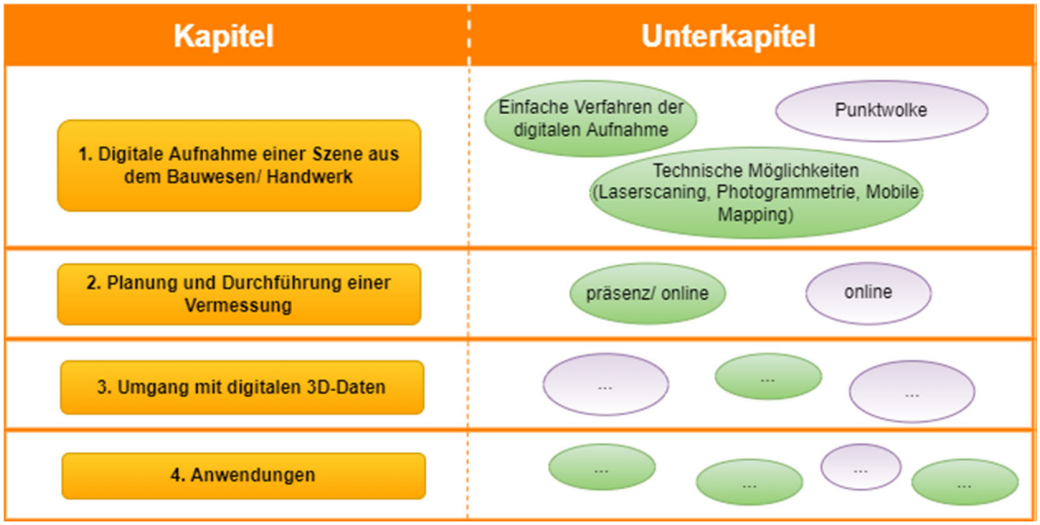


Abb. 6: Aufbau des Curriculums (lila Ellipse = Lehreinheit ausschließlich online, grüne Ellipse = Lehreinheit sowohl in Präsenz als auch online)

Eine Einheit, die beschreibt was man unter einer Punktwolke versteht wurde beispielsweise nur Online zur Verfügung gestellt, weil die Wahrscheinlichkeit hoch ist, dass dies nur für sehr wenige Teilnehmende komplett neues Wissen ist. Alle anderen können so diese bereits bekannten Themen überspringen. Kapitel, die für einen größeren Teil der Lernenden interessant sind, werden dagegen eher in Präsenz vermittelt und anschließend als Video in den Online-Kurs hochgeladen.
Im Verlauf des Projekts soll die Möglichkeit bestehen, einzelne Unterkapitel des aktuellen Curriculums zu streichen oder neue hinzuzufügen, allerdings nur, falls sich Themen ergeben, die für alle Teilnehmenden relevant sind.

5.2 Netzwerktreffen

Im Rahmen des 1. Netzwerktreffens hatten alle teilnehmenden Unternehmen die Möglichkeit sich vorzustellen. Bereits bei dieser Kurzvorstellung ergaben sich Fragestellungen zu einzelnen Anwendungsfällen und Themen, die bestimmte Branchen der KMU besonders interessieren: Zum einen sind mobile Systeme zur Erfassung von 3D-Daten, ihre Handhabung und die mit diesen Systemen erreichbaren Genauigkeiten ein wichtiges Thema, zum anderen der Umgang mit dreidimensionalen Punktwolken. Durch die detaillierte Vorstellung verschiedener Anwendungsfälle im 2. Netzwerktreffen konnten in der Diskussion bereits einzelne Best Practice Beispiele unternehmensgetrieben identifiziert werden. Durch die Hochschule wurden diese aufgegriffen und entsprechend zu den Themenfeldern sortiert und in Hintergrundvorträgen (praxisbezogene Lehre) verarbeitet. Aktuell erfolgt neben der Aufbereitung der Themen in den Lehreinheiten die Einbindung aufgezeigter Best Practice Beispiele in einzelne Unternehmensprozesse.
Für die KMU wird eine Marktübersicht über aktuell vorhandene mobile Systeme zur Erfassung von 3D-Daten erstellt und den KMU zur Verfügung gestellt. Darüber hinaus werden Anbieter (teilnehmende KMU) diverse Systeme im Rahmen der Netzwerktreffen vorstellen. Ingenieurbüros (teilnehmende KMU) geben anhand von Vorträgen Einblicke in ihre Arbeit mit der Erfassung von 3D-Daten und zeigen somit die vielfältigen Einsatzmöglichkeiten der Messsysteme auf.

5.3 Aktionen

Im Rahmen der ersten Aktion erfolgte eine Bedarfsanalyse anhand einer Online-Umfrage unter den Projektteilnehmenden. Die Ergebnisse und Erkenntnisse für das weitere Vorgehen wurden im Rahmen des ersten Netzwerktreffens vorgestellt. Exemplarische Anwendungsfälle wurden so identifiziert und präsentiert. Die dargestellten Anwendungsfälle werden im nächsten Schritt i. w. durch die FHWS bearbeitet.
In der ersten Anwendung erfolgte der Besuch eines Gips-Abbau-Unternehmens. Aus der Dokumentation des aktuellen Messszenarios im Untertagebau zur Abtragung von Gestein sowie über Tage zur Halden- und Haufvermessung werden Lösungsansätze zum Einsatz von 3D-Vermessung zur Prozessoptimierung entwickelt und im weiteren Verlauf anhand von Testmessungen auf ihre Eignung und Praktikabilität geprüft. Ein zweiter Anwendungsfall umfasst die generelle Umstellung eines Unternehmens von 2D- auf 3D-Vermessung. Das Unternehmen möchte herausfinden, wie groß der Mehraufwand einer 3D-Datenerfassung ist und welchen Mehrwert die erzeugten 3D-Bestandspläne als Datengrundlage bieten. Im Bereich Architektur und Visualisierung wurde innerhalb der bisherigen Netzwerktreffen die Weiterverarbeitung und -nutzung von 3D-Punktwolken als Problematik identifiziert. Teilweise stellt die Übertragung der

Punktwolkenformate in die verfügbare CAD-Software sowie das Extrahieren geeigneter Informationen aus den 3D-Punktwolken eine Herausforderung für die Unternehmen dar. Geplant ist hier, den aktuellen Prozessablauf zu analysieren und alternative Prozesse und Softwareprodukte im Rahmen des Projekts zu evaluieren.
Die Ergebnisse der Umsetzung der Prozessoptimierung für die genannten Anwendungsfälle werden im Rahmen von Netzwerktreffen allen interessierten teilnehmenden KMU präsentiert. Die Präsentationen sind für die zweite Hälfte der Projektlaufzeit geplant. Sie umfassen aktuell die Umstellung auf punktwolkenbasierte 3D-Datenerfassung in einem Unternehmen, den Umgang mit 3D-Punktwolken hinsichtlich der Datenformate und des Arbeitens in 3D-Punktwolken, die Konzeption eines Messsystems/Messaufbaus zur 3D-Erfassung für ein immer wiederkehrendes Messszenario sowie die Präsentation von Vergleichsmessungen mit unterschiedlichen Systemen.

6 Diskussion aktueller Ergebnisse und Ausblick

Bereits im ersten Netzwerktreffen wurde von den Beteiligten großes Interesse für das Thema mobiler 3D-Datenerfassung und Einbringen der 3D-Daten in die bestehenden Arbeitsprozesse bekundet. Dies bestätigen auch die nahezu vollständige Teilnahme aller Projektmitglieder und der rege Austausch im ersten Treffen. Durch die hybride Form der Netzwerktreffen ist es bis jetzt gelungen, die Teilnahmequote hoch zu halten.
Die Bedarfsanalyse anhand der umfangreichen Online-Umfrage im Vorfeld hat eine gute Datengrundlage für die Erstellung des Curriculums geliefert und wird auch im weiteren Verlauf bei der Aufbereitung der Projektthemen berücksichtigt. Jedes Unternehmen wurde berücksichtigt. Allerdings hätten noch mehr Mitarbeitende den Fragebogen ausfüllen können, da meist mehr als eine Person pro Unternehmen am Projekt teilnimmt. Dadurch könnten die Bedürfnisse der einzelnen Personen, die meist auch die direkten Anwender der neuen Methoden sind, besser berücksichtigt werden.
Bis jetzt wurde das erste Kapitel des Curriculums behandelt und die Lehreinheiten online zur Verfügung gestellt. Erste Erfahrungen zeigen, dass die Prinzipien des „Flipped Classroom“ hier nicht eins zu eins übernommen werden können, da die Ausgangssituation der Projekt-teilnehmenden nicht wie bei Studierenden oder Schülerinnen und Schülern ist. Es handelt sich bei dem Projekt um eine Weiterbildung der Teilnehmenden während der täglichen Arbeit, d.h. das Zeitkontingent für die Online-Lehreinheiten ist begrenzt. Die aktive Teilnahme kann nur erreicht werden, wenn die Lehrinhalte konkret auf die Bedürfnisse der KMU angepasst sind und didaktisch gut aufbereitet werden, d.h. kurzweilig und mit praktischen Beispielen.
Im weiteren Verlauf sollen die Online-Lehreinheiten in den Netzwerktreffen nochmals aufgegriffen und besprochen werden und so der direkte Zusammenhang mit den Themen der Netzwerktreffen verdeutlicht werden. Außerdem wird noch mehr darauf geachtet, die Online-Lehreinheiten zeitlich und inhaltlich auf die aktuellen Themen der Netzwerktreffen und auch Aktionen abzustimmen.
Betrachtet man die Gewichtung der Anbieter und Anwender im Projekt (Abb. 2) wird klar, dass ein großer Anteil der Beteiligten bereits Fachkenntnisse auf dem Gebiet mobiler 3D-Vermessung hat. Eine Herausforderung bei der Erstellung der Lehrinhalte liegt darin, trotz der bereits

vorhandenen Fachkenntnisse einzelner Personen, allen Beteiligten die Grundlagen der 3D-Vermessung zu vermitteln.
Das Projekt DiTeP4KMU richtet sich an KMU aus dem Bereich des Bauwesens und des Handwerks. Die Prozessoptimierung mittels moderner Formen der didaktischen Aufbereitung wird hier am Beispiel der Erfassung und Verarbeitung von 3D-Daten gezeigt. Die Vorgehensweise, d.h. die Einbindung der Hochschule (Koordination, didaktische Aufbereitung von Lehrinhalten durch Blended Learning etc.), der Austausch der Unternehmen über hybride Netzwerkveranstaltungen und die Prozessanalyse in einzelnen Unternehmen kann so aber auch auf andere Branchen und andere Prozesse von Unternehmen übertragen werden.
Danksagung: Dieser Beitrag wurde innerhalb des Projekts „Frühe Digitalisierung technischer Planungsprozesse im Bauwesen und Handwerk (DiTeP4KMU)“ durch den Europäischen Sozialfond (ESF) und das Bayerische Staatsministerium für Wissenschaft und Kunst gefördert.

7 Literaturverzeichnis

BRUNN, A., 2018: Blended Learning in Photogrammetrie und Bildverarbeitung für regionale KMUs–Ein Erfahrungsbericht. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **27**, 450-460.

BUND-LÄNDER-KOMMISSION FÜR BILDUNGSPLANUNG 1973: Bildungsgesamtplan 1973. Unterrichtung durch die Bundesregierung, 36-41.

E-TEACHING.ORG 2021: Inverted Classroom. e-teaching.org.

HODAM, H., RIENOW, A. & JÜRGENS, C., 2020: Bringing Earth Observation to Schools with Digital Integrated Learning Environments. Remote Sensing, **12** (3), 345, https://doi.org/10.3390/rs12030345.

HORN, M.B. & STAKER, H., 2011: The rise of K-12 blended learning. Innosight Institute, **5**, 1-17.

KICKLER, J., 2021: Punktwolken in der Praxis – Einsatz von der Denkmalpflege bis zum Neubau. In: 21. Internationale Geodätische Woche Obergurgl 2021. Weinold, T. (Hrsg.), Berlin, Wichmann, 71-81.

LASINGER, D., 2011: Die Leistung vor der Innovation: Ermittlung und Nutzung schwacher Signale von Chancen. Zugl.: Linz, Univ., Diss., 2010.

MEYER, T. & BRUNN, A., 2016: Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen durch „Blended Learning“. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **25**, 436-448.

SEIFERT, C. & C. PAPE 2014: Didactics of teaching networking courses with blended learning aspects. Unpublished, 13.

SELENSKI, J., WICH, V. & BRUNN, A., 2020: Erfahrungen mit einem Regelbetrieb der Erstellung von E-Learning-Angeboten für Photogrammetrie und Mobile 3D-Vermessung. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **29**, 411-421.

TEUCHER, M., GLÄßER, C. & SCHREINER, V., 2018: E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **27**, 465-470.

OSGUTHORPE, R. & GRAHAM, C., 2003: Blended Learning Environments: Definitions and Directions. Quarterly Review of Distance Education, **4** (3), 227-233.

WICH, V. & BRUNN, A., 2019: Methodik des Blended Learnings im Bereich der Aus- und Weiterbildung der mobilen 3D Vermessung. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **28**, 270-281.

Bewertung des Phänotyps in der landwirtschaftlichen Züchtungsforschung anhand der photogrammetrisch bestimmten Wuchshöhe und spektraler Indizes am Beispiel verschiedener Genotypen der Gerste

PAUL HERZIG[1], UWE KNAUER[2], PETER BORRMANN[1], DAVID KILIAS[2], HANS-CHRISTIAN KLÜCK[2], UDO SEIFFERT[2], KLAUS PILLEN[1] & ANDREAS MAURER[1]

Zusammenfassung: Durch den Fortschritt auf dem Gebiet der Pflanzengenomik hat sich die Phänotypisierung von Pflanzen im Freiland zu einem neuen Engpass in der Pflanzenzüchtung entwickelt. Es besteht ein steigender Bedarf an objektiver, präziser und zeitlich hochaufgelöster Charakterisierung eines sich entwickelnden Pflanzenbestandes. Die zerstörungsfreie und sensorgestützte Hochdurchsatz Phänotypisierung mittels Drohnen hat das Potential diese Lücke zu bedienen. In dieser Studie wird die Eignung zweier Sensorsysteme (Rot-Grün-Blau (RGB) Kamera und Multispektralkamera) unter der Verwendung unterschiedlicher Modellansätze zur Ertragsvorhersage einer Gerstenpopulation geprüft. Die an unbemannten Luftfahrzeugsystemen (UAVs) montierte Sensorik wurde wöchentlich zur Datenerfassung an 3 Standorten in Mitteldeutschland genutzt. Zusätzlich zu Vegetationsindizes wurden auch morphologische Merkmale wie Bestandshöhe, Bedeckungsgrad und Parameter der Wachstumsdynamik für die Ertragsvorhersage verwendet und mit Referenzwerten aus bodengestützten Phänotypisierungen verglichen. Die Präzision der Ertragsvorhersage beider Sensorsysteme lag auf dem gleichen Niveau und erreichte eine maximale Vorhersagegenauigkeit von r^2 = 0,82 mit einer kontinuierlichen Zunahme der Genauigkeit über die gesamte Vegetationsperiode.

1 Einleitung

Trotz zahlreicher Fortschritte auf dem Gebiet der Genetik und der Anwendung neuer molekularer Technologien in der Pflanzenforschung (GOWEN et al. 2013) hat sich der Ertragszuwachs durch züchterische Forschung bei den wichtigsten Nutzpflanzen in vielen Regionen der Welt stabilisiert oder stagniert sogar (ACRECHE et al. 2008; SADRAS & LAWSON 2011). Im Bereich der Pflanzenzüchtung ist die Phänotypisierung im Feld im Hinblick auf den hohen Durchsatz und die Qualität einer der größten Engpässe in Züchtungsprogrammen (ARAUS & CAIRNS 2014; TARDIEU et al. 2017).

Fernerkundungstechniken nutzen die Lichtreflexion des Pflanzenbestandes, um den Ertrag auf der Ebene des Genotyps zerstörungsfrei zu bewerten (REYNOLDS et al. 1999). Zu diesem Zweck werden mit passiven Sensoren Informationen über die Reflexion elektromagnetischer Wellen des Pflanzenbestandes gewonnen. Die gemessene Reflexion hängt von den chemischen und

[1] Institut der Agrar- und Ernährungswissenschaften, Professur für Pflanzenzüchtung, Martin Luther Universität Halle-Wittenberg, Betty-Heimann-Str. 3, D-06120 Halle; E-Mail: paul.herzig@hs-anhalt.de; peter.borrmann@julius-kuehn.de; [klaus.pillen, andreas.maurer]@landw.uni-halle.de

[2] Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF, Sandtorstraße 22, D-39106 Magdeburg; E-Mail: uwe.knauer@hs-anhalt.de; udo.seiffert@compolytics.com

morphologischen Eigenschaften der Oberflächen ab und ändert sich mit dem Pflanzentyp, der Biomasse, dem Entwicklungsstadium, der Vitalität und den physiologischen Eigenschaften wie dem Wassergehalt und den Pflanzenpigmenten (BABAR et al. 2006; SILVA-PEREZ et al. 2018). Bestimmte Pflanzenmerkmale sind mit der Absorption ganz bestimmter Wellenlängen der elektromagnetischen Strahlung verbunden. Darüber hinaus können mehreren spezifischen Wellenlängen durch Koeffizienten, die so genannten Vegetationsindizes (VI), zusammengefasst werden.

Bei vielen UAV-Fernerkundungsanwendungen werden die Lichtspektren im sichtbaren Bereich (400-700 nm), im nahen Infrarot (700-1200 nm) und im kurzwelligen Infrarot (>1200 nm) mit multispektralen, hyperspektralen oder herkömmlichen Rot-Grün-Blau (RGB)-Kameras gemessen. Multi- und vor allem hyperspektrale Sensoren haben eine hohe spektrale Auflösung, sind aber in der Regel teurer und schwerer als handelsübliche RGB-Kameras. Außerdem ist die spektrale Erfassung empfindlicher gegenüber den Lichtverhältnissen der Umgebung als bei der Farbfotografie. Multispektralsensoren sind die verdichtete Form eines Hyperspektralsensors durch (typischerweise) anwendungsspezifische Bandenauswahl, da mehrere, über Filter frei wählbare Wellenlängenbänder im sichtbaren und infraroten Spektrum erfasst werden können und dennoch die Nutzung eines breiten Spektralindexes ermöglichen (NEBIKER et al. 2016). Aber auch ultrahochauflösende Bilder von kostengünstigen RGB-Kameras mit einer allgemein hohen Qualität der werkseitigen Farbkalibrierung, bieten eine breite Palette von Anwendungen für die Phänotypisierung. Die hohe räumliche Auflösung ist der Hauptvorteil von RGB-Sensoren, die zur Rekonstruktion der 3D-Struktur des Blätterdachs auf der Grundlage des Structure from Motion-Algorithmus (SfM) (ULLMAN 1979) verwendet werden können und somit die Morphologie der Fläche sehr gut wiedergeben können. Die zusätzlichen Informationen über die Architektur des Blätterdaches können dann in die Modellierung der Ertragsbildung und Biomasse integriert werden (BENDIG et al. 2014; GARCÍA-MARTÍNEZ et al. 2020).

Landwirtschaftsmodelle, die sich auf Fernerkundungsdaten stützen, z. B. zur Ertragsschätzung, basieren zunehmend auf Methoden des maschinellen Lernens. Diese Methoden sind oft robuster und genauer als herkömmliche korrelative Methoden, da sie sich durch Training flexibel an die Komplexität der Daten anpassen können (LIAKOS et al. 2018). Darüber hinaus kann die Varianz einer vorherzusagenden Variable entweder durch parametrische oder nicht-parametrische Ansätze erklärt werden. Erstere sind aufgrund ihrer vordefinierten Struktur leichter zu interpretieren; letztere erfordern in der Regel mehr Training, sind aber auch anpassungsfähiger. Häufig verwendete Methoden sind die multivariate Regression (HAN et al. 2019; FERNANDEZ-GALLEGO et al. 2020), Entscheidungsbäume (HAN et al. 2019; LI et al. 2020), Support-Vektor-Maschinen (HAN et al. 2019) oder künstliche neuronale Netze (HAN et al. 2019; GARCÍA-MARTÍNEZ et al. 2020).

Trotz zahlreicher vielversprechender Fortschritte auf dem Gebiet der UAV-gestützten Hochdurchsatz-Phänotypisierung haben die Züchter in der Vergangenheit aufgrund der Komplexität und der Kosten gezögert, diese Technologien in ihre bereits erfolgreichen Züchtungspipelines zu integrieren. In dem Maße jedoch, wie das Wissen auf der Ebene der translationalen Forschung zwischen technologischen und biologischen Bereichen wächst, können sowohl die Komplexität der Datenerfassung und -verarbeitung als auch die Kosten der

Technologie reduziert werden, was die Akzeptanz erhöht. Darüber hinaus wird die Akzeptanz durch den Nachweis von Effizienzgewinnen bei der Erreichung von Zuchtzielen durch neue innovative Methoden erhöht (REYNOLDS et al. 2020).
Diese Studie vergleicht UAV-gestützte RGB- und Multispektralbild-Analysen im Hinblick auf für die Züchtung geeignete Merkmale wie Pflanzenhöhe (HEI), Vegetationsbedeckung (VCOV) und Ertragsvorhersagen unter Verwendung verschiedener Modellierungsansätze während der gesamten Vegetationsperiode. Das zeitliche Raster (wöchentlich) der Merkmalserfassung an drei Versuchsstandorten wurde für die Extraktion von wachstumsdynamischen Merkmalen genutzt. Mit Hilfe der bodengestützten Phänotypisierung war es möglich, die Genauigkeit der UAV-gestützten Phänotypisierung zu quantifizieren. Es wurden Wiederholbarkeits- und Genotyp-Assoziationsstudien durchgeführt, um die Eignung der Verwendung konventioneller und neuer Merkmale für die Gerstenzüchtung zu testen. Gerste wurde in dieser Studie als Versuchsorganismus ausgewählt, da sie als genetische und phänotypische Modellpflanze für gemäßigte Getreidearten wie Weizen, Dinkel, Roggen und Triticale dient und gleichzeitig eine hohe wirtschaftliche Bedeutung hat. Darüber hinaus konnten die in dieser Studie unter Verwendung der vielfältigen Gerstenkartierungspopulation S42IL (SCHMALENBACH et al. 2008; HONSDORF et al. 2017) gefundenen Ergebnisse als eine Art Validierung mit zahlreichen zuvor veröffentlichten Studien zu dieser Population verglichen werden. Dieser Artikel stellt die Kurzfassung der schon veröffentlichten Studie HERZIG et al. (2021) dar.

2 Material & Methoden

2.1 Pflanzenmaterial, Standort- und Anbaubedingungen

Das Pflanzenmaterial umfasste 49 Wildgersten-Introgressionslinien (ILs) der S42IL-Bibliothek und 11 Sommergersten-Sorten als Kontrollen (SCHMALENBACH et al. 2008). Die S42IL-Population wurde konventionell in mehreren Genotyp-Assoziationsstudien phänotypisiert, um signifikante Introgressionslinien für Pflanzenhöhe (SCHMALENBACH et al. 2009; WANG et al. 2010a; HOFFMANN et al. 2012), Trockenstress (HONSDORF et al. 2014b; HONSDORF 2015; HONSDORF et al. 2017) und Ertrag (ZAHN et al. 2020) zu identifizieren, die in dieser Studie vergleichend für die sensorgestützte Phänotypisierung verwendet werden können. Die Feldversuche im Jahr 2019 wurden an drei Standorten in Deutschland auf den Versuchsfeldern der MLU in Halle (51°29'43 "N 11°59'50 "E) und Merbitz (51°36'32 "N 11°53'43 "E) sowie auf dem Versuchsgelände des Instituts für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben (51°48'15 "N 11°15'02 "E) durchgeführt (Abb. 1). Die Feldversuche umfassten 360 Ertragsparzellen. Weitere Informationen sind in HERZIG et al. (2021) aufgeführt.

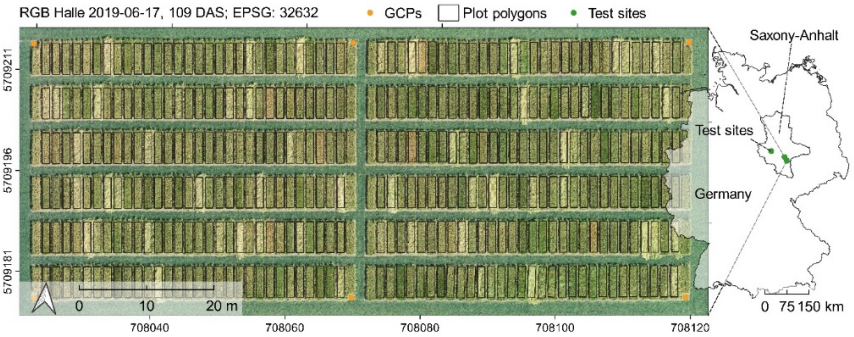


Abb. 1: Überblick über die drei Feldversuche im Jahr 2019 in Halle, Merbitz und Gatersleben

2.2 Bodenphänotypisierung

Die Phänotypisierung am Boden umfasste die in Tab. 1 beschriebenen agronomischen und entwicklungsrelevanten Merkmale, die in dieser Studie als Ground Truth (GT) verwendet wurden. Die Phänotypisierung wurde wöchentlich an allen Standorten parallel zu den UAV-Flügen durchgeführt.

Tab. 1: Liste und Beschreibung der untersuchten Merkmale

Merkmal	Abkürzung	Einheit	Detektionsinstrument	Messung
Schossen	SHO	Tage	Visuell	Anzahl der Tage von der Aussaat bis zum ersten Knoten, (BBCH 31)
Blüte	HEA	Tage	Visuell	Anzahl der Tage von der Aussaat bis zum Erscheinen der Grannenspitzen, (BBCH 49)
Reife	MAT	Tage	Visuell	Anzahl der Tage von der Aussaat bis zur Gelbreife, (BBCH 87)
Bestandshöhe	HEI_{GT} [a]/ HEI_{CHM} [b]/ HEI_{CHMred} [c]	cm	Visuell/ UAV (RGB)	Durchschnittliche Bestandshöhe aller Pflanzen einer Parzelle. UAV-RGB-Daten wurden verwendet, um digitale Höhenmodelle (DEM) zu erstellen, die zur Bestimmung von Wachstumsparametern führen (HEI_{GRi} [d], HEI_{GRd} [e], HEI_{MAX} [f])
Bedeckungsgrad	VCOV	%	UAV (RGB, Multispek)	Fläche einer Parzelle, die von Pflanzen bedeckt ist, die zur Bestimmung der Wachstumsparameter führen ($VCOV_{GRi}$ [g], $VCOV_{90}$ [h], $VCOV_{smoothed}$ [i])
Parzellenertrag	YLD	kg	Erntemaschine / UAV (RGB, Multispek)	Korngewicht bei der Ernte der gesamten Parzelle ($7.5m^2$)/ Modellierung auf der Grundlage von spektralen UAV-Daten, VCOV und HEI.

[a] Im Feld gemessene Bestandshöhe.
[b] Vom UAV-Pflanzenhöhenmodell abgeleitete Bestandshöhe.
[c] Vom UAV-Pflanzenhöhenmodell abgeleitete Höhe des Kronendachs, reduziert auf den HEIGT-Datensatz.
[d] Wachstumsrate der Wachstumsphase der zunehmenden Pflanzenhöhe.
[e] Wachstumsrate der Wachstumsphase der abnehmenden Pflanzenhöhe.
[f] Maximale Pflanzenhöhe aller gemessenen Zeitpunkte.
[g] Wachstumsrate der Wachstumsphase der zunehmenden Vegetationsbedeckung.
[h] Anzahl der Tage, an denen die Vegetationsbedeckung von 10 % auf 90 % zunahm.
[i] Geglättete Daten zur Vegetationsbedeckung.

2.3 UAV-Plattformen, Kamerasysteme und UAV-Kampagnen

Die wöchentlichen UAV-Flüge begannen Anfang April und endeten nach 11 Wochen in HAL und nach 12 Wochen in IPK und MER mit einem Flug kurz vor der Ernte. Es wurden zwei Kamerasysteme verwendet, die sich in Bezug auf Kosten, Gewicht, und Bilddaten unterscheiden.

Hochauflösende RGB-Bilder wurden mit einem Quadrocopter-System DJI Phantom 4 Professional (SZ DJI Technology Co. Ltd., Shenzhen, China) aufgenommen, dass standardmäßig mit einer Zenmuse X4S-Kamera (20-Megapixel-CMOS-Sensor (Complementary Metal Oxide Semiconductor)) ausgestattet war. Die Brennweite der Optik von 8,8 mm ergab eine Bodenauflösung von 0,82 cm/Pixel bei einer Flughöhe von 30 m und von 1,3 cm/Pixel bei einer Flughöhe von 50 m. Das Kreuzgitterflugschema bei Flughöhen von 30 m und 50 m wurde zur Erstellung eines detaillierten digitalen Höhenmodells (DEM) verwendet.

Die Multispektralbilder wurden mit der MACAW Multispektralkamera (Tetracam Inc., Chatsworth, CA, USA) aufgenommen, die auf der UAV-Plattform DJI Matrice 600 Pro (SZ DJI Technology Co. Ltd., Shenzhen, China) montiert war. Die Kamera besteht aus sechs unabhängigen CMOS-Bildsensoren mit einer Bildgröße von jeweils 1280 × 1024 Pixeln (1,3 MP) und einer radiometrischen Auflösung von 16 Bit. Die Brennweite der Optik von 9,6 mm ergab eine Bodenauflösung von 1,5 cm/Pixel bei einer Flughöhe von 30 m. In dieser Studie wurden fünf Filter mit einer Halbwertsbreite von 10 nm (FWHM) mit den Zentralwellenlängen 670, 700, 740, 780 und 900 sowie ein Filter mit einer FWHM von 40 nm und Zentralwellenlängen von 970 nm ausgewählt. Weitere Daten zur spezifischen Filtertransmission, zu den Sensorspezifikationen und zur spektralen Empfindlichkeit des CMOS sind in HERZIG et al. (2021) aufgeführt.

Zur Korrektur der gemessenen Reflexionswerte wurde die Globalstrahlung mit Hilfe eines incident light Sensors (ILS) mit derselben Filterkombination auf der Oberseite der Matrice 600 Pro (M600) angebracht, gemessen.

Die Georeferenzierung der UAV-Bilder erfolgte über sechs Bodenkontrollpunkte (GCPs), die gleichmäßig über die Feldversuche verteilt waren. Die Position der GCPs wurden mit einem Trimble R9s GNSS-Empfänger (Global Navigation Satellite System) (Trimble Ltd., USA) mit einer Genauigkeit von 0,02 m bestimmt.

2.4 UAV-Datenverarbeitung

Der gesamten Bearbeitungs-Pipeline ist in Abb. 2 dargestellt. Mit Ausnahme der anfänglichen Vorverarbeitung und der photogrammetrischen Verarbeitung wurden alle Arbeiten an den UAV-Daten in R (R CORE TEAM 2020) durchgeführt. Die verwendeten Spezialpakete werden in den entsprechenden Abschnitten erwähnt.

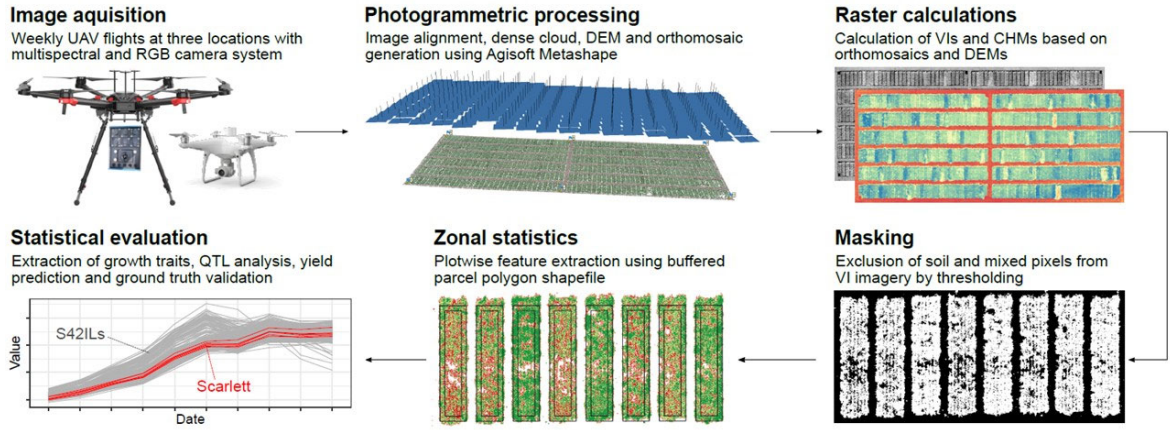


Abb.2: Schematische Darstellung der Bildaufnahme, Bildverarbeitung und Datenanalyse (HERZIG et al. 2021)

2.4.1 Vorverarbeitung der Multispektral-Daten

Die Vorverarbeitung der MACAW-Bilder wurde mit dem Bildbearbeitungsprogramm Pixel Wrench 2 (Tetracam Inc.; Chatsworth, CA, USA) durchgeführt und umfasste im Wesentlichen die Versatz-Korrektur (Drehung und Skalierung der Bilder des jeweiligen Sensors) und die Korrektur auf die einfallende Strahlung. In Anbetracht der geringen Flughöhe der UAVs und des daher angenommenen begrenzten atmosphärischen Einflusses wurde keine zusätzliche atmosphärische Korrektur durchgeführt.

2.4.2 Photogrammetrische Verarbeitung

Die photogrammetrische Datenverarbeitung der gewonnenen UAV-Bilder erfolgte in Agisoft Metashape Professional (Version 1.5.2.7838, Agisoft LLC, St. Petersburg, Russland). Der Arbeitsablauf zur Erstellung des Orthomosaiks bestand aus vier Hauptverarbeitungsschritten: Bildausrichtung, Erstellung einer dichten Punktwolke, Erstellung eines digitalen Höhenmodells (DEM) und abschließende Erstellung eines Orthomosaiks. Die detaillierten Einstellungen für die Verarbeitung der Datensätze sind in HERZIG et al. (2021) zusammengefasst.

2.4.3 Berechnung des Pflanzenhöhenmodells (CHM) und des Vegetationsindex (VI)

Die Pflanzenhöhenmodelle (CHM) wurden aus hochauflösenden RGB-DEMs abgeleitet. Dabei diente die erste kalkulierte DEMs (Aufnahme direkt nach der Aussaat) als T_0-Modell von dem die Pixelwerten der nachfolgenden DEMs subtrahiert wurden, um die absolute HEI_{CHM} aller Parzellen zu erhalten. HEI_{CHMred} stellt eine Teilmenge von HEI_{CHM} dar, die auf die 107 Parzellen von HEI_{GT} reduziert wurde. 30 VIs wurden berechnet und im Hinblick auf ihre Sensitivität in folgende Gruppen unterteilt: Einzelband, Pigment, Wassergehalt und Physiologie. Die Liste der VIs findet sich in HERZIG et al. (2021). Alle Rasterberechnungen wurden mit dem Rasterpaket von HIJMANS (2020) durchgeführt.

2.4.4 Maskierung des Bodens

Um eine Boden-Pflanzensegmentierung gerade in den ersten Entwicklungsstadien der Pflanzen zu vorzunehmen, wurde die Otsu-Methode auf kontrastreiche NDVI- und EG-Bilder angewendet. Die Otsu-Methode ist ein adaptiver, nichtparametrischer und daher nicht überprüfter Schwellenwertalgorithmus für die Bildsegmentierung auf der Grundlage von Graustufenhistogrammen (OTSU 1979), der sich als geeignet für die Trennung von Pflanzen- und Bodenpixeln erwiesen hat (MARCIAL-PABLO et al. 2018). Die Methode wurde mit der Funktion auto_thresh des Pakets von LANDINI et al. (2017) implementiert. Die resultierenden Masken wurden dann auf die entsprechenden Spektralbänder und VIs derselben Datensatzes angewendet.

2.4.5 Bestimmung der Vegetationsbedeckung (VCOV)

Für die Bodenmaskierung zur Bestimmung der VCOV wurde nicht die Methode von Qtsu gewählt, da dieses Verfahren einen hohen Anteil an Misch-Pixel (Boden-Pflanze) zulässt, die zwar für die Erfassung der spektralen Merkmale, nicht aber für die Schätzung des relativen Vegetationsanteils erwünscht sind. Daher war eine strengere Diskriminierung erforderlich. Nur Pixel der RGB-Bilder, die über einem festen GCC-VI-Schwellenwert von 0,38 lagen, wurden als Pflanzenpixel behandelt. Im Gegensatz dazu wurden die multispektralen Bilder auf Basis des k-means Algorithmus mit vier Clusterzentren auf die NDVI-Bilder angewendet und maskiert, was zu einem dynamischeren Schwellenwert führte. Nur Pixel, die zum höchsten NDVI-Cluster gehören, wurden als Pflanzenpixel behandelt.

2.4.6 Parzellenweise Merkmalsextraktion

Die parzellenweise Merkmalsextraktion wurden mit Hilfe des RSAGA-Pakets von BRENNING (2008) realisiert, das den Zugang zu den Geoberechnungs- und Geländeanalysefunktionen des geographischen Informationssystems ermöglicht. Es wurde die SAGA-Version 6.2.0 verwendet. Die Shapefiles der Feldversuche wurden mit der Software MiniGIS von Geo-Konzept (Adelschlag, Deutschland) erstellt. Nach Anwendung eines Puffers von -0,2 m auf jedes Parzellenpolygon-Shapefile, um Randeffekte zu berücksichtigen, wurden zusammenfassende Statistiken wie das arithmetische Mittel und das 95 %-Quantil der Parzellen extrahiert.

2.5 Statistische Analyse

Die weitere Datenanalyse erfolgte anhand des 95 %-Quantils des Parzellen Polygons der HEI_{CHM}-Orthomosaike und des Mittelwerts der VI- und VCOV-Orthomosaike.

2.5.1 Validierung durch Ground Truth (GT) Daten

Zur Bewertung der Genauigkeit der durch die UAV-Bilder ermittelten Bestandshöhe (HEI_{CHM}) wurden der Korrelationskoeffizient (r) von Pearson und der mittlere quadratische Fehler (RMSE) von HEI_{CHM} und HEI_{GT} für jeden Zeitpunkt und über den gesamten Vegetationszeitraum berechnet (Formel 2).

2.5.2 Modellierung der Wachstumsrate

Durch multitemporale Messungen wurde die Berechnung der Wachstumsraten von HEI- und VCOV-Daten durch Anwendung der im R-Paket growthPheno (BRIEN et al. 2020) ermöglicht.

Dieses Verfahren ermöglicht die Glättung von Wachstumskurven und anschließend die Extraktion von Merkmalen (SET). Details sind in HERZIG et al. (2021) zu finden.

2.5.3 Ertragsvorhersage

Unter Verwendung der spektralen Merkmale (Tab. 2) sowie die räumlichen Merkmale HEI und VCOV der beiden vorgestellten bildgebenden Systeme wurde die Vorhersagegenauigkeit des Parzellenertrags bewerten. Dabei wurden beide Systeme unabhängig voneinander betrachtet, um ihre Eignung zur Vorhersage zu vergleichen. Da HEI_{CHM} nur aus RGB-Bildern abgeleitet werden konnte, wurde keine Bestandshöhen bei der multispektralen Ertragsvorhersage verwendet. Die Modellierung des Parzellenertrags umfasste drei Ansätze:

Ansatz 1: Anwendung eines einzigen linearen Regressionsmodells für jedes Merkmal zu einem einzigen Zeitpunkt

Ansatz 2: Einbeziehung aller gemessenen und abgeleiteten Merkmale desselben Zeitpunkts als Prädiktoren in eine Mehrfachregression

Ansatz 3: Erweiterung des Mehrfachregressionsmodells, das durch eine gestapelte Datenbasis aller vorrangegangenen Messtermin zur Vorhersage führt

Für die Mehrfachregression in den Ansätzen 2 und 3 wurden ein verallgemeinertes lineares Modell (GLM) mit elastischem Netz und ein Random-Forest-Modell angepasst. Beide wurden mit Hilfe des ‚Tidymodels'- Package (KUHN & WICKHAM 2020) implementiert. Der Algorithmus für das regularisierte GLM mit elastischem Netz kombiniert linear L1- (Lasso) und L2-Regularisierungsstrafen (Ridge) in einem mehrfachen linearen Modell.
Der Random-Forest-Algorithmus ist eine robuste Lernmethode zur Klassifizierung und Regression (BREIMANN 2001). Die beiden Hyperparameter ntree und mtry wurden konstant bei ntree = 1000 und mtry = npredictors/3 gehalten.
Die Zusammenhänge zwischen einzelnen Merkmalen und Erträgen wurden durch Berechnung des Bestimmtheitsmaßes (r^2) bewertet. Bei der multiplen linearen Regression wurde auch der RMSE erfasst. Das r^2 und der RMSE wurden wie folgt berechnet:

(1) $$r^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

und

(2) $$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}}$$

wobei x_i der tatsächliche Parzellenertrag, $\bar{x}$ der durchschnittliche tatsächliche Parzellenertrag, y_i der vorhergesagte Parzellenertrag und n die Anzahl der Beobachtungen ist. Mehrfache Regressionsparameter wurden extrahiert und über eine zehnfache kreuzvalidierte Modellanpassung gemittelt.
Die Schätzungen der GLM-Variablen wurden aus der besten kreuzvalidierten Anpassung extrahiert und ihre entsprechenden Prozentwerte wurden als Bruchteil der gesamten absoluten Schätzungssumme berechnet.

Die für die Genotyp-Assoziationsstudie verwendeten vorhergesagten Erträge wurden über eine zehnmal wiederholte zehnfache GLM-Kreuzvalidierung ermittelt, wobei die Vorhersagen der unabhängigen Validierungsuntergruppe beibehalten und für jede Parzelle gemittelt wurden.

2.5.4 Genotyp-Assoziationsstudie

Vor der Genotyp-Assoziationsstudie wurde die Wiederholbarkeit (R^2) als Maß für den Anteil der genotypischen Variation an der gesamten phänotypischen Varianz mithilfe des lme4-Pakets (BATES et al. 2018) für jedes Merkmal zu jedem Zeitpunkt wie folgt berechnet:

$$(3)\quad R^2 = \frac{V_G}{V_P/r}$$

wobei V_G und V_P die genotypische und die phänotypische Gesamtvarianz sind und r die Anzahl der Wiederholungen (r = 6 in dieser Studie) pro Genotyp ist. Die Genotyp-Assoziationsstudie wurde dann als Dunnett-Test (DUNNETT 1955) durchgeführt, wobei das durch die UAV ermittelte Merkmal jedes Genotyps mit der Kontrollsorte Scarlett unter Verwendung des glht-Pakets für allgemeine lineare Hypothesen (HOTHORN et al. 2008) miteinander verglichen wurde. Zusätzlich wurde die Bonferroni-Anpassung auf die Modellzusammenfassung angewendet und die Signifikanzschwelle auf $p < 0{,}01$ festgelegt.

3 Ergebnisse

3.1 Bestimmung der Bestandshöhe

Die Korrelationen von HEI_{CHM} und HEI_{GT} über alle Zeitpunkte hinweg sind in Abb. 3 dargestellt. In den drei Umgebungen HAL, IPK und MER wurden Korrelationen von 0,82, 0,89 und 0,91 erreicht, was zu einer Gesamtkorrelation von r = 0,87 unter Verwendung des 95 %igen CHM-Plot-Quantils führt. Trotz der in erster Linie starken Korrelationen wurde HEI_{GT} von der UAV im Allgemeinen unterschätzt. Eine Detaillierte Auswertung mit dem Bezug zu einzelnen Messterminen wird in HERZIG et al. (2021) gegeben.

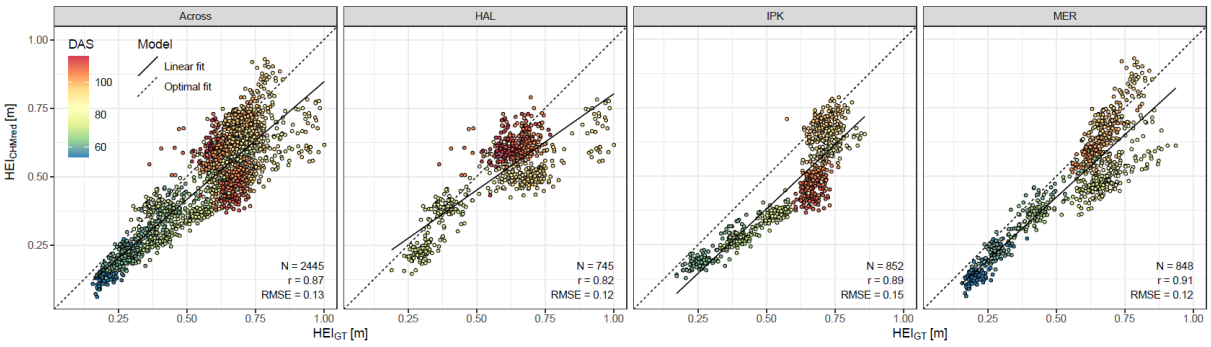


Abb.3: Pearson-Korrelation der Bestandshöhenmessungen im Feld (HEI_{GT}) und photogrammetrischer UAV-Pflanzenhöhenbestimmung (HEI_{CHMred}) innerhalb und über die Umgebung hinweg. Die Farben geben den Zeitpunkt der Messung in Tagen nach der Aussaat (DAS) an (HERZIG et al. 2021)

Die Wiederholbarkeiten von HEI_{CHM}, HEI_{CHMred} und HEI_{GT} zeigten hohe Werte zwischen den einzelnen Zeitpunkten von bis zu $R^2 = 0,98$ fast über die gesamte Wachstumsperiode (HERZIG et al. 2021). Vor dem SHO war die Wiederholbarkeit gering, begann jedoch mit dem Höhenwachstum um SHO herum zuzunehmen, um ein Maximum und ein Plateau beim HEA zu erreichen, bis zum Zeitpunkt der Reife, wo die Wiederholbarkeit wieder abnahm.
Die im Dunnett-Test ermittelten größten signifikanten Unterschiede wurden im Zeitraum zwischen SHO und HEA gefunden, wobei die Bestandshöhe zu dieser Zeit auch die größten Unterschiede aufwies. Im Allgemeinen erreichte der HEI_{CHM}-Datensatz die höchste Signifikanz, und mehrere Linien wurden über den Messzeitraum hinweg konstant festgestellt. Im Vergleich dazu zeigten die reduzierten HEI_{GT}- und HEI_{CHMred}-Datensätze weniger signifikante Ergebnisse, was den großen Vorteil der UAV-Phänotypisierung zeigt, die einen größeren Stichprobenumfang abdecken kann.

3.2 Bestimmung der Vegetationsbedeckung

Die Schätzung der VCOV durch die RGB-Bilder ergab eine höhere Kontinuität zwischen den aufeinanderfolgenden Flügen mit einer späteren Sättigung und einer geringeren Rate an Ausreißern, was einen zuverlässigeren Datensatz im Vergleich zu den mit multispektralen Bildern ermittelten VCOV-Werten ergab. Die VCOV-Werte des multispektralen Datensatzes schienen aufgrund der größeren Streuung und Ausreißerquote eine schlechtere Qualität aufzuweisen. In MER wiesen die VCOV-Werte die größten Schwankungen zwischen den beiden Bestimmungsmethoden auf. Dennoch lässt sich über alle Datensätze hinweg feststellen, dass der Populationsdurchschnitt der S42ILs leicht über der Leistung von Scarlett liegt.

3.3 Ertragsvorhersage

Der gemessene Parzellenertrag war der Zielwert für die Ertragsmodellierung. Die höchsten Erträge wurden am IPK mit einem Durchschnittsertrag von 4,43 kg/7,5m^2 (≙59,1 dt/ha) gefunden, gefolgt von MER mit 4,01 kg/7,5m^2 (≙53,5 dt/ha) und HAL mit 3,73 kg/7,5m^2 (≙49,7 dt/ha). HAL zeigte den höchsten Variationskoeffizienten mit CV=0,27, gefolgt von MER mit CV = 0,16 und IPK mit CV = 0,12. Im Folgenden werden nur die Ergebnisse der GLM-Modelle beschrieben; die Ergebnisse der entsprechenden RF-Modelle, die im Allgemeinen die gleichen Trends widerspiegeln, sind in HERZIG et al. (2021) enthalten.
Die Modelle aus Ansatz 1, die auf einem einzigen Merkmal zu einem Zeitpunkt reduziert wurden, zeigte die lineare Regression starke Unterschiede in der Vorhersageleistung zwischen den Modellen (Abb. 4a). Für Abb. 4a wurden die beiden VIs mit den höchsten r^2-Werten pro Umgebung und Bilddatensatz ausgewählt und über alle Umgebungen hinweg aufgetragen, was zu vier bzw. fünf VIs für die multispektralen und RGB-Bilddatensätze führte. Die höchsten r^2-Werte des RGB-Datensatzes wurden in allen drei Umgebungen kurz vor MAT mit $r^2 = 0,55$, $r^2 = 0,37$ und $r^2 = 0,21$ in HAL, MER und IPK ermittelt (Abb. 4a). Im multispektralen Datensatz wurden sowohl um HEA als auch kurz vor MAT hohe r^2-Werte festgestellt, die in der gleichen Größenordnung lagen wie im RGB-Datensatz, mit Ausnahme der hohen Vorhersagekraft der multispektralen Daten bei 47 DAS in MER.
Bei Ansatz 2 wurden für die Ertragsvorhersagemodelle die Summe aller Merkmale aus dem RGB- oder Multispektraldaten inklusive der vorhandenen morphologischen Merkmale (HEI und VCOV)

zu einem jeweiligen Zeitpunkt gebildet und berücksichtigt. Die Vorhersagekraft nahm zu, unterlag aber immer noch großen Schwankungen zwischen den Zeitpunkten (Abb. 4b). Die Kurven der Vorhersagegenauigkeit zwischen dem RGB- und dem multispektralen Datensatz zeigten ähnliche Schwankungen in einem Bereich zwischen r^2=0,03 und r^2=0,6. Die höchste Vorhersagegenauigkeit wurde an allen drei Standorten im letzten Drittel der Vegetationsperiode mit Werten von r^2 = 0,6, r^2 = 0,5 und r^2 = 0,29 in den RGB-Bildern in HAL, MER bzw. IPK (Abb. 4b) erreicht.

Ansatz 3 bestand darin, alle Merkmale wie in Ansatz 2 zu verwenden, welche sich zusätzlich mit jedem voranschreitenden Messtag addierten, sodass alle zuvor ermittelten Merkmale in die Modellierung des aktuellen Modells miteinflossen. Bei diesem Ansatz nahm die Vorhersagegenauigkeit mit jedem zusätzlichen Messtag zu und übertraf die Vorhersagegenauigkeit am Ende der Vegetationsperiode im Vergleich zum zweiten Ansatz deutlich (Abb. 5). Die Vorhersagegenauigkeit aller Modelle zur Ertragsschätzung zeigte im Allgemeinen einen asymptotischen Trend. Die Trends in der Vorhersagegenauigkeit zwischen den Standorten und den RGB- und Multispektraldatensätzen wurden von Ansatz 2 übernommen, allerdings auf einem höheren Niveau. Hohe r^2-Werte wurden ebenfalls vor MAT festgestellt und erreichten den höchsten Wert in HAL mit r^2 = 0,82, gefolgt von r^2 = 0,65 in MER und r^2 = 0,55 in IPK. Dennoch hatten die hohen r^2-Werte in HAL einen größeren RMSE als in MER oder IPK. Der RMSE der Regressionsmodelle des multispektralen Datensatzes lag zwischen 0,24 und 0,39 kg/Plot (≙ 6,4-10,4%) in HAL, zwischen 0,27 und 0,35 kg/Plot (≙ 6,7-8,7%) in MER und zwischen 0,26 und 0,33 kg/Plot (≙ 5,9-7,5%) in IPK. Der RMSE der RGB-basierten Regressionsmodelle lag in HAL zwischen 0,23 und 0,45 kg/Plot (≙ 6,1-12,1%), in MER zwischen 0,24 und 0,32 kg/Plot (≙ 6,1-8,0%) und in IPK zwischen 0,23 und 0,33 kg/Plot (≙ 5,2-7,6%). Obwohl IPK den niedrigsten r^2-Wert aufwies, ist der absolute Schätzfehler am geringsten, weil der CV des gemessenen Ertrags am kleinsten ist.

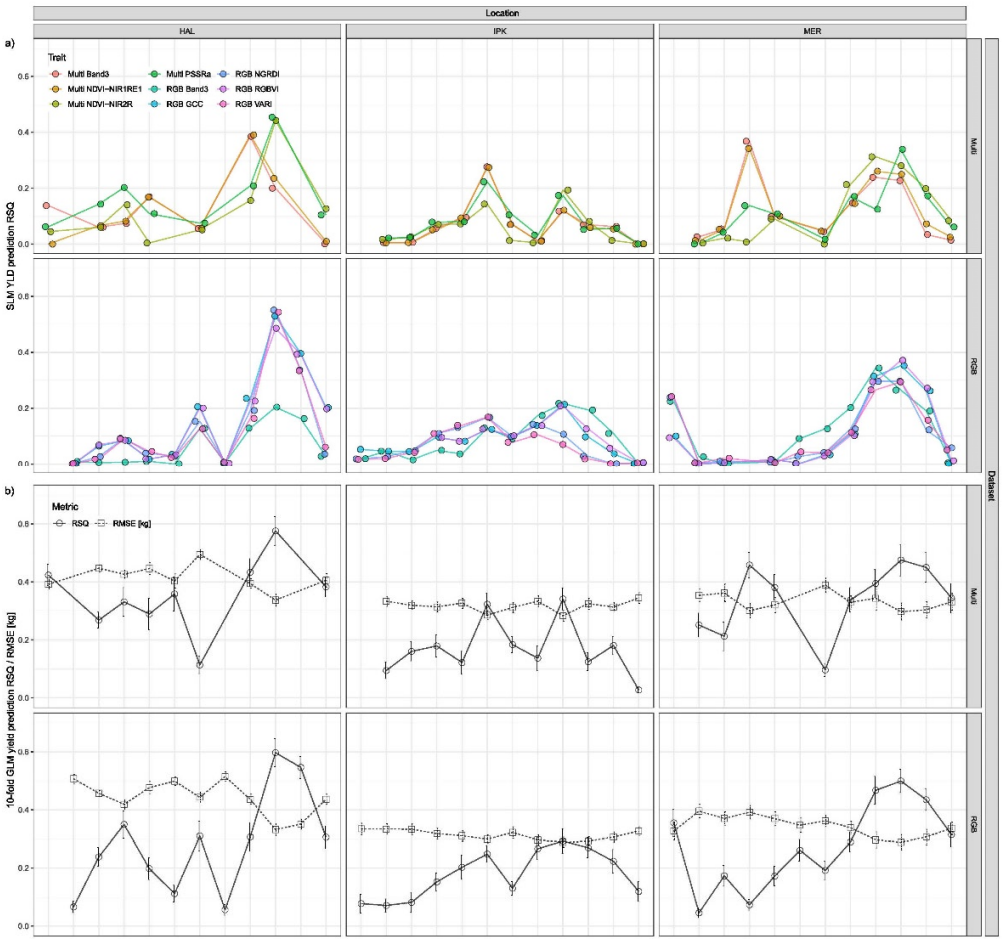


Abb.4: Vergleich der Ertragsvorhersagemodelle. (a) Vorhersagegenauigkeit eines einzelnen Merkmals mit linearer Regression (SLM). (b) Ertragsvorhersageleistung der GLM-Modelle, basierend auf den täglich zusammengeführten Datensätzen. RSQ (r2) und RMSE (root mean square error) wurden als Qualitätskriterium für die Vorhersagegenauigkeit verwendet

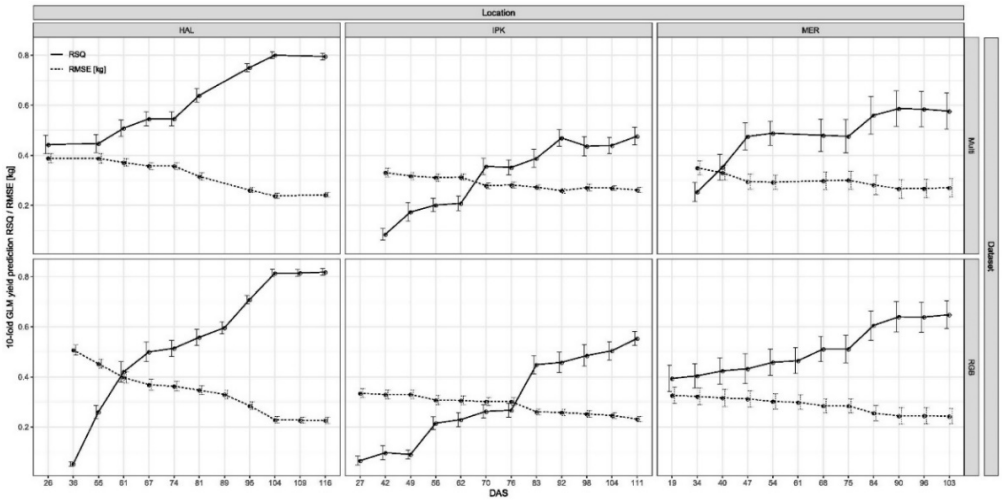


Abb. 5: Ergebnisse der Ertragsvorhersage des GLM-Modells, basierend auf Ansatz 3

4 Diskussion

4.1 Bestimmung der Bestandshöhe

Die Korrelationen von HEI_{CHM} und HEI_{GT} lagen geringfügig unter den in ähnlichen Studien (BENDIG et al. 2014; BARETH et al. 2016) beobachteten Korrelationen, was darauf zurückzuführen sein könnte, dass die Pflanzenhöhe beim Einsatz von Pflanzenwachstumsregulatoren eine relativ geringe Variation aufwies. Ohne den Einsatz von Pflanzenwachstumsregulatoren, wie es bei HAL der Fall war, litten die Parzellen unter Lagerbildung und die Bestimmung der Höhe war sowohl mit der Drohne als auch mit dem Lineal schwierig, was zu einer ungenauen Messung führte. Zudem können die HEI_{GT}-Messungen im Feld aufgrund der Subjektivität der Messung nicht als allgemeiner Standard für die Bestimmung der Kronenhöhe angesehen werden (PASK et al. 2012).
Die Unterschätzung der Pflanzenhöhe durch die Bestimmung des SfM-Algorithmus ist ein bereits bekanntes Problem (GRENZDÖRFFER 2014). Dieser Effekt tritt eher mit zunehmender Bestandshöhe und unter windigen Bedingungen auf. Darüber hinaus ist der Bestand in der Phase der Triebverlängerung zwischen SHO und HEA sehr dünn, was den größten Effekt auf die Unterschätzungen des HEICHM haben dürfte. Dies deutet darauf hin, dass ein dünnerer Pflanzenbestand die Wahrscheinlichkeit verringert, Tie-Points an der Spitze des Bestandes zu finden. Es ist wahrscheinlicher, dass Tie-Points in einer niedrigeren Ebene gefunden werden können, wo die Blätter ein dichteres Blätterdach bilden.
Beim Vergleich der Wiederholbarkeiten zwischen HEI_{CHM}, HEI_{GT} und HEI_{CHMred} zeigte HEI_{CHM} die höchsten Werte, was darauf hindeutet, dass es möglich ist, die Bestandshöhe zu einzelnen Zeitpunkten mit hoher Genauigkeit mittels UAV zu bestimmen. Die absolute HEI-Bestimmung hingegen ist mit verschiedenen Fehlerquellen behaftet und daher nur mit geringerer Genauigkeit möglich. Darüber hinaus deutet die höhere Wiederholbarkeit von HEI_{CHM} im Vergleich zu HEI_{CHMred} darauf hin, dass ein größerer Stichprobenumfang die statistische Aussagekraft erhöht, die durch den Einsatz von UAVs zur Bestimmung der Bestandshöhe in der Praxis erreicht werden kann.
Die Ergebnisse des Dunnett-Tests zeigen eine starke Überlappung der gefundenen signifikanten Linien zwischen HEI_{CHM}, HEI_{GT} und HEI_{CHMred}, wobei die meisten dieser Linien bereits aus früheren Veröffentlichungen bekannt waren (SCHMALENBACH et al. 2009; WANG et al. 2010b; HOFFMANN et al. 2012; HONSDORF et al. 2014b; HONSDORF et al. 2014a).

4.2 Vegetationsbedeckung

Eine frühe Bestandsetablierung wirkt sehr kompetitiv und in der Lage Beikräuter zu verdrängen, was für den ökologischen Landbau von besonderem Interesse ist. Dieses Merkmal korreliert in hohem Maße mit dem physiologischen Zustand der Pflanze, wie z. B. dem Blattflächenindex, der Pflanzenvitalität, der Biomasse, dem Ertrag und der Reaktion auf erhöhte Temperaturen (LIU & PATTEY 2010; GRIEDER et al. 2015; LIEBISCH et al. 2015; BALLESTEROS et al. 2018; KIM et al. 2019; GARCÍA-MARTÍNEZ et al. 2020).
In dieser Studie scheinen die Ergebnisse der RGB-ermittelten VCOV zuverlässiger zu sein, was teilweise durch eine höhere Bildauflösung erklärt werden kann, die mit der geringeren Flughöhe, aber auch mit einer besseren räumlichen Auflösung des RGB-Kamerasensors zusammenhängt. Darüber hinaus ist die schlechtere Qualität der multispektral ermittelten VCOV möglicherweise

darauf zurückzuführen, dass einzelne Messtermine entfernt werden mussten, weil die Multispektralkamera zu Beginn der Saison in MER und IPK ausfiel.
Die Ergebnisse des anschließend durchgeführten Dunnett-Tests bestätigten, dass die Bestimmung von VCOV-Merkmalen durch die Verwendung von an Drohnen montierten RGB-Kameras möglich war und dass die Genauigkeit der Ergebnisse für weitere quantitative genetische Analysen ausreichend ist. Allerdings waren die hier vorgestellten Ergebnisse durch die Anzahl der Beobachtungen zum Zeitpunkt der Merkmalszunahme stark eingeschränkt, denn um die Wachstumsdynamik aus Bilddaten zu extrahieren, ist eine zeitlich hochauflösende Datenerfassung entscheidend (AASEN et al. 2020). Die wöchentliche Überwachung der Merkmale in der Umgebung ergab nur 2-3 relevante Beobachtungen, die für die Wachstumsmodellierung verwendet werden konnten, bis eine Sättigung erreicht war, was nicht ausreichte, um die kleinen Merkmalsvariationen innerhalb der Population zu erkennen. Dennoch zeigen die Ergebnisse das Potenzial dieser Methode zur Erfassung der VCOV-Merkmale, insbesondere wenn die Überwachungsintervalle auf alle 2-3 Tage erhöht werden. Wenn nur die VCOV-Merkmale von Interesse sind, könnte eine kommerzielle RGB-Kamera mit ihrer einfacheren Datenverarbeitung und konsistenteren Ergebnissen eine bessere Wahl sein.

4.3 Ertragsvorhersage

Der Variationskoeffizient des im Feld bestimmten Ertrags war in HAL am höchsten und in IPK am niedrigsten, und auch die Ertragsvorhersage von Ansatz 1 zeigte die höchste Vorhersagekraft in HAL und die niedrigste in IPK (Abb. 4a), was darauf hindeutet, dass eine größere Variation des vorhergesagten Merkmals höhere r^2-Werte ermöglichen kann. Ähnliche Korrelationen wurden auch in anderen Studien zur Ertragsmodellierung festgestellt (OEHLSCHLAGER et al. 2018; GARCÍA-MARTÍNEZ et al. 2020; TAO et al. 2020). Ansatz 2, der aus einer Datenfusion mehrerer VIs, HEI- und VCOV-Informationen bestand, führte zu besseren Ertragsvorhersagemodellen in den multispektralen und auch RGB-Datensätze, was im Einklang mit den Ergebnissen weiterer Studien (GEIPEL et al. 2014; BENDIG et al. 2015; RISCHBECK et al. 2016) steht. Ansatz 3 zeigte die beste Vorhersagegenauigkeit mit der höchsten Genauigkeit in HAL, was mit einer höheren Variation der untersuchten Merkmale zusammenhängen könnte. Zum Beispiel war HAL die einzige Umgebung, in der keine Wachstumsregulatoren eingesetzt wurden, was zu einer größeren Merkmalsvariation in der Bestandshöhe und dem Zielmerkmal Parzellenertrag führte. In HERZIG et al. (2019) wurde ebenfalls festgestellt, dass die Variation des Zielmerkmals die Vorhersagegenauigkeit von Regressionsmodellen beeinflusst. Eine zunehmende Vorhersagegenauigkeit des Ertrags über die Vegetationsperiode hinweg konnte in einer Reihe von Studien ebenfalls beobachtet (YUE et al. 2017; HASSAN et al. 2019; TAO et al. 2020). Eine vertiefende Diskussion über die Ertragsmodelle und dessen Auswahl von Merkmalsgruppen und dessen Effektgrößen folgt in HERZIG et al. (2021).
Die anschließend durchgeführte Genotyp-Assoziationsstudie mit den Daten der Ertragsvorhersage konnte nur bei IPK signifikante Linien detektieren. Sechs der neun in den YLD_{GT}-Daten entdeckten signifikanten Linien konnten auch in ZAHN et al. (2020) gefunden werden. Die Überlappung der detektierten signifikanten Linien zwischen YLD_{GT} und dem multispektralen Datensatz nimmt mit höherem r^2 zu und falsch-positive Linien wurden nur zwischen dem Beginn

und der mittleren Phase der Vegetationsperiode entdeckt, was darauf hindeutet, dass eine zunehmende Vorhersagegenauigkeit falsch-positive Entdeckungen verringert. Die Ergebnisse der Assoziationsstudie, insbesondere die der Modelle am Ende der Vegetationsperiode mit weniger falsch-positiven Linien, können für die Pflanzenzüchtung genutzt werden. Eine robuste und verbesserte Vorhersageleistung über mehrere Versuchsjahre und Umgebungen hinweg könnte genauere Assoziationsstudie mit weniger falsch-negativen Ergebnissen ermöglichen.

5 Schlussfolgerungen

Es wurde prognostiziert, dass die Fernerkundung mittels UAV ein wichtiges neues Instrument zur Unterstützung von Züchtern und Landwirten in der Präzisionslandwirtschaft sein wird, aber die vielversprechenden UAV-Anwendungen werden bisher nur langsam angenommen (FREEMAN & FREELAND 2015). Daher wurde in dieser Studie ein einfacher praktischer Ansatz zur objektiven und präzisen Phänotypisierung traditioneller und neuer Merkmale mit hohem Durchsatz verwendet, der auf dem methodischen Vergleich von zwei unterschiedlichen Kamerasystemen ausgestatteten UAVs beruht. Die hochauflösenden RGB-Bilder konnten, mit Hilfe des SfM-Algorithmus, dazu verwendet werden, um die relative Bestandshöhe mit hoher Präzision über die gesamte Vegetationsperiode zu extrahieren. Die Ergebnisse der Genotyp-Assoziationsstudie mit der hohen Überlappung signifikanter Linien von UAV-geschätzter Höhe und im Feld gemessener Höhe zeigen den praktischen Nutzen für die Pflanzenzüchtung. Die Bestimmung der Vegetationsbedeckung zu Beginn der Vegetationsperiode ist nur mit objektiver Sensortechnik möglich. Daher wurden RGB- und Multispektralbilder verwendet, um den prozentualen Anteil der Pflanzendecke innerhalb einer Parzelle zu ermitteln. Die Ergebnisse des RGB-Datensatzes zeigten eine präzisere und konsistentere Extraktion der VCOV über mehrere aufeinanderfolgende Flüge hinweg, so dass dieser für das Merkmal VOCV geeigneter ist als der multispektrale Datensatz.
Darüber hinaus ermöglicht die nicht-invasive UAV-Phänotypisierung Wachstumsanalysen und damit die Identifizierung von Genotyp-Assoziationen, was für die Merkmale Bestandshöhe und Vegetationsgrad erfolgreich durchgeführt wurde. Für diese Art der Anwendung ist es von entscheidender Bedeutung, das richtige Zeitintervall für die Phänotypisierung je nach Merkmal zu bestimmen. Das Phänotypisierungsintervall von 7 Tagen für das Merkmal Vegetationsbedeckung erwies sich als zu groß. Nichtsdestotrotz waren die Vegetationsbedeckung und die modellierte Bestandshöhe für die Ertragsprognose von Bedeutung. Die Fusion von spektralen Vegetationsindizes und räumlichen Merkmalen lieferte ergänzende Informationen für die Ertragsschätzung, was in zahlreichen Studien bestätigt wurde.
Insbesondere die auf den RGB-Bildern basierenden Ertragsmodelle konnten durch die räumlichen Merkmale verbessert werden, was dazu führte, dass die Vorhersageleistung beider Kamerasysteme auf dem gleichen Niveau lag.
Die Modellierung auf der Grundlage einzelner Zeitpunkte führte zu großen Unterschieden in der Vorhersageleistung über die Vegetationsperiode, während die Akkumulation von Zeitpunkten eine stetige Verbesserung der Vorhersagemodelle zeigte. Im Gegensatz dazu führte die Fusion von RGB- und Multispektraldatensätzen nicht zu einer Verbesserung der Ertragsvorhersage, was auf Redundanz in den Datensätzen hindeutet. Die Zusammenführung der Datensätze über

verschiedene Umgebungen hinweg führte zu einem allgemeinen Vorhersagemodell mit durchschnittlicher Genauigkeit für die drei Umgebungen. Um das volle Potenzial von UAV-Anwendungen auszuschöpfen und die Akzeptanz zu erhöhen, ist die Übertragbarkeit von Modellen auf unterschiedliche Jahre, Populationen und Standorte erforderlich. Um diese Qualität in der UAV-Fernerkundung zu erreichen, ist ein umfassendes und standardisiertes Datenerfassungs- und Modellierungsverfahren nötig, um allgemeine und genaue Ertragsvorhersagemodelle zu erstellen.

6 Literaturverzeichnis

ASEN, H., KIRCHGESSNER, N., WALTER, A. & LIEBISCH, F., 2020: PhenoCams for Field Phenotyping: Using Very High Temporal Resolution Digital Repeated Photography to Investigate Interactions of Growth, Phenology, and Harvest Traits. Front Plant Sci, **11**: 593.

ACRECHE, M. M., BRICENO-FELIX, G., SANCHEZ, J. A. M. & SLAFER, G. A., 2008: Physiological bases of genetic gains in Mediterranean bread wheat yield in Spain. Eur J Agron, **28**: 162-170.

ARAUS, J. L. & CAIRNS, J. E., 2014: Field high-throughput phenotyping: the new crop breeding frontier. Trends in plant science, **19**: 52-61.

BABAR, M. A., REYNOLDS, M. P., VAN GINKEL, M., KLATT, A. R., RAUN, W. R. & STONE, M. L., 2006: Spectral reflectance indices as a potential indirect selection criteria for wheat yield under irrigation. Crop Sci, **46**: 578-588.

BALLESTEROS, R., ORTEGA, J. F., HERNANDEZ, D. & MORENO, M. A., 2018: Onion biomass monitoring using UAV-based RGB imaging. Precis Agric, **19**: 840-857.

BARETH, G., BENDIG, J., TILLY, N., HOFFMEISTER, D., AASEN, H. & BOLTEN, A., 2016: A Comparison of UAV- and TLS-derived Plant Height for Crop Monitoring: Using Polygon Grids for the Analysis of Crop Surface Models (CSMs). Photogramm Fernerkun: 85-94.

BATES, D., KLIEGL, R., VASISHTH, S. & BAAYEN, H., 2018: Parsimonious Mixed Models. arXiv preprint arXiv:1506.04967.

BENDIG, J., BOLTEN, A., BENNERTZ, S., BROSCHEIT, J., EICHFUSS, S. & BARETH, G., 2014: Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. Remote Sensing, **6**: 10395-10412.

BENDIG, J., YU, K., AASEN, H., BOLTEN, A., BENNERTZ, S., BROSCHEIT, J., GNYP, M. L. & BARETH, G., 2015: Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int J Appl Earth Obs, **39**: 79-87.

BREIMANN, L., 2001: Random Forests. Machine Learning, 45.

BRENNING, A., 2008: Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, **19**(23-32), 410

BRIEN, C., JEWELL, N., WATTS-WILLIAMS, S. J., GARNETT, T. & BERGER, B., 2020: Smoothing and extraction of traits in the growth analysis of noninvasive phenotypic data. Plant Methods, **16**: 36.

DUNNETT, C. W., 1955: A Multiple Comparison Procedure for Comparing Several Treatments with a Control. J Am Stat Assoc, **50**: 1096-1121.

FERNANDEZ-GALLEGO, J. A., LOOTENS, P., BORRA-SERRANO, I., DERYCKE, V., HAESAERT, G., ROLDAN-RUIZ, I., ARAUS, J. L. & KEFAUVER, S. C., 2020: Automatic wheat ear counting using machine learning based on RGB UAV imagery. The Plant journal for cell and molecular biology, **103**: 1603-1613.

FREEMAN, P. K. & FREELAND, R. S., 2015: Agricultural UAVs in the U.S.: potential, policy, and hype. Remote Sensing Applications: Society and Environment, **2**: 35-43.

GARCÍA-MARTÍNEZ, H., FLORES-MAGDALENO, H., ASCENCIO-HERNÁNDEZ, R., KHALIL-GARDEZI, A., TIJERINA-CHÁVEZ, L., MANCILLA-VILLA, O. R. & VÁZQUEZ-PEÑA, M. A., 2020: Corn Grain Yield Estimation from Vegetation Indices, Canopy Cover, Plant Density, and a Neural Network Using Multispectral and RGB Images Acquired with Unmanned Aerial Vehicles. Agriculture, **10**(7), 277.

GEIPEL, J., LINK, J. & CLAUPEIN, W., 2014: Combined Spectral and Spatial Modeling of Corn Yield Based on Aerial Images and Crop Surface Models Acquired with an Unmanned Aircraft System. Remote Sensing, **6**: 10335-10355.

GOWEN, A.A., TAGHIZADEH, M. & O'DONNELL, C.P., 2013: Using Hyperspectral Imaging for Quality Evaluation of Mushrooms. Hyperspectral imaging for food quality analysis and control, 403-430.

GRENZDÖRFFER, G. J., 2014: Crop height determination with UAS point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XL**-1: 135-140.

GRIEDER, C., HUND, A. & WALTER, A., 2015: Image based phenotyping during winter: a powerful tool to assess wheat genetic variation in growth response to temperature. Funct Plant Biol, **42**: 387-396.

HAN, L., YANG, G., DAI, H., XU, B., YANG, H., FENG, H., LI, Z. & YANG, X., 2019: Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. Plant Methods, **15**: 10.

HASSAN, M. A., YANG, M., RASHEED, A., YANG, G., REYNOLDS, M., XIA, X., XIAO, Y. & HE, Z., 2019: A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. Plant Sci, **282**: 95-103.

HERZIG, P., BACKHAUS, A., SEIFFERT, U., WIREN, N. VON, PILLEN, K. & MAURER, A., 2019: Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population. Plant Sci, **285**: 151-164.

HERZIG, P., BORRMANN, P., KNAUER, U., KLÜCK, H.-C., KILIAS, D., SEIFFERT, U., PILLEN, K. & MAURER, A., 2021: Evaluation of RGB and Multispectral Unmanned Aerial Vehicle (UAV) Imagery for High-Throughput Phenotyping and Yield Prediction in Barley Breeding. Remote Sensing, **13**: 2670.

HIJMANS, R. J., 2020: Package 'raster' - Geographic Data Analysis and Modeling.

HOFFMANN, A., MAURER, A. & PILLEN, K., 2012: Detection of nitrogen deficiency QTL in juvenile wild barley introgression lines growing in a hydroponic system. BMC genetics, **13**: 88.

HONSDORF, N., MARCH, T. J., BERGER, B., TESTER, M. & PILLEN, K., 2014a: High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. PloS one, **9**: e97047.

HONSDORF, N., MARCH, T. J., HECHT, A., EGLINTON, J. & PILLEN, K., 2014b: Evaluation of juvenile drought stress tolerance and genotyping by sequencing with wild barley introgression lines. Mol Breeding, **34**: 1475-1495.

HONSDORF, N., MARCH, T. J. & PILLEN, K., 2017: QTL controlling grain filling under terminal drought stress in a set of wild barley introgression lines. PloS one, **12**: e0185983.

HOTHORN, T., BRETZ, F. & WESTFALL, P., 2008: Simultaneous inference in general parametric models. Biom J, **50**: 346-363.

KIM, S. L., CHUNG, Y. S., JI, H., LEE, H., CHOI, I., KIM, N., LEE, E., OH, J., KANG, D. Y., BAEK, J., LEE, G. S., KWON, T. R. & KIM, K. H., 2019: New Parameters for Seedling Vigor Developed via Phenomics. Applied Sciences, **9**(9), 1752.

KUHN, M. & WICKHAM, H., 2020: Tidymodels: Easily Install and Load the 'Tidymodels' Packages. R package version 0.1.2.

LANDINI, G., RANDELL, D. A., FOUAD, S. & GALTON, A., 2017: Automatic thresholding from the gradients of region boundaries. J Microsc, **265**: 185-195.

LI, B., XU, X., ZHANG, L., HAN, J., BIAN, C., LI, G., LIU, J. & JIN, L., 2020: Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. ISPRS Journal of Photogrammetry and Remote Sensing, **162**: 161-172.

LIAKOS, K. G., BUSATO, P., MOSHOU, D., PEARSON, S. & BOCHTIS, D., 2018: Machine Learning in Agriculture: A Review. Sensors, **18**(8), 2674.

LIEBISCH, F., KIRCHGESSNER, N., SCHNEIDER, D., WALTER, A. & HUND, A., 2015: Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. Plant Methods, **11**: 9.

LIU, J. G. & PATTEY, E., 2010: Retrieval of leaf area index from top-of-canopy digital photography over agricultural crops. Agr Forest Meteorol, **150**: 1485-1490.

MARCIAL-PABLO, M. D. J., GONZALEZ-SANCHEZ, A., JIMENEZ-JIMENEZ, S. I., ONTIVEROS-CAPURATA, R. E. & OJEDA-BUSTAMANTE, W., 2018: Estimation of vegetation fraction using RGB and multispectral images from UAV. Int J Remote Sens, **40**: 420-438.

NEBIKER, S., LACK, N., ABÄCHERLI, M. & LÄDERACH, S., 2016: Light-Weight Multispectral Uav Sensors and Their Capabilities for Predicting Grain Yield and Detecting Plant Diseases. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XLI**-B1: 963-970.

OEHLSCHLAGER, J., SCHMIDHALTER, U. & NOACK, P. O., 2018: UAV-Based Hyperspectral Sensing for Yield Prediction in Winter Barley: 1-4.

OTSU, N., 1979: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, **9**: 62-66.

PASK, A., PIETRAGALLA, J., MULLAN, D. M. & REYNOLDS, M. P., 2012: Physiological Breeding II: A Field Guide to Wheat Phenotyping.

R CORE TEAM, 2020: R: A Language and Environment for Statistical computing, Foundation for Statistical Computing.

REYNOLDS, M., CHAPMAN, S., CRESPO-HERRERA, L., MOLERO, G., MONDAL, S., PEQUENO, D. N. L., PINTO, F., PINERA-CHAVEZ, F. J., POLAND, J., RIVERA-AMADO, C., SAINT PIERRE, C. & SUKUMARAN, S., 2020: Breeder friendly phenotyping. Plant Sci, **295**: 110396.

REYNOLDS, M. P., RAJARAM, S. & SAYRE, K. D., 1999: Physiological and Genetic Changes of Irrigated Wheat in the Post-Green Revolution Period and Approaches for Meeting Projected Global Demand. Crop Sci, **39**: 1611-1621.

RISCHBECK, P., ELSAYED, S., MISTELE, B., BARMEIER, G., HEIL, K. & SCHMIDHALTER, U., 2016: Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. Eur J Agron, **78**: 44-59.

SADRAS, V. O. & LAWSON, C., 2011: Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. Crop Pasture Sci, **62**: 533-549.

SCHMALENBACH, I., KORBER, N. & PILLEN, K., 2008: Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik, **117**: 1093-1106.

SCHMALENBACH, I., LEON, J. & PILLEN, K., 2009: Identification and verification of QTLs for agronomic traits using wild barley introgression lines. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik, **118**: 483-497.

SILVA-PEREZ, V., MOLERO, G., SERBIN, S. P., CONDON, A. G., REYNOLDS, M. P., FURBANK, R. T. & EVANS, J. R., 2018: Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. Journal of experimental botany, **69**: 483-496.

TAO, H., FENG, H., XU, L., MIAO, M., YANG, G., YANG, X. & FAN, L., 2020: Estimation of the Yield and Plant Height of Winter Wheat Using UAV-Based Hyperspectral Images. Sensors, **20**(4), 1231.

TARDIEU, F., CABRERA-BOSQUET, L., PRIDMORE, T. & BENNETT, M., 2017: Plant Phenomics, From Sensors to Knowledge. Curr Biol, **27**: R770-R783.

ULLMAN, S., 1979: The interpretation of structure from motion. Proc R Soc Lond B Biol Sci, **203**: 405-426.

WANG, G., SCHMALENBACH, I., KORFF, M. VON, LEON, J., KILIAN, B., RODE, J. & PILLEN, K., 2010a: Association of barley photoperiod and vernalization genes with QTLs for flowering time and agronomic traits in a BC2DH population and a set of wild barley introgression lines. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik, **120**: 1559-1574.

WANG, J. M., YANG, J. M., MCNEIL, D. L. & ZHOU, M. X., 2010b: Identification and molecular mapping of a dwarfing gene in barley (Hordeum vulgare L.) and its correlation with other agronomic traits. Euphytica, **175**: 331-342.

YUE, J. B., YANG, G. J., LI, C. C., LI, Z. H., WANG, Y. J., FENG, H. K. & XU, B., 2017: Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. Remote Sensing, **9**(7), 708 9.

ZAHN, S., KOBLENZ, B., CHRISTEN, O., PILLEN, K. & MAURER, A., 2020: Evaluation of wild barley introgression lines for agronomic traits related to nitrogen fertilization. Euphytica, **216**.

Eschentriebsterben – Erkennung der Eschenvitalität mit hyperspektralen Fernerkundungsdaten

UWE KNAUER[1,2], KATHARINA HOLSTEIN[2], BONITO THIELERT[2], DENNY SCHMELZ[2] & INA EHRHARDT[2]

Zusammenfassung: Der Beitrag beschreibt die Erfassung, Verarbeitung und Auswertung hyperspektraler Bilddaten zur Ableitung von Kenngrößen, die für ein Monitoring des Eschentriebsterbens geeignet sind. Die Darstellung beschreibt laufende Forschungsarbeiten und versteht sich als Diskussionsbeitrag zur gewählten Methodik. Die Methoden umfassen die flugzeuggestützte Datenaufnahme mit HYSPEX-Hyperspektralkameras, die Datenaufbereitung mit PARGE und ATCOR, die Ableitung eines DOM aus simultan erhobenen RGB-Bilddaten sowie die Berechnung von spektralen Indices für Korrelationsuntersuchungen zu terrestrisch erhobenen Schadstufen in den Eschenbeständen. Es wird ein Ausblick auf die weiteren Arbeiten im Projektverlauf gegeben.

1 Einleitung

Das Vorkommen der Gemeinen Esche im deutschen Wald sowie in Europa geht seit einigen Jahren kontinuierlich zurück. Dieser Rückgang ist auf das Eschentriebsterben, hervorgerufen durch den Pilz *Hymenoscyphus fraxineus* (Falsches Weißes Eschenstengelbecherchen), zurückzuführen. Der drastische Rückgang der Eschenbestände hat sowohl forstwirtschaftliche Folgen, als auch eine starke Auswirkung auf die Biodiversität im Wald. Zur Rettung der Gemeinen Esche wurde ein Verbundprojekt aus fünf Forschungsverbünden mit 27 Teilprojekten gestartet (FRAXFORFUTURE 2021). Ziel ist die Entwicklung praktikabler Lösungen zum Erhalt der Esche.

Im Beitrag wird auf Ansätze und neue Verfahren für das Monitoring des Gesundheitszustands der Gemeinen Esche durch begleitende Überwachung der Schadentwicklung mittels Fernerkundung, speziell der Hyperspektraltechnik eingegangen. Ziel der Arbeiten ist, auf Basis von hyperspektralen Fernerkundungsdaten Methoden zur Unterscheidung der Esche von anderen Baumarten sowie zur Bestimmung des Vitalitätszustandes innerhalb der Eschenpopulation zu entwickeln. Hierfür sollen aus den Spektren der Bestände mittels Methoden des maschinellen Lernens und der statistischen Datenverarbeitung geeignete Indices zur Erfassung der Esche und deren Vitalität abgeleitet werden.

Über das Bundesgebiet verteilt wurden ca. 20 Monitoringflächen mit erhöhter Beobachtungsintensität als Basis zur Datengewinnung für den Forschungsverbund eingerichtet. Auf diesen Waldflächen findet ein detailliertes Monitoring der räumlichen und zeitlichen Dynamik der Ausbreitung des Eschentriebsterbens sowie des Gesundheitszustandes von Eschenbeständen auf Einzelbaum- und Bestandesebene statt. Darüber hinaus werden anhand retrospektiver und

[1] Hochschule Anhalt, Digitale Technologien in der Pflanzenproduktion, Strenzfelder Allee 28, D-06406 Bernburg, E-Mail: uwe.knauer@hs-anhalt.de
[2] Fraunhofer IFF, Joseph-von-Fraunhoferstr. 1, D-39106 Magdeburg, E-Mail: [katharina.holstein, bonito.thielert, ina.ehrhardt]@iff.fraunhofer.de

aktueller Monitoring- und Fernerkundungsdaten Verfahren entwickelt, um das Eschensterben großflächig zu erfassen und auszuwerten. Im Rahmen der Intensivbeobachtung werden phänologische, physiologische, biochemische und holzanatomische Parameter untersucht, auf deren Basis Frühindikationen für Befallsdispositionen von Eschen abgeleitet werden sollen (PETERS et al. 2021).
Im projektbezogenen fernerkundlichen Monitoring des Gesundheitszustandes von Eschen finden u.a. Befliegungen mit räumlich und (hyper-)spektral hochauflösender Messtechnik statt. Die Befliegungszeiträume liegen im Frühjahr sowie im Spätsommer, um Daten unter unterschiedlichen Austriebsbedingungen der Bäume zu gewinnen. Die Befliegungen ergänzend erfolgt die bodenbasierte Intensivbeobachtung der ausgewählten Flächen, bei der neben der georeferenzierten Erfassung des Standorts aller Einzelbäume auch deren Vitalitätszustand wiederholt beurteilt und systematisch dokumentiert wird (FRAXMON 2021).

2 Lösungsansatz

Zur Untersuchung der Vitalitätsunterschiede und zur Entwicklung eines auf Hyperspektraldaten basierenden Monitorings werden mit flugzeuggestützten Sensoren erhobene Spektraldaten zu georeferenzierten und hinsichtlich atmosphärischer Einflüsse korrigierten Hyperspektralbildern aufbereitet. Diese bilden die Datengrundlage für weitergehende Analysen. Die Zuordnung von terrestrisch erhobenen Referenzdaten erfolgt über die Geokoordinaten der Einzelbäume. Für die Untersuchungsgebiete wurden diese im Rahmen der Flächeneinrichtung zuvor terrestrisch eingemessen (FRAXFORFUTURE 2020).

2.1 Datenerhebung

Die Erfassung der Hyperspektraldaten erfolgt über ein Dual-Kamerasystem des norwegischen Herstellers Norsk Elektro Optikk. Es kommen eine Kamera Hyspex VNIR 1800 (NEO-VNIR 2022) in Kombination mit einer Kamera Hyspex SWIR 384 (NEO-SWIR 2022) zum Einsatz. Tabelle 1 beschreibt wesentliche optische und spektrale Eigenschaften des Sensorsystems. Die gegenüber Vorgängermodellen erhöhte Aufnahmerate erlaubt den Betrieb in geringeren Flughöhen und bei höheren Fluggeschwindigkeiten ohne, dass die Gefahr einer Unterabtastung besteht. Dies erfordert den gleichzeitigen Einsatz eines Inertialmesssystems (IMU) mit vergleichbarer zeitlicher Auflösung. Eingesetzt wurde das IMU/GNSS Applanix POS-AV (APPLANIX 2022). Begrenzend auf die Wahl der Aufnahmerate wirkt allerdings die für einen guten Signal-Rausch-Abstand erforderliche Belichtungszeit, um die zur Verfügung stehende hohe radiometrische Auflösung von 16 Bit effektiv auszunutzen. Als Kompromiss zwischen der durch den Kameraöffnungswinkel und Flughöhe begrenzten Flächenleistung und hoher räumlicher Auflösung wurde eine Flughöhe von 1000 m über der Erdoberfläche für die Befliegungen geplant.

Tab. 1: Übersicht zu den Eigenschaften des Sensorsystems

Eigenschaften	VNIR 1800	SWIR 384
Wellenlängenbereich	400-1000 nm	930-2500 nm
Kameraöffnungswinkel	17°	16°
Anzahl spektraler Bänder	186	288
Anzahl räumlicher Pixel	1800	384
Spektrale Auflösung	3,26 nm	5,45 nm
Aufnahmerate (Zeilen pro Sekunde)	260 fps	400 fps
Radiometrische Auflösung	16 Bit	16 Bit

2.2 Radiometrische Korrektur der Hyperspektraldaten

Die radiometrische Korrektur der erhobenen Spektraldaten erfolgt einerseits durch Berücksichtigung von herstellerseitig bereitgestellten Kalibrierdaten für die Sensoren und andererseits durch Umrechnung der Radianzwerte in Reflektanzwerte mittels der Software ATCOR 4 (SCHLÄPFER & RICHTER 2022).

Im ersten Schritt erfolgt die Umrechnung der pro Pixel der Sensormatrix gemessenen Intensitätswerte (Einheit DN für Digital Number) in die Radianz. Dem liegt eine unter Laborbedingungen durchgeführte Referenzmessung des Herstellers zugrunde. Hierbei werden die Sensitivitätsunterschiede auf dem Sensorchip durch Verwendung einer homogenen Lichtquelle sowie einer Ulbricht-Kugel registriert und eine Matrix mit Korrekturwerten abgeleitet.

Im zweiten Schritt erfolgt die Umrechung der am Sensor ermittelten Radianzwerte in Reflektanzwerte an der Oberfläche. Hierzu werden anhand der Spektraldaten die zum Messzeitpunkt herrschenden atmosphärischen Bedingungen geschätzt. Dies beinhaltet die Quantifizierung des Anteils diffuser und direkter Strahlung im Messsignal, die Berücksichtigung von Sonnenstand und Intensität der Sonneneinstrahlung und unterschiedliche Aerosolgehalte. Die verwendete Software ATCOR 4 verwendet dazu einen modellbasierten Ansatz, der die spektrale Charakteristik typischer Oberflächen bereits berücksichtigt. Zusätzlich können Referenzflächen mit bekannter spektraler Charakteristik in die Atmosphärenkorrektur einbezogen werden. Diese Information kann über geoferenzierte Feldmessungen mit Spektrometern oder in der aufgenommenen Szene platzierte Kalibrierfelder gewonnen werden.

Operationell wird eine automatische Atmosphärenkorrektur eingesetzt. Bei wolkenfreiem Himmel und somit hinreichend gut vorhersagbarer spektraler Charakteristik der Lichteinstrahlung (TOA) anhand eines Emissionsmodells der Sonne wird aus den Navigationsdaten für jeden Flugstreifen die Hauptflugrichtung und damit die Ausrichtung zur Sonne ermittelt. Anhand des GPS-Zeitstempels erfolgt die Bestimmung von Zenith- und Azimuthwinkel der Sonne. Der Wasserdampfgehalt der Atmosphäre zum Messzeitpunkt wird anhand der Wasserabsorptionsbanden in den gemessenen Spektraldaten geschätzt und bildet den wesentlichen Parameter zur Abschätzung des Streulichtanteils am Messsignal. Die Sensoreigenschaften werden durch die bandbezogenen Parameter mittlere Wellenlänge und Halbwertsbreite (FWHM) beschrieben und sind für die verwendeten Kameras hinterlegt. Wesentliche Unterschiede für die modellbasierte Atmosphärenkorrektur bestehen zwischen maritimen, unversiegelten natürlichen und urban geprägten Landschaften. Für das Monitoring von Waldflächen werden in ATCOR 4 daher Modelle für unversiegelte natürliche Oberflächen (Modus: *rural*) vorausgewählt.

2.3 Geometrische Korrektur der Hyperspektraldaten

Hierzu werden die Positionsdaten der GNSS-Antenne/des Inertialmesssystems unter Berücksichtigung des Antennen-Sensor-Leverarm auf die raumzeitlichen Koordinaten der beiden Hyperspektralkameras umgerechnet. Hierfür kommt die Software POSPAC (TRIMBLE 2022) zum Einsatz. Die Berechnung erfolgt im Postprocessing unter Berücksichtigung korrgierter Bahndaten der Navigationssatelliten sowie der differentiellen Korrektur des GNSS-Fehlers hinsichtlich regionaler Laufzeitfehler und Veränderungen des Trägerphasensignals. Hierzu wird eine virtuelle Basisstation für das Untersuchungsgebiet durch Interpolation der Korrektursignale eines globalen Stationsnetzwerks rechnerisch bestimmt. Alternativ können auch die Daten einer realen, im Untersuchungsgebiet zu platzierenden Basisstation genutzt werden. Aufgrund der Befliegung mehrerer, räumlich weit verteilt liegender Monitoringflächen an einem Tag wurde auf virtuelle Basisstationen zurückgegriffen. Anhand von Triggersignalen der Kameras, die über einen Trigger-Eingang der IMU gemeinsam mit dem Datenstrom des Navigationssystems gespeichert werden, kann für jeden Zeitpunkt der Aufnahme einer Bildzeile eine Lageinformation des Sensors ermittelt werden. Dies erfolgt mittels der Software HYSPEX NAV, die vom Kamerahersteller bereitgestellt wird und die mit POSPAC korrigierten Daten als Eingabe erhält.

Die geometrische Korrektur der Zeilenscannerdaten erfolgt mittels der Software PARGE (SCHLÄPFER 2022). Hierzu werden die mit HYSPEX NAV zu jedem Hyperspektralbild generierten Navigationsdaten zusammen mit dem Hyperspektralbild importiert. Außerdem benötigt das Programm Informationen zu den geometrischen Eigenschaften des Sensors. Diese werden durch einen Vektor beschrieben, dessen Elemente die räumlichen Beobachtungswinkel für jeden Bildpunkt des Hyperspektralsensors darstellen. Für den NADIR-Einsatz und den geringen Öffnungswinkel des Systems (SWIR-FOV=16°) wird eine äquidistante Unterteilung des Intervalls [-8°;+8°] entsprechend der Pixelanzahl der Sensoren (VNIR-Kamera=1800 Pixel, SWIR-Kamera=384 Pixel) verwendet. Wichtig ist die Berücksichtigung der Einbaulage der Kameras zur Flugrichtung. Ggf. muss eine sich im Vorzeichen unterscheidende Konfiguration genutzt werden, die für PARGE als Sensorkalibrationsdatei bereitgestellt werden muss.

Anhand der bereitgestellten Sensorcharakteristik und der Navigationsdaten erfolgt eine geografische Eingrenzung des Befliegungsgebiets. Für dieses Gebiet kann ein digitales Oberflächenmodell (DOM) entweder aus einem globalen Geländemodell generiert oder aus anderer Quelle importiert werden. Die Verwendung eines globalen Modells bietet den Vorteil der Datenverfügbarkeit für jedwede Befliegung und die Nachteile einer geringen räumlichen Auflösung sowie mangelnder Aktualität. Da für eine Anwendung im Forst aufgrund der variierenden Wuchshöhen und bspw. gegenüber Agrarlandschaften kleinräumig großen Höhenunterschieden, varrierender Belaubung, der Bewirtschaftung, möglichen Sturmschäden und Kalamitäten ein aktuelles DOM erforderlich ist, wird dieses aus simultan erhobenen RGB-Bilddaten für die jeweilige Befliegung berechnet. Näheres zu diesem Vorverarbeitungsschritt ist in Abschnitt 3.3 beschrieben.

Der nächste Verarbeitungsschritt ist die sogenannte Boresight-Kalibrierung. Darunter versteht man die Ermittlung der Sensororientierung im Referenzkoordinatensystem. Die Boresight-Kalibrierung kann entweder passpunktbasiert oder automatisiert durch Auswertung einer Doppelkreuzbefliegung desselben Gebietes erfolgen. Der passpunktbasierte bietet den Vorteil, die

nach der Boresight-Kalibrierung erzielte Lagegenauigkeit gut einschätzen zu können. Ihr Nachteil liegt in der Erfordernis, hinreichend viele Messmarken im Gelände zu verteilen bzw. hinreichend viele geeignete Landmarken präzise zu vermessen. Ein pragmatischer Lösungsansatz ist die Verwendung von Passpunkten, deren Lage in einem bereits georefenzierten aktuellen Ortholuftbild (z.B. aus Befliegungen des amtlichen Vermessungswesens) hinreichend genau bestimmt werden kann. Kandidaten für Passpunkte sind diejenigen Landmarken, deren Lage sowohl im Ortholuftbild als auch im Hyperspektralbild sicher identifiziert werden kann. Geeignet sind bspw. die Eckpunkte von Fahrbahnmarkierungen, Fußpunkte von Pfählen, etc. Demgegenüber bietet die automatisierte Boresightkalibrierung den Vorteil, keinen Auswahlprozess für die Passpunkte zu durchlaufen und damit eine Auswertung ohne zusätzliche interaktive Komponente durchführen zu können. Eine mögliche Kombination beider Ansätze, die automatisierte Suche von geeigneten Passpunkte in den Hyperspektralbildern und im hochauflösenden digitalen Ortholuftbild, wird in (KNAUER 2013) beschrieben. Die Boresightkalibrierung wird im passpunktbasierten Verfahren für einen einzelnen Flugstreifen und im automatisierten Verfahren für die Flugstreifen des Befliegungsmusters durchgeführt. Die Ergebnisse werden in Form von Korrekturwinkelangaben in die Prozessierung der weiteren Flugstreifen übernommen. Innerhalb der Software PARGE wird dazu der Status des aktuellen Projekts in einer Datei gesichert, aus der die Boresight-Parameter später direkt geladen werden können.

Nach Durchführung dieser vorbereitenden Schritte erfolgt die eigentliche Berechnung eines hyperspektralen Orthofotos über parametrische Geokodierung in einem vorgegebenen Raster. Die Rastergröße wird als Parameter frei gewählt. Das zugrundeliegende Verfahren ist ein Raytracing-Algorithmus. Die Position auf der Oberfläche wird als Schnittpunkt des Strahls zwischen Position des Pixels und DOM ermittelt. Die in den einzelnen Bändern enthaltene spektrale Information wird den korrespondierenden Bändern des Orthofotos an der entsprechenden Position hinzugefügt. PARGE bietet Wahlmöglichkeiten bzgl. der Berechnung. Es ist zwischen hoher räumlicher Genauigkeit und spektraler Genauigkeit abzuwägen. Die schnelle bilineare Nächste-Nachbarn-Interpolation als Resampling-Verfahren stellt einen guten Kompromiss dar und wird daher vorrangig eingesetzt.

Nach der Prozessierung der einzelnen Flugstreifen erfolgt die Erstellung eines Orthomosaiks. Hierzu werden die einzubeziehenden Flugstreifen eines Befliegungsgebiets ausgewählt. Die Reihenfolge der Auswahl bestimmt die Art der Überlagerung. Es besteht die Wahlmöglichkeit der überlappenden Darstellung nach der Reihenfolge und alternativ die Auswahl nach der Nächster-Nachbar-Regel (MODUS: *center cropped*). Die Sichtbarkeit eines Bildpunkts wird hier durch die Lage der Ebene in dieser Reihenfolge bestimmt. In der Regel bietet sich aufgrund visuell besserer visueller Ergebnisse die Wahl der überlappenden Darstellungsweise an. Das Orthomosaik wird i.d.R. mittels der Angabe von Geokoordinaten auf einen rechteckigen Bereich, der die Monitoringfläche komplett enthält begrenzt. Dieses wird für die weitere Verarbeitung im ENVI-Format gespeichert.

Im Fall einer unzureichenden Georeferenzierung, die sich insbesondere in den Überlappungsbereichen der Flugstreifen bei der Mosaikberechnung zeigen kann, muss entweder der Gesamtprozess inklusive Boresight-Kalibrierung erneut durchlaufen werden oder es kann eine

nachträgliche Minimierung der geometrischen Abweichung zwischen den Flugstreifen im Gesamtverband erfolgen. Hierzu wird aus den Spektralbildern der einzelnen Flugstreifen jeweils ein einkanaliges Bild eines spektralen Index (für den VNIR-Bereich wird der NDVI verwendet) aufgrund höherer Robustheit gegenüber Schattenwurf und BRDF-Effekten genutzt. Für diese Bilder wird in der Abfolge der Flugstreifen die paarweise Korrelationsfunktion durch schnelle zweidimensionale Fouriertransformation, Multiplikation der Fourierspektren und Rücktransformation in den Ortsbereich berechnet. Entsprechend des Korrelationstheorems kann die Verschiebung zwischen den Flugstreifen als Maximalwertbestimmung im Ergebnisbild bestimmt werden.

2.4 Berechnung des Digitalen Oberflächenmodells

Das für die Georeferenzierung der hyperspektralen Zeilenscannerdaten erforderliche DOM wird mit der Software Agisoft Metashape Professional (AGISOFT 2022) erstellt. Eine geeignete Alternative stellt die Software Pix4Dmapper (PIX4D 2022) dar. Simultan in einem festen Zeitintervall während des Fluges aufgezeichnete RGB-Bilder werden genutzt, um eine dichte Punktwolke und durch Vermaschung ein 3D-Modell des Kronendaches zu erzeugen. Anhand dieser Daten wird das DOM als zweidimensionale Karte der Höheninformation im Referenzkoordinatensystem erzeugt und als GEOTIFF gespeichert. Dieses Format kann in PARGE importiert werden.

2.5 Berechnung vitalitätsbezogener spektraler Kenngrößen

Das atmosphärenkorrigierte und georeferenzierte Hyperspektralbild bildet die wesentliche Grundlage für die Berechnung spektraler Indizes. Diese werden als Rasterbilder entsprechend der räumlichen Auflösung des Hyperspektralbildes erzeugt. In der Literatur sind eine Reihe spektraler Indices beschrieben (JONES & VAUGHAN 2010), die sich zur Erfassung von Vitalitätsunterschieden der Vegetation eignen. Weite Verbreitung hat die Auswertung des NDVI (ROUSE 1974), der auf Blattebene zum Chlorophyllgehalt und auf Kronenebene gut zur Biomasse und damit dem Belaubungsgrad korreliert ist. Ergänzend wird häufig der PRI ausgewertet, der Stresseffekte insbesondere auch durch seine Sensitivität gegenüber Carotinoiden im Bereich des grünen Lichts charakterisieren hilft (GAMON et al. 1992). Neben dem forstlichen Monitoring werden spektrale Indices auch in der Fernerkundung für landwirtschaftliche Anwendungen und in den Umweltwissenschaften genutzt, so dass zahlreiche weitere Indices definiert wurden (HEINRICH 2022).

Darüber hinaus bietet die Aufnahme der Hyperspektralbilder die Möglichkeit, für die Vitalitätsansprache verschiedener Schädigungsgrade der Esche neue spezifische Indices zu entwickeln. Hierbei werden zwei Ansätze verfolgt. Erstens die Bestimmung einzelner Wellenlängen und Wellenlängenkombinationen mit hoher Relevanz für die Vitalitätsbewertung und deren Verrechnung zu einem neuen skalaren Indexwert (MAHLEIN et al. 2013). Zweitens das Trainieren eines Klassifikations- oder Regressionsmodells zur Vorhersage eines Vitalitätsindex, der als Referenzwert zur Verfügung gestellt wird (BANI et al. 2017). Der Vitalitätsindex ergibt sich dann als Wert der Modellantwort auf neue Spektraldaten. Eine Kombination beider Verfahren ist ebenfalls möglich (BENDEL et al. 2020). Mit geeigneten Verfahren erfolgt zunächst eine

Bewertung der Relevanz einzelner Prädiktorvariablen (KNAUER et al. 2017). Damit kann aus komplexen Modellen, die anhand der vollständigen Hyperspektraldaten trainiert wurden, eine Teilmenge an relevanten spektralen Bändern identifiziert werden. Für diese wird dann im zweiten Schritt untersucht, wie sie sich z.B. in Form eines differentiellen Indexwertes miteinander kombinieren lassen, um eine robuste Bewertung der Eschenvitalität zu ermöglichen. Aufbauend können verschiedene Modelle zur Vitalitätsansprache, aber auch zur Erkennung der in einem Ensemble-Klassifikator zusammengefasst werden, um das Vorhersageergebnis zu verbessern (KNAUER et al. 2019).

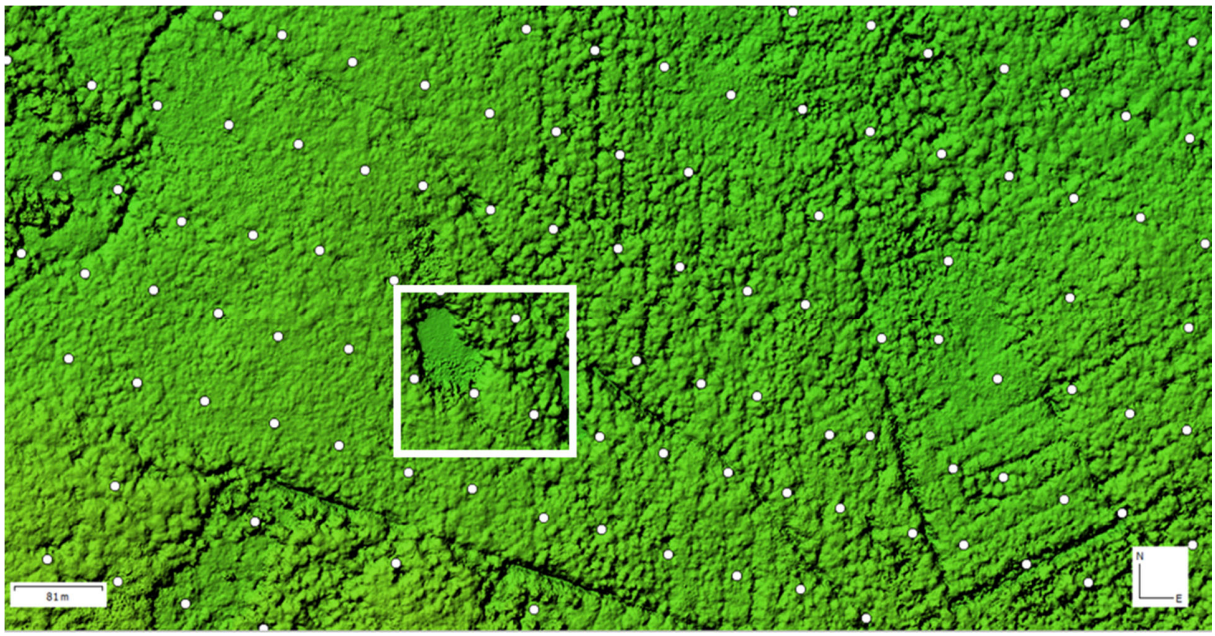


Abb. 1: Ausschnitt aus dem Oberflächenmodell eines Befliegungsgebiets in Thüringen mit einer Rastergröße von 11 cm. Der Kreis zeigt die Lage der forstlichen Monitoringfläche. Die Punkte über dem Gebiet zeigen die Koordinaten der Bildaufnahmepositionen an. Der zugrundliegende Messflug erfolgte am 22.7.2021.

3 Erste Ergebnisse

Es wurden Messflüge am 11.9.2020, 22.07.2021 und 9.9.2021 durchgeführt. Entsprechend der zuvor beschriebenen Methodik wurden DOM für die einzelnen Befliegungsgebiete berechnet. Abb. 1 zeigt beispielhaft das Oberflächenmodell des Waldgebiets Ettersberg in Thüringen. Der Kreis kennzeichnet in Näherung die Lage der Beobachtungsfläche. Die Lagekorrektur erfolgte unter Einbeziehung von 7 Passpunkten, die gleichmäßig im Befliegungsgebiet platziert wurden.
Unter Nutzung des DOM sowie der Softwarepakete PARGE und ATCOR 4 erfolgt die Georeferenzierung der Hyperspektraldaten. Abb. 2 zeigt beispielhaft den Ausschnitt ein aus dem Hyperspektralbild für das Untersuchungsgebiet abgeleiteten spektralen Index. Die exakte Georeferenzierung und Verbesserung der Lagegenauigkeit ist Gegenstand der aktuellen Arbeiten. Arbeiten hinsichtlich der Baumartenerkennung und der spektralen Charakterisierung des Eschentriebsterbens konnten auf Basis der beschriebenen Methodik aufgenommen werden. Die

georeferenzierten Karten werden in ein GIS-System importiert und dort weiter ausgewertet. Parallel werden Trainingsdatensätze für die Baumartenklassifikation und die Modellierung von Vitalitätsunterschieden aufgebaut. Die Prozessierung dieser Datensätze erfolgt in einem separaten, speziell für das maschinelle Lernen mit Hyperspektraldaten entwickelten Umgebung des Fraunhofer IFF auf Basis von MATLAB.

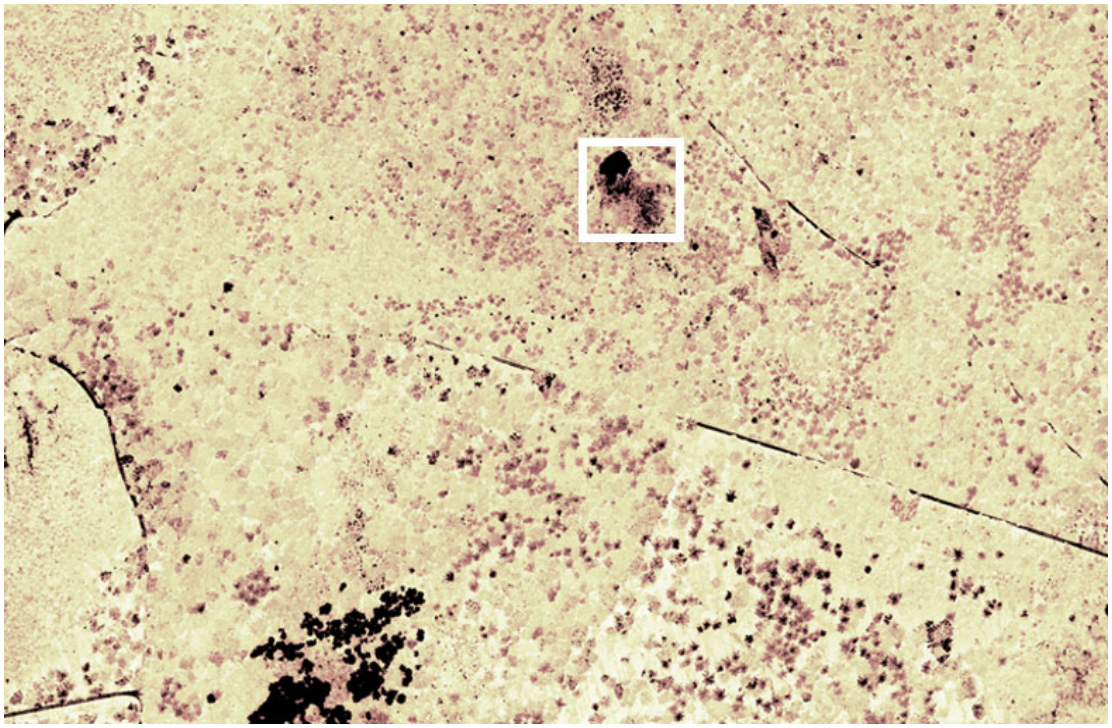


Abb. 2: Kartenausschnitt mit Visualisierung des spektralen Index NDVI zur Sichtbarmachung von Vitalitätsunterschieden. Dunklere Flächen entsprechen Bereichen mit geringerem Anteil vitaler Vegetation. Der zugrundeliegende Messflug erfolgte am 11.9.2020. Eine Korrespondenz zu Abb. 1 wird durch die markierte Fläche als Landmarke hergestellt.

4 Fazit & Ausblick

Der Artikel beschreibt Methodik und Vorgehensweise bei Datenerfassung und Datenauswertung zum Monitoring des Eschentriebsterbens mittels hyperspektraler Fernerkundung. Es werden die methodischen Schritte der Datenverarbeitung von der Bildaufnahme bis zur Ableitung von Vitalitätsindices beschrieben. Erste Ergebnisse illustrieren den gewählten Ansatz. Es sind jedoch umfangreiche weitere Arbeiten erforderlich. Daher sind weitere Messflüge geplant, um die erforderliche Datengrundlage für die Ableitung von robusten Modellen zur Vitalitätsvorhersage zu schaffen. Die Flüge finden begleitend zu umfangreichen terrestrischen Bonituren/Ansprachen auf Einzelbaumebene statt, so dass geeignete Referenzdaten zur Verfügung stehen. Als prinzipielle Herausforderung sind Unterschiede in der Bewertung des Eschenschädigungsgrades durch einen Beobachter am Boden und durch die Bewertung der Krone aus der Luft zu nennen. Trotz dieser Differenzen bieten die Fernerkundungsmethoden jedoch ein hohes Potential, da sie durch hohe

räumliche Auflösung gleichzeitig eine flächenhafte Erfassung und eine Analyse auf Einzelbaumebene erlauben. Die hohe spektrale Auflösung von Hyperspektralkamerasystemen kann unter Einbeziehung räumlicher Beziehungen genutzt werden, verschiedene Schädigungsstufen abzubilden.

5 Danksagung

Die Arbeiten werden als Teil des Verbundprojektes FraxForFuture von der Fachagentur Nachwachsende Rohstoffe aus Mitteln des Waldklimafonds unter dem Förderkennzeichen 2019WK20D4 gefördert. Besonderer Dank gilt den Kolleginnen und Kollegen des Forstlichen Forschungs- und Kompetenzzentrums Gotha für die Erhebung von Referenzdaten. Gleichermaßen gilt unser Dank den Partnern im Teilprojekt FraxMon für die gute Zusammenarbeit und den intensiven fachlichen Austausch.

6 Literaturverzeichnis

AGISOFT, 2022: Features Professional Edition. https://www.agisoft.com/features/professional-edition/, letzter Zugriff am 22.01.2022.

APPLANIX, 2022: POS-AV. https://www.applanix.com/products/posav.htm, letzter Zugriff am 22.01.2022.

BANI, G., SEIFFERT, U., BIEHL, M. & MELCHERT, F., 2017: Adaptive basis functions for prototype-based classification of functional data. 12. International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization, 1-8, https://doi.org/10.1109/WSOM.2017.8020020.

BENDEL, N., KICHERER, A., BACKHAUS, A., KLÜCK, H.-C., SEIFFERT, U., FISCHER, M., VOEGELE, R. T. & TOEPFER, R., 2020: Evaluating the Suitability of Hyper- and Multispectral Imaging to Detect Foliar Symptoms of the Grapevine Trunk Disease Esca in Vineyards. Plant Methods, **16**, Artikel 142, https://doi.org/10.1186/s13007-020-00685-3.

FRAXFORFUTURE, 2020: Flächeneinrichtung. https://www.fraxforfuture.de/untersuchungsflaechen/flaecheneinrichtung, letzter Zugriff am 22.01.2022.

FRAXFORFUTURE, 2021: FraxForFuture - Gemeinsam für den Erhalt der Esche. https://www.fraxforfuture.de, letzter Zugriff am 22.01.2022.

FRAXMON, 2021: FraxMon. https://www.fraxforfuture.de/boniturmerkmale, letzter Zugriff am 22.01.2022.

GAMON, J., PENUELAS, J. & FIELD, C., 1992: A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment, **41**, 35-44.

HEINRICH, V., 2022: Index DataBase – A database for remote sensing indices, https://www.indexdatabase.de/, letzter Zugriff am 22.01.2022.

JONES, H. G. & VAUGHAN, R. A., 2010: Remote Sensing of Vegetation: Principles, Techniques, and Applications, Oxford University Press, 1. Auflage, ISBN 978-0199207794.

KNAUER, U. & SEIFFERT, U., 2013: An Approach for Automated Registration of Hyperspectral Images for Boresight Calibration. Publikationen der Deutschen Gesellschaft für

Photogrammetrie, Fernerkundung und Geoinformation e.V., Seyfert, E. (Hrsg.), Band **22**, 271-278.

KNAUER, U., MATROS, A., PETROVIC, T., ZANKER, T., SCOTT, E. S. & SEIFFERT, U., 2017: Improved classification accuracy of powdery mildew infection levels of wine grapes by spatial-spectral analysis of hyperspectral images. Plant Methods, **13**(1), 1-15, https://doi.org/10.1186/s13007-017-0198-y.

KNAUER, U., REKOWSKI, C., STECKLINA, M., KROKOTSCH, T., PHAM MINH, T., HAUFFE, V., KILIAS, D., EHRHARDT, I., SAGISCHEWSKI, H., CHMARA, S. & SEIFFERT, U., 2019: Tree Species Classification Based on Hybrid Ensembles of a Convolutional Neural Network (CNN) and Random Forest Classifiers. Remote Sensing, **11**(23), 2788, https://doi.org/10.3390/rs11232788.

MAHLEIN, A.-K., RUMPF, T., WELKE, P., DEHNE, H.-W., PLÜMER, L., STEINER, U. & OERKE, E.-C., 2013: Development of spectral indices for detecting and identifying plant diseases. Remote Sensing of Environment, **128**, 21-30.

NEO-VNIR, 2022: HySpex VNIR-1800. https://www.hyspex.com/hyspex-products/hyspex-classic/hyspex-vnir-1800/, letzter Zugriff am 22.01.2022.

NEO-SWIR, 2022: HySpex SWIR-384. https://www.hyspex.com/hyspex-products/hyspex-classic/hyspex-swir-384/, letzter Zugriff am 22.01.2022.

PETERS, S., LANGER G. & KÄTZEL, R., 2021: Bonitur geschädigter Eschen im Kontext des Eschentriebsterbens. AFZ-Der Wald, **76**(12), 28-31.

PIX4D, 2022: Pix4Dmapper. https://www.pix4d.com/de/produkt/pix4dmapper-photogrammetrie-software, letzter Zugriff am 22.01.2022.

ROUSE, J. W., HAAS, R. H., SCHELL, J. A. & DEERING, D. W., 1974: Monitoring Vegetation Systems in the Great Plains with ERTS. 3. ERTS-1 Symposion der NASA, Washington DC, 309-317.

SCHLÄPFER, D., 2022: PARGE Image Rectification. https://www.rese-apps.com/software/parge/, letzter Zugriff am 22.01.2022.

SCHLÄPFER, D. & RICHTER, R., 2022: ATCOR 4 - for Airborne Remote Sensing Systems. https://www.rese-apps.com/software/atcor-4-airborne/index.html, letzter Zugriff am 22.01.2022.

TRIMBLE, 2022: Applanix POSPac MMS. https://de.geospatial.trimble.com/products-and-solutions/applanix-pospac-mms, letzter Zugriff am 22.01.2022.

Potenzial der Fernerkundung für den Moorschutz - Bewertung von Methoden zur Erkennung von stark bewachsenen Wasserflächen

KATRIN KRZEPEK[1], JAKOB SCHMIDT[1] & DOROTA IWASZCZUK[1]

Zusammenfassung: Die Seen in Mooren sind typischerweise von bis zu vier Meter hohen Schilfpflanzen bedeckt, was die tachymetrische Vermessung der Uferlinie nahezu unmöglich macht. Vor diesem Hintergrund wurde in der vorliegenden Arbeit die Vermessung der freien und der von Pflanzen überwachsenen Wasserfläche in einem Moor mittels Fernerkundung untersucht. Genutzt wurden C-Band SAR-Daten in VH- und VV-Polarisation sowie der modifizierte Normalized Differenced Water Index (MNDWI) und der Normalized Differenced Vegetation Index (NDVI). Als Klassifikationsverfahren dienten das Schwellwertverfahren und der Random Forest Algorithmus. Es wird gezeigt, dass die Seefläche trotz der Pflanzenbedeckung eindeutig detektiert werden kann. Das C-Band ist geeignet zur Durchdringung der Schilfvegetation. Durch die Kombination der Radar- und Multispektraldaten kann die Fehlklassifizierung der Wasserflächen im restlichen Bildausschnitt deutlich verringert werden.

1 Einleitung und Motivation

Der Klimawandel gilt als größte Herausforderung der heutigen Zeit. Die Europäische Union hat sich für 2050 das Ziel gesetzt Klimaneutralität zu erreichen. Der Ausstoß klimaschädlicher Gase muss somit im Gleichgewicht mit deren Speicherung stehen. Ein wichtiger Baustein bildet dabei die Vegetation. Beim Wachstum von Pflanzen wird Kohlenstoffdioxid (CO_2) aus der Luft gebunden. Werden die Pflanzen zersetzt, gelangt das Gas in die Atmosphäre zurück. Im wasserreichen Ökosystem Moor hingegen können die abgestorbenen Pflanzen aufgrund der sauerstoffarmen Umgebung nicht oder nur teilweise zersetzt werden (BUNDESAMT FÜR NATURSCHUTZ 2022). Dabei wird der Kohlenstoff im Torf gespeichert. Obwohl Moore nur ca. 3 % der Landfläche der Erde bedecken, binden sie doppelt so viel Kohlenstoff wie alle Wälder weltweit (JOOSTEN 2008).

Die Faktoren Wasserhaushalt und Vegetationswachstum spielen somit eine wichtige Rolle für die Ökosystemdienstleistungen eines Moors. Die Detektion offener Wasserflächen mittels Fernerkundungsdaten wurde unter anderem zur Überwachung von Flutereignissen (z.B. PULVIRENTI et al. 2011; MARTINIS et al. 2015; BONI et al. 2016; CAZALS et al. 2016) oder zur Bestimmung der Landnutzung und Landbedeckung (*land use land cover,* LULC) (z.B. HAQUE & BASAK 2017; BAGWAN & SOPAN 2021) in zahlreichen Studien untersucht. LIU et al (2019) stellten eine Methode für die LULC-Detektion vor, in der Multispektral- und SAR-Daten kombiniert werden und die eine Korrektheit (*User-Accuracy*) von bis zu 100 % erreicht.

[1] Technische Universität Darmstadt, Institut für Geodäsie,
Fachgebiet Fernerkundung und Bildanalyse, Franziska-Braun-Str. 7, D-64287 Darmstadt,
E-Mail: [katrin.krzepek, jakob.schmidt1, dorota.iwaszczuk]@tu-darmstadt.de

In dieser Arbeit wird untersucht, ob sich die Kombination von Multispektral- und SAR-Daten dazu eignet, stark bewachsene Wasserflächen zu detektieren.
Genutzt werden C-Band SAR-Daten sowie der modifizierte Normalized Differenced Water Index (MNDWI). Zur Abgrenzung der Wasserfläche wird zum einen ein Fusionsbild erzeugt, auf das der Schwellwertansatz angewandt wird. Zum anderen dienen die Daten als Inputparameter für eine Klassifizierung der Wasserfläche mittels des Random Forest Algorithmus. Weitergehend wird aus der erkannten Wasserfläche unter Verwendung des Normalized Differenced Vegetation Index (NDVI) die Fläche des Sees bestimmt, die mit vitaler Vegetation bedeckt ist.
Die Fernerkundung bietet Vorteile wie eine kostengünstige Erfassung und hohe zeitliche Auflösung der Daten. Zusätzlich werden Satellitendaten weltweit erhoben und sind dadurch vergleichbar. Dieser Punkt ist besonders relevant, da Klimaschutzleistungen aus dem Sektor Land- und Forstwirtschaft zukünftig auf einem CO_2-Markt anerkannt und entlohnt werden sollen (LEHMANN 2021).
Der Fokus dieser Arbeit liegt auf der Ausarbeitung einer Methode, die auch von Akteuren angewendet werden kann, die nicht auf Fernerkundung spezialisiert sind.

2 Messung der Seefläche

Zur Messung der Seefläche wird der Radar-Rückstreukoeffizient σ^0 (C-Band, VH- und VV-Polarisation) und der MNDWI nach XU (2006) verwendet (Formel 1). Der MNDWI ist eine Weiterentwicklung des Normalized Differenced Water Index (NDWI) von MCFEETERS (1996) und verwendet anstelle des nahen Infrarots (NIR) das mittelnahe Infrarot (MIR).

$$MNDWI = \frac{(Green - MIR)}{(Green + MIR)} \qquad (1)$$

Die Verwendung dieser Bänder geht auf die charakteristische Spektralkennlinie von Wasser zurück. Der physikalische Hintergrund der Bestimmung der Seefläche über SAR-Daten ergibt sich aus dem Verhalten der Rückstreuung des Radarsignals (σ^0) abhängig von Oberflächeneigenschaft und Dielektrizitätskonstante (ALBERTZ 2009).
Es werden zwei verschiedene Klassifikationsverfahren genutzt: Zum einen wird der Schwellwertansatz auf ein Fusionsbild aus den drei Datensätzen (σ^0-VH, σ^0-VV und MNDWI) angewandt. Zum anderen dienen diese drei Datensätze als Inputparameter für einen Random Forest Algorithmus, der die Klassifikation vornimmt.
Der Schwellwertansatz wird im ersten Schritt auf die drei Datensätze einzeln angewandt. Zur Festlegung der σ^0-Schwellwerte wird ein Histogramm aus den Werten der Pixel, die im Seeausschnitt liegen, gebildet und der Schwellwert beim Minimum der zwei sich zeigenden Spitzen festgelegt. Pixel, die einen niedrigeren Wert als den Schwellwert aufweisen, werden als Seefläche klassifiziert (siehe Abb. 1).

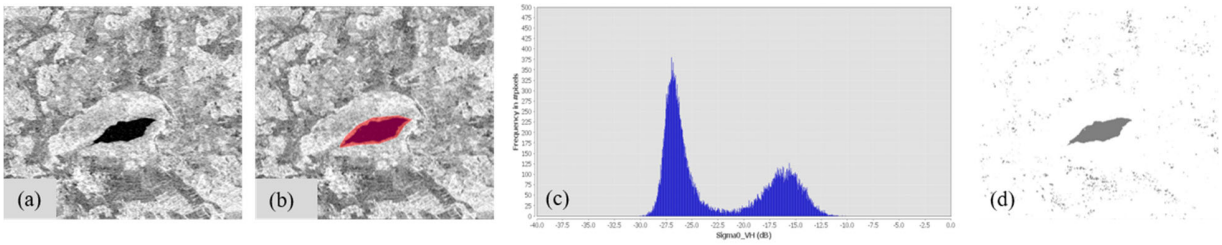


Abb. 1: Vorgehen Schwellwertansatz (a) σ^0-VH März (b) rot: Seeausschnitt (c) Histogramm der Pixel aus dem Seeausschnitt (d) Seemaske

Die Festlegung des MNDWI-Schwellwerts erfolgt auf zwei unterschiedlichen Wegen. Zum einen werden wie zuvor die Histogramme ausgewertet und das Minimum zwischen den zwei Spitzen der See- und Landpixel ermittelt. Zum anderen wird ein fester Grenzwert von -0,25 festgelegt.
Im Folgenden wird aus den Daten ein Fusionsbild erstellt. Aufgrund der Merkmalseigenschaften ergibt sich, dass beim MNDWI hohe Werte charakteristisch für die Seefläche sind, während bei den σ^0-Werten Wasserflächen einen niedrigen Wert aufweisen. Zur sinnvollen Kombination müssen die Datensätze skaliert und ein Teil der Datensätze invertiert werden. Das Fusionsbild wird im Anschluss wie folgt errechnet:

$$Fusionsbild = G_{VH} * \sigma^0_{VH} + G_{VV} * \sigma^0_{VV} + G_{MNDWI} * MNDWI \qquad (2)$$

Bei G handelt es sich um die Gewichtung der einzelnen Datensätze, die aus den vorherigen Ergebnissen ermittelt wird. Auf diese wird in Kapitel 4.2 näher eingegangen. Für das Fusionsbild wird ebenfalls eine Klassifikation der Seefläche mittels Schwellwertverfahren durchgeführt.
Das Klassifikationsverfahren Random Forest wurde in den 1990er Jahren vorgestellt und vielfach erweitert (BREIMAN 2001). Der Algorithmus wird beim Machine Learning eingesetzt und hat den Vorteil, dass es ein einfach anzuwendendes Verfahren ist, welches nur relativ kurze Trainingszeiten benötigt.
Des Weiteren wird mittels der zuvor detektierten Seefläche und unter Verwendung des NDVIs die mit Vegetation bedeckte Fläche des Sees ermittelt. Diese Größe kann Hinweis auf das Kohlenstoffbindungspotenzial des Moors geben. Als Bedingung wird gesetzt, dass die Fläche als See klassifiziert wurde und einen NDVI-Wert vitaler Vegetation aufweist. Genutzt wird ein NDVI-Grenzwert von 0,1 in Anlehnung an BALDENHOFER (2022). Der NDVI wird wie folgt errechnet:

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \qquad (3)$$

3 Anwendung im Untersuchungsgebiet

Zur Anwendung der vorgestellten Methode wurde das Federseeried in Baden-Württemberg gewählt. Inmitten des Moors liegt der ca. 1,4 km² große Federsee (WERNICKE & BOSCH 2022). Der Federsee ist im Mittel nur einen Meter tief und umgeben von einem bis zu 100 m breiten Schilfgürtel, der an weitläufige Streu- und Feuchtwiesen, extensiv genutztes Grünland und naturnahe Moorwälder grenzt (WERNICKE & BOSCH 2022). Der Schilfgürtel bedeckt auch große Teile der Wasserfläche (siehe Abb. 2).

Das NABU-Naturschutzzentrum Federsee erfasst täglich die Höhe des Seepegels. Als Untersuchungszeitraum wurde das Jahr 2019 gewählt.
Im Rahmen des Projekts wurden Satellitenbilder der Sentinel-Satelliten der European Space Agency (ESA), Sentinel-1 und Sentinel-2, verwendet. Die Sentinel-1-Satelliten messen mittels C-Band SAR-Instrumenten in einfacher oder dualer Polarisation. Es wurden Ground Range Detected (GRD)-Daten genutzt, die mittels Interferometic wide swath mode (IW) aufgenommen wurden.
Für das Untersuchungsgebiet wurde für jeden Monat des Jahres 2019 ein Sentinel-1 Produkt heruntergeladen, sowohl in VH- als auch in VV-Polarisation. Die Rückstreudaten wurden mit der open-source Software SNAP vorverarbeitet, was unter anderem die Entfernung des thermischen Rauschens, die radiometrische Kalibrierung und die Geländekorrektur umfasste. Als Speckle Filter wurde der refined Lee-Filter verwendet.
Zudem wurde ein Sentinel-2 Level 2A Produkt pro Monat verwendet. Der MNDWI wurde unter Verwendung von Band 3 (559,8 nm) und Band 11 (1.613,7 nm) errechnet. Für eine möglichst gute Fusion und Vergleichbarkeit der Datensätze wurden Aufnahmezeitpunkte gewählt, die zeitlich nah beieinander lagen. Bei den Multispektralbildern konnten auf Grund von starker Wolkenbedeckung keine Bilder für die Monate Januar und November ausgewählt werden. Bei den SAR-Daten zeigte der VV-Datensatz des Märzes starke Auffälligkeiten und musste bei der Auswertung der Daten verworfen werden.

4 Ergebnisse und Diskussion

4.1 Analyse der Schwellwerte

Die Histogramme der SAR-Daten des Seeausschnitts zeigen in der Regel zwei Spitzen, die gut voneinander unterschieden werden können. Die Spitze links im Histogramm wird hervorgerufen durch niedrige σ^0-Werte von ca. -22 bis -32 dB (M: -22,4 dB; SD: 1,1 dB) für die VH-Polarisation und ca. -25 bis -15 dB (M: -16,1 dB; SD: 1,9 dB) für die VV-Polarisation und ist der Wasserfläche zuzuordnen. Die rechte Spitze ist der Landfläche zuzuordnen, die einen höheren σ^0-Wert von ca. -20 bis - 12,5 dB für die VH-Polarisation und ca. -15 bis -5 dB für die VV-Polarisation aufweist.
Auffällig ist, dass die Spitzen der Histogramme der VV-Polarisation weniger eindeutig voneinander abzugrenzen sind und der Schwellwert schlechter zu bestimmen ist. In den Monaten Januar und März ist es nicht möglich, aufgrund des beschriebenen Vorgehens einen Schwellwert für die VV-Polarisation festzulegen, da sich keine zwei Spitzen ermitteln lassen.
Allgemein muss festgehalten werden, dass die Wasserfläche auf den Sentinel-1-Bildern gut erkennbar ist. Eine Ausnahme bildet hier das Bild der VV-Polarisation vom 15.03.2019. Analysen der äußeren Einflüsse wie Niederschlag, Temperatur, Windgeschwindigkeit etc. konnten die körnige Struktur des Bildes nicht erklären. Es wird vermutet, dass ein Sensorproblem vorlag.
Die Histogramme sind in der Regel klar zu interpretieren. Menschliche Fehler beim Ablesen des Schwellwerts sind jedoch nicht auszuschließen. Fraglich ist zusätzlich, ob das Minimum zwischen den beiden Spitzen tatsächlich die Grenze zwischen See- und Landpixeln darstellt. Unklar ist dabei, wie die Methode mit dem Uferbereich umgeht, der ökologisch schwer dem See- oder Landgebiet zugeordnet werden kann, da die Übergänge des ökologischen Systems nicht scharf voneinander abgegrenzt sind.

Zur Bestimmung der Seegröße mittels optischer Daten wurde der MNDWI berechnet. Da die Bestimmung eines Minimums zwischen See- und Landpixel aus dem Histogramm der MNDWI-Werte nicht eindeutig möglich war, wurde auf Basis einer visuellen Auswertung ein Grenzwert von -0,25 festgelegt. Die erzeugten Masken wurden für die Erstellung der Fusionsbilder verwendet.
Werden die MNDWI-Bilder mit den Luftbildern verglichen, stellt sich heraus, dass die Bereiche mit den Werten -0,25 bis 0,25 größtenteils die bewachsene Seefläche darstellen (siehe Abb. 2).

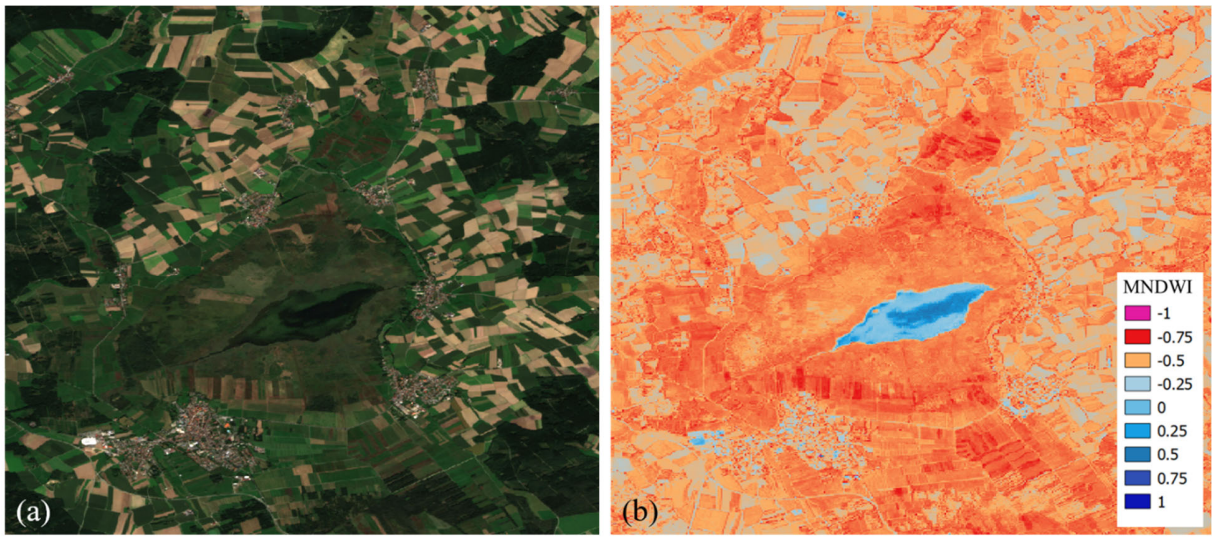


Abb. 2: (a) Luftbild September (B) MNDWI September

4.2 Bewertung der Datenquellen

Die Mittelwerte sowie die Standardabweichungen der Flächenmessungen ergeben sich wie folgt:

Tab. 1: Mittelwert (M) und Standardabweichung (SD) der gemessenen Seeflächen

Methode	M [km²]	SD [km²]
σ^0-VH	1,414	0,017
σ^0-VV	1,401	0,020
MNDWI	1,374	0,037
Fusionsbild	1,386	0,016
Random Forest	1,362	0,031

Die Seegrößen, die mittels Schwellwertverfahren aus den Fusionsbildern ermittelt wurden, weisen die geringste Standardabweichung auf. Eine Methode mit einer geringen Standardabweichung ist jedoch nicht grundsätzlich genau. Schwankungen in den Messungen der Wasserflächen können auch durch tatsächlich stattfindende Veränderungen der Seefläche hervorgerufen werden. Da keine Referenzdaten für die Seefläche vorlagen, wurde die Genauigkeit der Seegrößen wie folgt bewertet:
Zum einen wurden die Korrelationen mit dem Seepegel bestimmt. Das NABU-Naturschutzzentrum Federseeried misst täglich die Höhe des Seepegels. Der Seepegel am Tag der

Radaraufnahme kann somit exakt ermittelt werden. Dieses Vorgehen wird für alle zwölf Monate durchgeführt. Es wird angenommen, dass bei einem steigenden Seepegel auch die Seeoberfläche zunimmt. Datenreihen bei denen der Korrelationskoeffizient zwischen Seepegel und Seefläche hoch ausfällt, werden somit als besser geeignet für die Messung gewertet als Datenreihen mit einem niedrigen Korrelationskoeffizient Dieses Bewertungsverfahren eignet sich nicht für Verfahren, die Datensätze der verschiedenen Satelliten kombinieren, da die Aufnahmen hier in der Regel nicht am selben Tag erfolgen.
Es ergeben sich folgende Korrelationskoeffizienten (r) nach Pearson:
σ^0-VH: r = 0,72; σ^0-VV: r = 0,46; MNDWI: r = 0,61.
Aufgrund des hohen Korrelationskoeffizienten von r = 0,72, der visuell guten Erkennbarkeit des Sees auf den σ^0-VH-Bildern und der einfachen Festlegung der Schwellwerte in den Histogrammen liegt der Schluss nahe, dass sich die VH-Polarisation am besten von den betrachteten Datensätzen zur Messung der Fläche des hier betrachteten Moorsees eignet.
Die ermittelten Korrelationskoeffizienten wurden zur Erstellung des Fusionsbilds genutzt. Erstellt wurden neun Fusionsbilder, da nur in neun von zwölf Monaten alle drei Datenquellen genutzt werden konnten. Die Gewichtung *G* der Datensätze (Formel 2) wurde festgesetzt als *1+Korrelationskoeffizient*. Das Endergebnis der Methode ist abhängig von der Gewichtung der Datensätze. Nachteilig ist dabei, dass die Gewichtungen der einzelnen Datensätze über das Jahr hinweg konstant angenommen und des Weiteren auf Grundlage der Korrelationskoeffizienten bestimmt wurden. Die feste Gewichtung ist insofern problematisch, da die Datensätze je nach betrachtetem Zeitpunkt sehr unterschiedlich geeignet sind, Wasserflächen klar erkennen zu können. Auch die Korrelationskoeffizienten sind kritisch zu betrachten, da eine geringe Korrelation nicht zwangsläufig mit der Güte des Messverfahrens zusammenhängen muss. Sie kann z.B. auch dadurch entstehen, dass der See ein unregelmäßiges Profil hat, sodass bei gleichmäßig steigendem Pegel die Fläche nicht in entsprechender Relation zunimmt. Auch sind zehn bis zwölf Wertepaare aus Seegröße und Seepegel eine zu geringe Stichprobe, um eine allgemeingültige Aussage zu treffen.
Im Vergleich zum Fusionsbild bietet der Random Forest Ansatz den Vorteil, dass der Algorithmus individuell mit den Datensätzen trainiert werden kann. Der Algorithmus gibt zusätzlich zum Klassifikationsergebnis auch eine Karte der Confidence an. Auf dieser sind Bereiche hervorgehoben, deren Klassifikationsergebnis Unsicherheiten aufweisen. Auffällig ist, dass die Unsicherheit in den Uferbereichen besonders hoch ist (siehe Abb. 3).

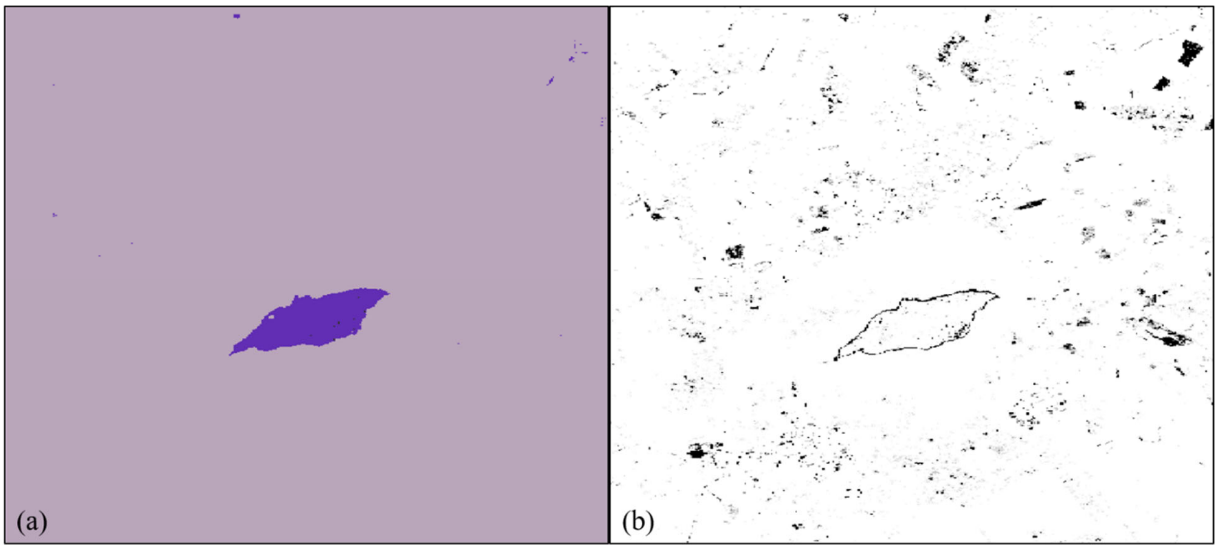


Abb. 3: Klassifikationsergebnis des Random Forest Algorithmus für den Monat September (a) Labelbild (b) Confidence

Dies ist zum einen auf den Mangel an Trainingsdaten aus diesem Bereich zurückzuführen. Die Trainingsdaten mussten händisch erstellt werden und es wurden nur Bereiche aus der Mitte des Sees gewählt, da diese zweifelsfrei dem See zuzuordnen sind. Dem Algorithmus fehlten somit Referenzdaten im Bereich des Ufers zum Trainieren. Zum anderen kann die Unsicherheit darauf zurückgeführt werden, dass auch aus ökologischen Gesichtspunkten eine Zuordnung im Uferbereich uneindeutig ist und dies nicht durch eine Schwäche der Methode begründet ist.

Für die Bewertung der Verfahren wurde zum anderen die von den Methoden fehlklassifizierte Fläche ermittelt. Das gesamte Bild umfasst 84,163 km². Es wird ein Ausschnitt um den See festgelegt (siehe Abb. 1). Liegen die von der Methode als See klassifizierte Flächen im Seeausschnitt, gelten sie als richtig klassifiziert. Liegen sie außerhalb, werden sie als fehlklassifiziert eingestuft.

Zur Bewertung werden die fehlklassifizierten Flächen übers Jahr gemittelt. Es ergeben sich folgende Anteile: VH: M = 6,981 %, VV: M = 6,626 %, MNDWI: M = 2,234 %, Fusion: M = 0,232 %, Random Forest: M = 0,228 %.

Insbesondere bei der Betrachtung der fehlklassifizierten Flächen zeigt sich die Kombination von SAR-Daten und Multispektralbildern als deutlich überlegen gegenüber der Nutzung einzelner Datenquellen. Dies lässt sich darauf zurückführen, dass jeder der drei Datensätze Stärken und Schwächen bei der Erkennung von Wasserflächen aufweist. Das C-Band zeigt eine gute Durchdringung der Vegetation auf dem Wasser. Die Radardaten werden jedoch auch durch andere Faktoren beeinflusst und zeigen auch außerhalb der Wasserfläche niedrige Werte. Durch die Kombination der verschiedenen Datenquellen und die Festlegung eines gemeinsamen Schwellwertes wird die Methode stabilisiert und zeigt einen sehr geringen Anteil fehlklassifizierter Flächen im gesamten Gebiet (siehe Abb. 4).

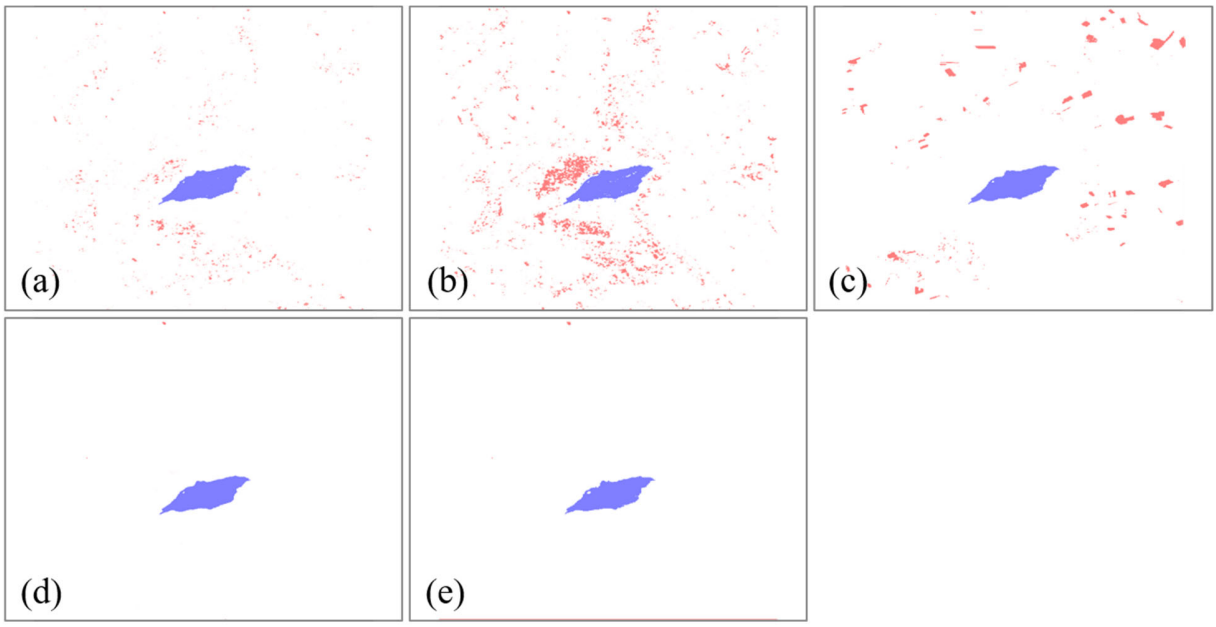


Abb. 4: Als See klassifizierte Flächen im Monat Juli, blau: im Seeausschnitt, rot: außerhalb des Seeausschnitts (a) Schwellwertverfahren σ^0-VH (b) Schwellwertverfahren σ^0-VV (c) Schwellwert-verfahren MNDWI (d) Schwellwertverfahren Fusionsbild (e) Random Forest Algorithmus mit σ^0-VH, σ^0-VV und MNDWI

Der Vergleich der Methoden zeigt, dass der Anteil der fehlklassifizierten Flächen des Random Forest Algorithmus mit 0,228 % und des Schwellwertverfahrens angewandt auf das Fusionsbild mit 0,231 % ähnlich gering ausfallen. Die Ergebnisse sind vergleichbar trotz der unterschiedlich aufwendigen Methoden. Dies ist ein positives Zeichen in Bezug auf die praktische Anwendung der Methode, da beim Random Forest Ansatz der vor Ort gemessene Seepegel nicht benötigt wird und die Schwellwertsuche entfällt.

4.3 Beobachtung der bewachsenen Seefläche

Die Kombination des NDVI mit der zuvor erstellten Seemaske erweist sich zur Detektion der von Pflanzen überwachsenen Seefläche als geeignet.
Die erstellten Masken zeigen eine Entwicklung der vitalen Pflanzenbedeckung im Jahresverlauf auf bis zu 93 % (siehe Abb. 5). Der Federsee ist in den Monaten Mai bis Oktober zum größten Teil bis fast vollständig bewachsen. Da jedoch keine Referenzdaten vorliegen, kann das Ergebnis nur anhand der visuellen Interpretation der Luftbilder beurteilt werden.
Ein Vorteil der Erstellung einer Maske ist die Quantifizierung der Fläche. Die Flächenmessungen können über mehrere Jahre vorgenommen werden und so Daten zur biologischen Entwicklung des Moorsees liefern.

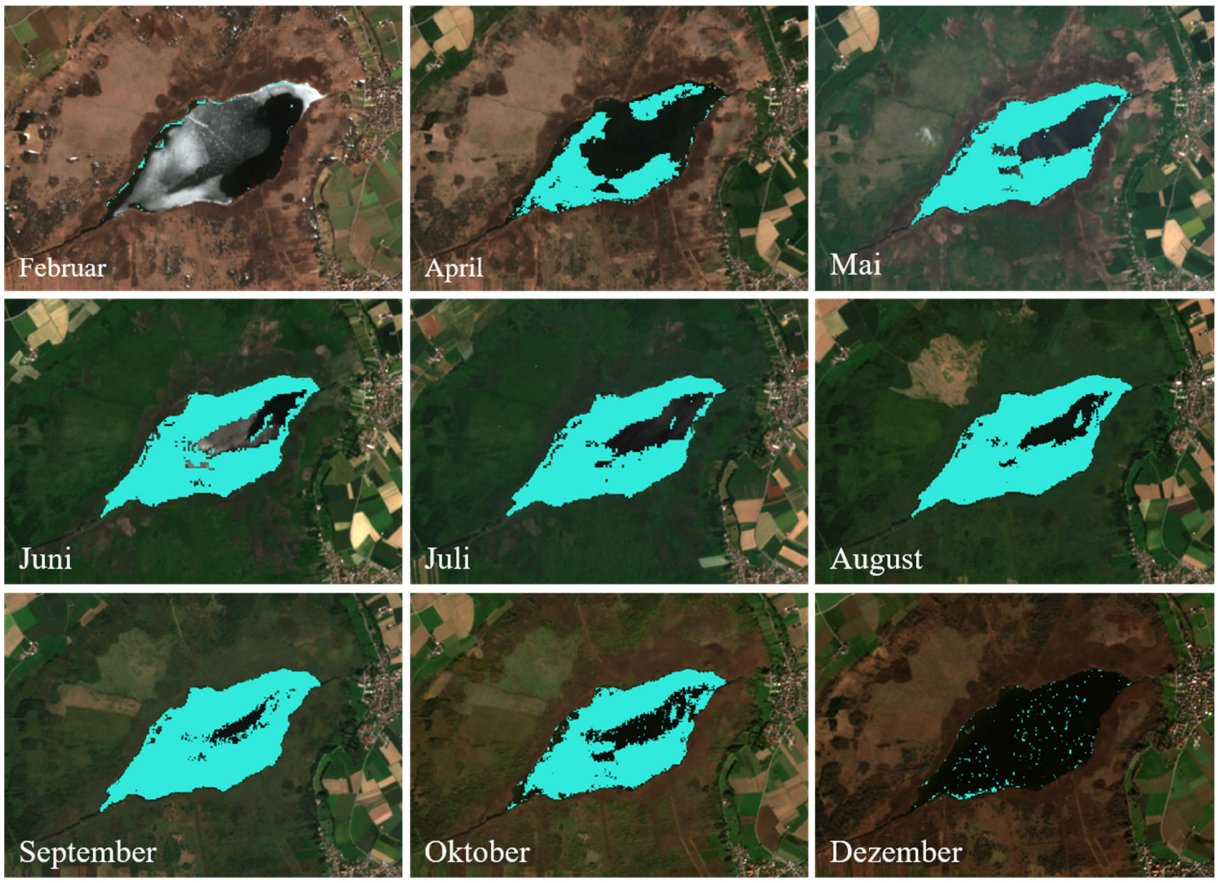


Abb. 5: Maske der bewachsenen Seefläche. Bedingung: Als Seefläche klassifiziert nach Fusionsbild-Schwellwertverfahren und NDVI > 0,1. Prozentualer Anteil vom See: Februar: 2 %; April: 45 %; Mai: 74 %; Juni: 78 %; Juli: 78 %; August: 85 %; September: 93 %; Oktober: 76 %; Dezember: 4 %

5 Fazit & Ausblick

Diese Arbeit zeigt, dass die Fernerkundung einen wichtigen Beitrag zur Erzeugung von Datengrundlagen für die Erfassung der Klimaschutzleistungen von Mooren leisten kann. Bei der Analyse der Ergebnisse zeigt sich, dass die Flächen, die über den Schwellwertansatz und die VH-polarisierten SAR-Daten gewonnen wurden mit einem r-Wert von 0,72 die höchste Korrelation mit dem Seepegel aufweisen. Das C-Band zeigt sich allgemein als geeignet für die Durchdringung der Moorvegetation auf dem Wasser. Wird der Anteil fehlklassifizierter Flächen betrachtet, sind mit ca. 0,2 % die Verfahren stark überlegen, die die Radardaten und den Index aus dem Multispektralbild kombinieren.

Auch die Abgrenzung der Vegetationsfläche an Land von der Vegetationsfläche auf dem Wasser erzielt durch die Kombination der Fernerkundungsdaten gute Ergebnisse. Die Größe dieser Fläche kann Hinweis auf das Kohlenstoffbindungspotenzial des Moors geben. Durch die Verwendung von frei verfügbaren Sentinel-Daten, open-Source Software SNAP und QGIS und des einfach anzuwendenden Random Forest Algorithmus kann die Methode auch von Akteuren reproduziert

werden, deren Schwerpunkt im Bereich Ökologie und Naturschutz und nicht in der Fernerkundung liegt.
Die zeitliche Auflösung der Fernerkundungsverfahren ist mit Wiederholzeiten von ca. drei bis fünf Tagen generell als gut zu bewerten. Probleme ergeben sich jedoch daraus, dass die Aufnahmezeitpunkte von Sentinel-1 und Sentinel-2 in der Regel nicht übereinstimmen und Sentinel-2-Bilder bei Wolkenbedeckung nicht genutzt werden können. Die Daten der Satelliten können aufgrund der unterschiedlichen Aufnahmezeitpunkte nur unter Vorbehalt kombiniert werden. Die räumliche Auflösung der genutzten Fernerkundungsdaten ist bezogen auf das kleinräumig sehr variable Ökosystem Moor gering. Zu Ungenauigkeiten kann es beispielsweise aufgrund von Mischpixeln kommen, wenn verschiedene Landbedeckungen in einem Pixel enthalten sind. Weiterer Forschungsbedarf und Potenziale liegen in der Anwendung der Methodik auf andere Moorseen, der Verbesserung der Auflösung und der genaueren Betrachtung des Uferbereichs.

6 Literaturverzeichnis

ALBERTZ, J., 2009: Einführung in die Fernerkundung - Grundlagen der Interpretation von Luft- und Satellitenbildern. WBG, Darmstadt, 4. Auflage. ISBN: 9783534231508.

BAGWAN, W.A., SOPAN GAVALI, R., 2021: Dam-triggered Land Use Land Cover change detection and comparison (transition matrix method) of Urmodi River Watershed of Maharashtra, India: a Remote Sensing and GIS approach. Geology, Ecology, and Landscapes, 1-9, https://doi.org/10.1080/24749508.2021.1952762.

BALDENHOFER, K. G., 2022: Lexikon der Fernerkundung, NDVI. https://www.fe-lexikon.info/lexikon-n.htm#ndvi, letzter Zugriff am 06.01.2022.

BONI, G, FERRARIS, L, PULVIRENTI, L, SQUICCIARINO, G., PIERDICCA, N., CANDELA, L., PISANI, A. R., ZOFFOLI, S., ONORI, R., PROIETTI, C., PAGLIARA, P., 2016: A Prototype System for Flood Monitoring Based on Flood Forecast Combined With COSMO-SkyMed and Sentinel-1 Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **9**(6), 2794–2805, https://doi.org/10.1109/JSTARS.2016.2514402.

BREIMAN, L., 2001: Random Forests. Machine Learning, **45**(1), 5-32, https://doi.org/10.1023/A:1010933404324.

BUNDESAMT FÜR NATURSCHUTZ, 2022: Moore, Entstehung und Zustand. https://www.bfn.de/entstehung-und-zustand, letzter Zugriff 06.01.2022.

CAZALS, C., RAPINEL, S., FRISON, P.-L., BONIS, A., MERCIER, G., MALLET, C., CORGNE, S., RUDANT, J.-P., 2016: Mapping and Characterization of Hydrological Dynamics in a Coastal Marsh Using High Temporal Resolution Sentinel-1A Images. Remote Sensing, **8**(7), https://doi.org/10.3390/rs8070570.

JOOSTEN, H., 2008: Peatlands and Carbon. Assessment on Peatlands, Biodiversity and Climate Change, Main Report, Global Environment Centre, Kuala Lumpur and Wetlands International, Wageningen, ISBN: 978-983-43751-0-2.

HAQUE, I., BASAK, R., 2017: Land cover change detection using GIS and remote sensing techniques: A spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh. The

Egyptian Journal of Remote Sensing and Space Science, **20**(2), 251–263, https://doi.org/10.1016/j.ejrs.2016.12.003.

LEHMANN, N., 2021: EU-Kommission will Landwirte für Kohlenstoffbindung belohnen. https://www.agrarheute.com/politik/eu-kommission-will-landwirte-fuer-kohlenstoffbindung-belohnen-588500, letzter Zugriff 06.01.2022.

LIU, S., QI, Z., LI, X. & Y, A., 2019: Integration of Convolutional Neural Networks and Object-Based Post-Classification Refinement for Land Use and Land Cover Mapping with Optical and SAR Data. Remote Sensing, **11**(6), https://doi.org/10.3390/rs11060690.

MARTINIS, S., KERSTEN, J., TWELE, A., 2015: A fully automated TerraSAR-X based flood service. ISPRS Journal of Photogrammetry and Remote Sensing, **104**, 203–212, https://doi.org/10.1016/j.isprsjprs.2014.07.014.

MCFEETERS, S. K., 1996: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, **17**(7), 1425-1342, https://doi.org/10.1080/01431169608948714.

PULVIRENTI, L., PIERDICCA, N., CHINI, M., GUERRIERO, L., 2011: An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. Natural Hazards and Earth System Sciences, **11**(2), 529–540, https://doi.org/10.5194/nhess-11-529-2011.

WERNICKE, K. & BOSCH, S., 2022: Weihen überm Wackelwald. https://www.nabu.de/natur-und-landschaft/schutzgebiete/nabu-schutzgebiete/baden-wuerttemberg/05702.html, letzter Zugriff am 06.01.2022.

XU, H., 2006: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, **27**(14), 3025-3033, https://doi.org/10.1080/01431160600589179.

3D-Deformationsanalyse und Rissdetektion in multitemporalen Voxeldaten von Röntgentomographen

FRANK LIEBOLD[1] & HANS-GERD MAAS[1]

Zusammenfassung: Der Beitrag stellt ein neues Verfahren zur 3D-Deformationsanalyse vor, das sich für spröde Materialien auf Basis mehrerer Voxeldatensätze aufeinanderfolgender Lastzustände anwenden lässt. Zunächst wird zwischen den Volumendaten des Nulllastzustands ohne Verformungen und dem verformten Zustand ein subvoxel-genaues dreidimensionales Verschiebungsvektorfeld mit Hilfe der dreidimensionalen Kleinsten-Quadrate-Zuordnung berechnet. Durch Tetrahedrisierung der Punkte des Verschiebungsfeldes entsteht ein Tetraedernetz, in dem im Folgenden die Änderungen der inneren Geometrie der Tetraeder analysiert werden. Dabei wird für jeden Tetraeder ein Deformationsvektor berechnet, dessen Betrag als skalares Deformationsmaß verwendet wird. Durch Anwendung des Schwellwertverfahrens auf das Skalarfeld werden Risse detektiert. Weiterhin lassen sich Rissbreiten und Scherungsbewegungen von den Deformationsvektoren ableiten.

1 Einleitung

Zur photogrammetrischen Deformationsanalyse in der Materialprüfung im Bauwesen sind in den vergangenen Jahren zahlreiche Publikationen erschienen. Viele dieser Beiträge basieren auf Bildsequenzauswertungen: Z.B. untersuchten FRASER & RIEDEL (2000) mit Hilfe eines trinokularen Kamerasystems thermische Verformungen an Stahlbalken. BENNING et al. (2004) und LANGE et al. (2006) verwendeten auch eine trinokulare Kameraanordnung bei Untersuchungen im bautechnischen Versuchswesen. Dabei wurden die Koordinaten der Punkte eines regelmäßigen Rasters signalisierter Kreismarken auf Oberflächen der Probekörper verfolgt. Streckenänderungen zu benachbarten Rasterpunkten wurden zur Deformationsmessung genutzt. Auch Rissbreiten wurden abgeleitet. HAMPEL & MAAS (2009) nutzten Bildkorrelationsverfahren zur Berechnung dichter subpixelgenauer Verschiebungsvektorfelder für künstliche sowie natürliche Punkte. Rissbreiten und -positionen wurden durch Analyse von Profilen in den Koordinatenrichtungen ermittelt. BARAZZETTI & SCAIONI (2010) verwendeten ein monokulares Kamerasystem zur Beobachtung eines Betonbalkens bei einem Belastungstest und verfolgten signalisierte Punkte mit Hilfe von Kreuzkorrelation und der Kleinsten-Quadrate-Zuordnung. In weiteren Versuchen wurde natürliche Textur genutzt. Dabei wurde eine Bildvorverarbeitung mit dem Wallis-Filter durchgeführt und dann ein Interestoperator zur Ermittlung der zu verfolgenden Punkte angewandt. KOSCHITZKI et al. (2012) nutzten auch ein monokulares Kamerasystem und ebenfalls einen Interestoperator. Durch Triangulation der Punkte wurde ein Dreiecksnetz generiert. Flächenverhältnisse zwischen der verformten Epoche und der Referenzepoche unter Nulllast dienten der Lokalisierung deformierter Bereiche und Risse. LIEBOLD & MAAS (2016) ging zur Berechnung von Hauptdehnungen

[1] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstr. 10, D-01069 Dresden, E-Mail: [frank.liebold, hans-gerd.maas]@tu-dresden.de

über und stellte Verfahren zur Filterung des Dehnungsfeldes vor. Auch GEERS et al. (1996) beschrieb ein Verfahren zur Berechnung von Dehnungen aus 2D-Verschiebungsfeldern. Später wurde der relative Translationsvektor als alternatives Deformationsmaß für die monokulare Bildsequenzanalyse vorgestellt, was eine Ableitung von Rissbreiten ermöglicht (LIEBOLD & MAAS 2018, 2020). Der Algorithmus kann auch für unebene (zylindrische) Oberflächen angewandt werden und Deformationen tangential zur Oberfläche auf Basis von Stereobildsequenzen bestimmt werden (LIEBOLD et al. 2019). LIEBOLD et al. (2020) zeigte zudem ein Verfahren, das auch vertikale Versätze berücksichtigt und berechnete 3D-Rissöffnungsvektoren.
Die eben genannten Methoden bezogen sich auf die Vermessung von Oberflächendeformationen. Die digitalen Bildkorrelationsverfahren lassen sich zur digitalen Volumenkorrelation erweitern, sodass auch Beiträge zur Analyse von sequentiellen Volumendaten existieren: BAY et al. (1999) untersuchten Proben, die mit Röntgentomographie erzeugt wurden. Sie nutzten die 3D-Kreuzkorrelation zur Berechnung von 3D-Vektorverschiebungsfeldern und berechneten auf dieser Basis Dehnungsfelder zur Deformationsanalyse. MAAS et al. (1994) nutzten zur Berechnung der 3D-Verschiebungsfelder die subvoxel-genaue Kleinste-Quadrate-Zuordnung, bei der eine 12-Parameter-Affintransformation als Modell diente. LIEBOLD et al. (2021) griff darauf zurück und wandte es auf In-situ-Röntgentomographiedaten an. Die Punkte des 3D-Verschiebungsfeldes wurden zu einem Tetraedernetz trianguliert und zur Deformationsanalyse Dehnungen für diese Tetraeder berechnet. Der vorliegende Beitrag baut auf den Arbeiten von LIEBOLD & MAAS (2018, 2020) sowie LIEBOLD et al. (2020, 2021) auf, erweitert diese Methoden und wendet sie auf Volumendaten an.

2 Deformationsanalyse und Rissdetektion bei Voxeldaten

Für die Analysen dieses Beitrags wurde ein Datensatz eines in-situ-Experiments mit einem SHCC-Probekörper (SHCC: strain-hardening cement-based composites; faserbewährter Zementkomposit) unter Zugbelastung aus LORENZONI et al. (2020) genutzt. Bei acht verschiedenen Laststufen wurden Volumendaten mit Hilfe von Röntgentomographie (Mikro-CT) erstellt und die ersten sieben Datensätze wurden zur Auswertung verwendet. Die Datenaufnahme erfolgte mit dem GE-VTomeX-CT-Scanner bei der Bundesanstalt für Materialforschung und –prüfung. Bei der Rekonstruktion entstanden 2024 Schichten à 750 × 1250 Pixeln mit einer Voxelgröße von $(13\ \mu m)^3$. In Abb. 1 sind drei Querschnitte des Volumens im Nulllastzustand dargestellt.

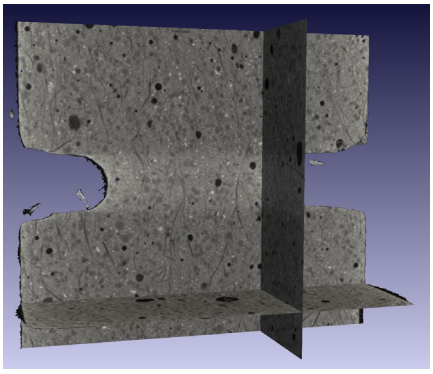


Abb. 1: 3D-Ansicht des Referenzzustandes unter Nulllast mit drei Querschnitten durch den Volumendatensatz.

2.1 Berechnung eines Verschiebungsvektorfelds

Für den Referenzvolumendatensatz erfolgt die Festlegung einer Punktmenge, die verfolgt werden soll. Dazu wird ein Raster im 10-Voxelabstand festgelegt, wobei Punkte ausgeschlossen werden, deren Standardabweichung der Grauwerte in der Punktumgebung einen Grenzwert (hier 5,5) unterschreiten. Abb. 2 zeigt beispielhaft ein ausgedünntes Punktraster mit einem Punktabstand von 40 Voxeln (vx). Einschlüsse, Fasern sowie Luftlöcher bieten einen ausreichenden Kontrast für Bildkorrelationsverfahren.

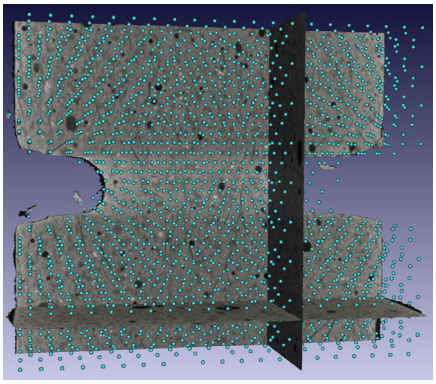


Abb. 2: Punktmenge, die zur Erstellung eines Verschiebungsvektorfeldes verfolgt werden soll. Der Punktabstand beträgt in der Abbildung 40 Voxel.

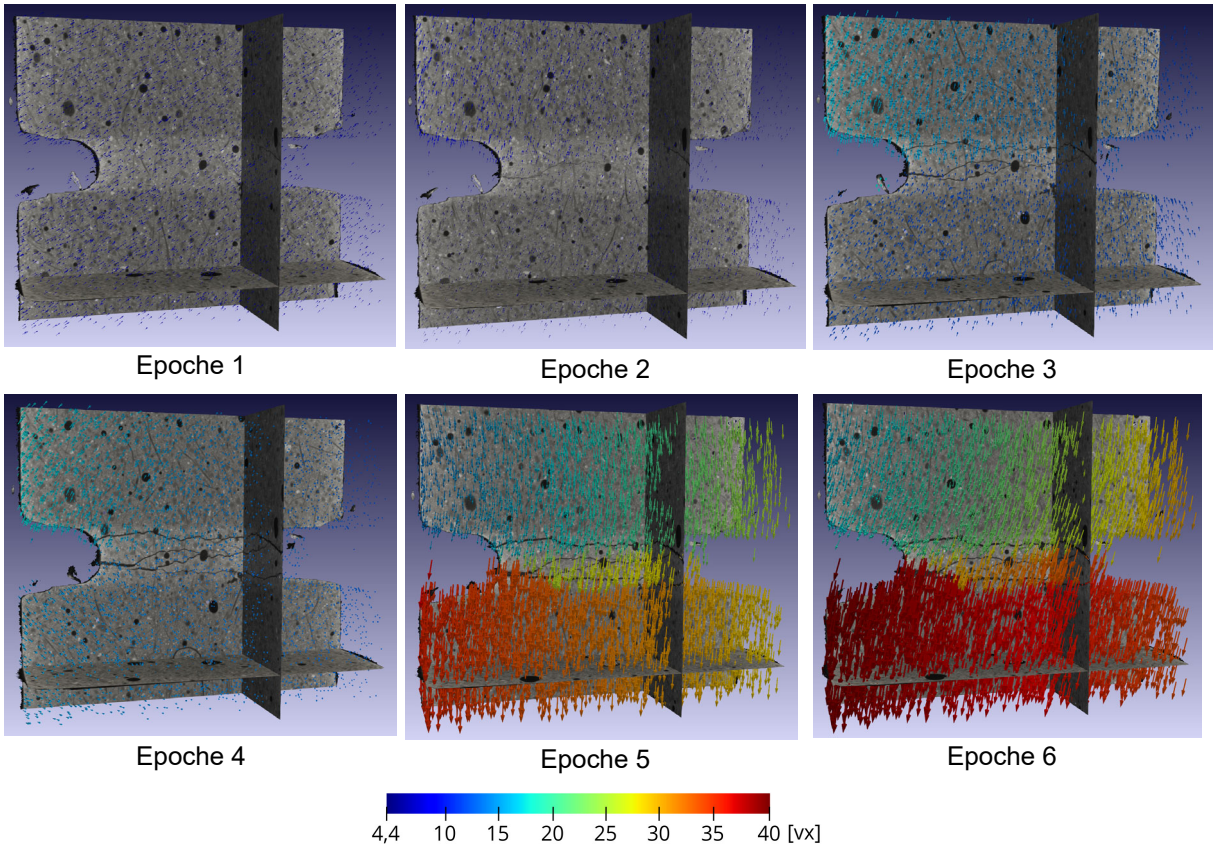


Abb. 3: Verschiebungsvektorfelder (ausgedünnt) für sechs deformierte Epochen. Die Längen der Pfeile entsprechen den Verschiebungen mit einem Vergrößerungsfaktor von 1,5.

Es folgt eine subvoxel-genaue Verschiebungsvektorfeldberechnung mit Hilfe der 3D-Kleinsten-Quadrate-Zuordnung (Maas et al 1994; Liebold et al. 2021). Um jeden Punkt im Referenzvolumens wird ein Subvolumen von 15×15×15 vx ausgewählt, die für den Algorithmus genutzt werden. Um Näherungswerte zu beschaffen, wird das 3D-Kreuzkorrelationsverfahren verwendet (Bay et al. 1999). In Abb. 3 sind die Verschiebungsvektorfelder für sechs deformierte Epochen zu sehen. Bei Rissen sind Diskontinuitäten zwischen den Farben der Pfeile zu erkennen. Die Standardabweichungen der Verschiebungen liegen bei wenigen hundertstel Voxeln.

2.2 Tetraederanalyse

2.2.1 Tetrahedrisierung

Die Punkte des Verschiebungsfelds, bei denen erfolgreich ein Versatz mit der 3D-Kleinsten-Quadrate-Zuordnung berechnet werden konnte, werden zu einem Tetraedernetz vermascht. Dabei werden zur Netzerstellung die Koordinaten im Referenzzustand genutzt. Letztlich liegen für die Tetraedereckpunkte zwei Koordinatensätze (Referenzzustand und deformierter Zustand) zur weiteren Auswertung vor. Abb. 4 zeigt das Tetraedernetz mit den Koordinaten des Referenzzustands, wobei hier das Netz zur besseren Sichtbarkeit für eine Punktemenge mit einem Abstand von 40 vx dargestellt ist.

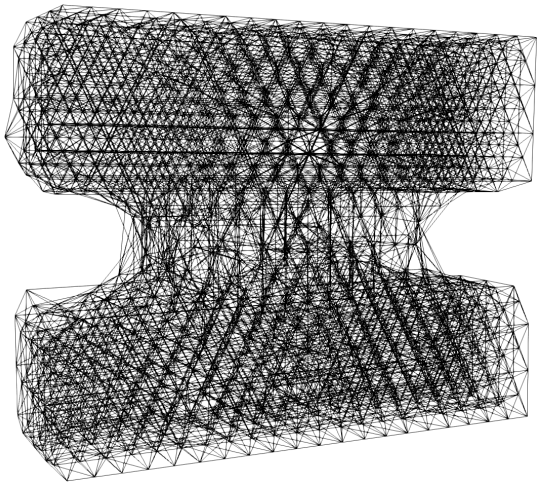
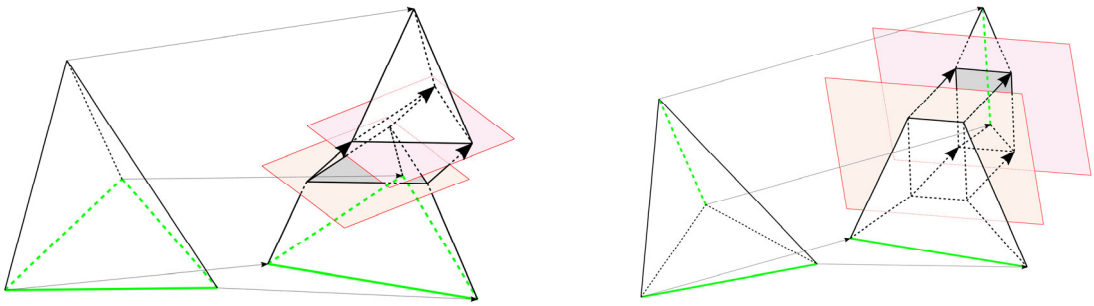


Abb. 4: Tetraedernetz mit Punktabstand von 40 vx und den Koordinaten des Referenzzustands.

Da es bei jeder Epoche eine unterschiedliche Anzahl erfolgreicher Zuordnungen (Verschiebungsvektoren) für das Verschiebungsvektorfeld gibt, erfolgt die Tetrahedrisierung für jede Epoche neu.

2.2.2 Mathematisches Modell

Ein Riss durch einen Tetraeder kann mit einer vereinfachten Aufteilung des Tetraeders in zwei Teile modelliert werden, wobei sich der eine Teil mit einer Relativtranslation verschiebt. Dabei können zwei mögliche Fälle auftreten: Fall 1 kennzeichnet sich dadurch, dass sich die drei Kantenlängen einer Seitenfläche des Tetraeders nicht ändern. Diese sind in Abb. 5a grün gekennzeichnet. Die Aufteilung ist durch die zwei Trennebenen (rot) dargestellt. Hingegen sind bei Fall 2 zwei gegenüberliegende Seitenlängen des Tetraeders unverändert, siehe Abb. 5b (grüne Kanten).



(a) (b)

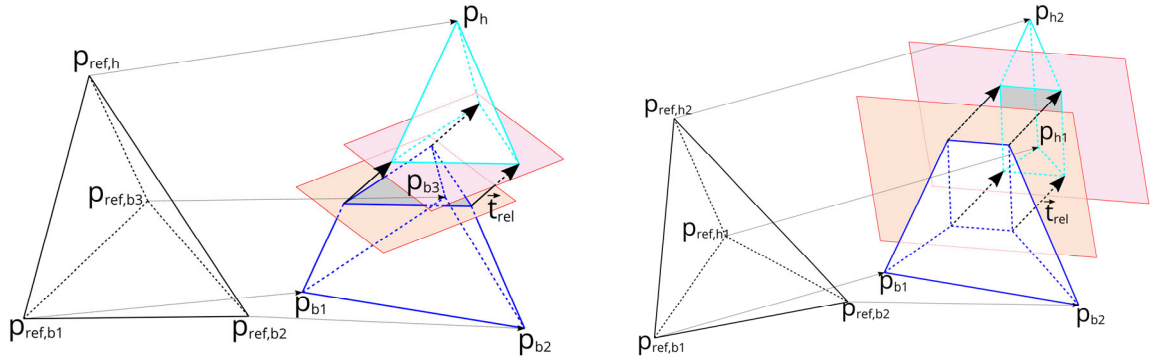
Abb. 5: Mögliche Fälle bei Tetraederteilung: a) drei Eckpunkte (ein undeformiertes Dreieck) auf einer Seite des Risses, einer auf der anderen; b) jeweils zwei Eckpunkte (eine undeformierte Kante) auf beiden Seiten des Risses.

Als mathematisches Modell für die Tetraederauswertung kann das Modell für die Analyse von 3D-Oberflächendaten von LIEBOLD et al. (2020) übernommen werden:

$$\vec{p}_i = \begin{cases} \vec{t} + \boldsymbol{R} \cdot \vec{p}_{ref,i}, & \text{für } i \in M_1 \\ \vec{t} + \boldsymbol{R} \cdot \vec{p}_{ref,i} + \vec{t}_{rel}, & \text{für } i \in M_2 \end{cases} \quad (1)$$

mit $\vec{t}$ .. Translationsvektor
$\boldsymbol{R}$ .. Rotationsmatrix
$\vec{t}_{rel}$ .. Relativtranslationsvektor
$\vec{p}_{ref,i}$ .. Referenzkoordinaten des i. Eckpunkts des Tetraeders
$\vec{p}_i$ .. Koordinaten des i. Eckpunkts des Tetraeders im deformierten Zustand

Die Mengen M_1 und M_2 ergeben sich bei Fall 1 wie folgt: $M_1 = \{b_1, b_2, b_3\}$ und $M_2 = \{h\}$. Für Fall 2 gilt: $M_1 = \{b_1, b_2\}$ und $M_2 = \{h_1, h_2\}$. Abb. 6a (Fall 1) und Abb. 6b (Fall 2) enthalten die die Punktbezeichnungen und den Relativtranslationsvektor $\vec{t}_{rel}$.



(a) (b)

Abb. 6: Tetraederteilung für Fall 1 (a) und Fall 2 (b) mit Punktbezeichnungen und dem relativen Translationsvektor.

In folgendem Schema ist eine mögliche Fallunterscheidung dargestellt:

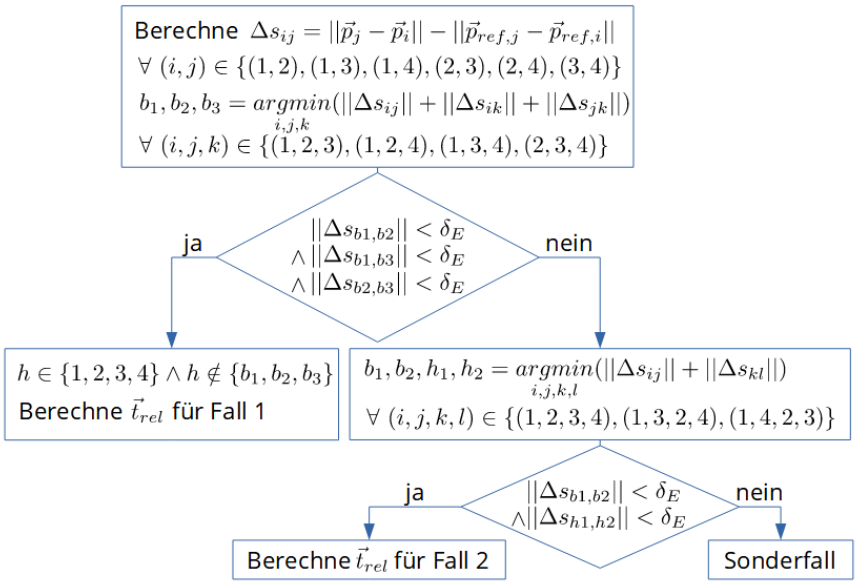


Abb. 7: Schema zur Fallunterscheidung für die Tetraederdeformation.

Darin muss der folgender Schwellwert δ_E für die Kantenlängenänderung festgelegt werden, ab dem eine Kante als deformiert angesehen wird. Der Fall eines undeformierten Tetraeders wird bei diesem Schema Fall 1 zugeordnet. Da die Verschiebungsvektoren mit Subvoxelgenauigkeit bestimmt werden können, liegt auch der Schwellwert δ_E im Subvoxelbereich. Für das gezeigte Beispiel wird $\delta_E = 0{,}2\ vx$ verwendet. Der Sonderfall tritt z.B. dann ein, wenn alle Kanten des Tetraeders als deformiert erkannt werden.

2.2.3 Berechnung des Deformationsvektors

Die Berechnung der Parameter $\vec{t}$, $\boldsymbol{R}$ und $\vec{t}_{rel}$ kann in einer Ausgleichung erfolgen und wird für jeden Tetraeder einzeln durchgeführt. Als Beobachtungen dienen die vier Koordinatenvektoren der Eckpunkte im deformierten Zustand. Die Referenzkoordinaten sind auf der rechten Seite der Gleichung enthalten und als Näherungswerte für die Rotationsparameter genügen meist die einer Nullrotation. Eine detaillierte Beschreibung zur Ausgleichung ist bei Liebold et al. (2020) zu finden.

Für Fall 1 ist es alternativ möglich, die Referenz- sowie deformierten Koordinaten der drei Punkte der Menge $M_1 = \{b_1, b_2, b_3\}$ zu verwenden, um die Parameter einer Starrkörpertransformation ($\vec{t}$ und $\boldsymbol{R}$) zu berechnen. Der Relativtranslationsvektor ergibt sich dann mit Hilfe des vierten Punktes der zweiten Menge zu:

$$\vec{t}_{rel} = \vec{p}_h - \vec{t} - \boldsymbol{R} \cdot \vec{p}_{ref,h} \tag{2}$$

$\vec{t}_{rel}$ wird als Deformationsvektor und dessen Betrag $\left\|\vec{t}_{rel}\right\|$ als skalares Deformationsmaß verwendet, welches die Einheit Voxel besitzt. Der Wert $\left\|\vec{t}_{rel}\right\|$ ist unabhängig von der Größe des Tetraeders. Abb. 8a zeigt eine farbkodierte Darstellung, wobei die Tetraeder nach ihrem Wert $\left\|\vec{t}_{rel}\right\|$ eingefärbt sind.

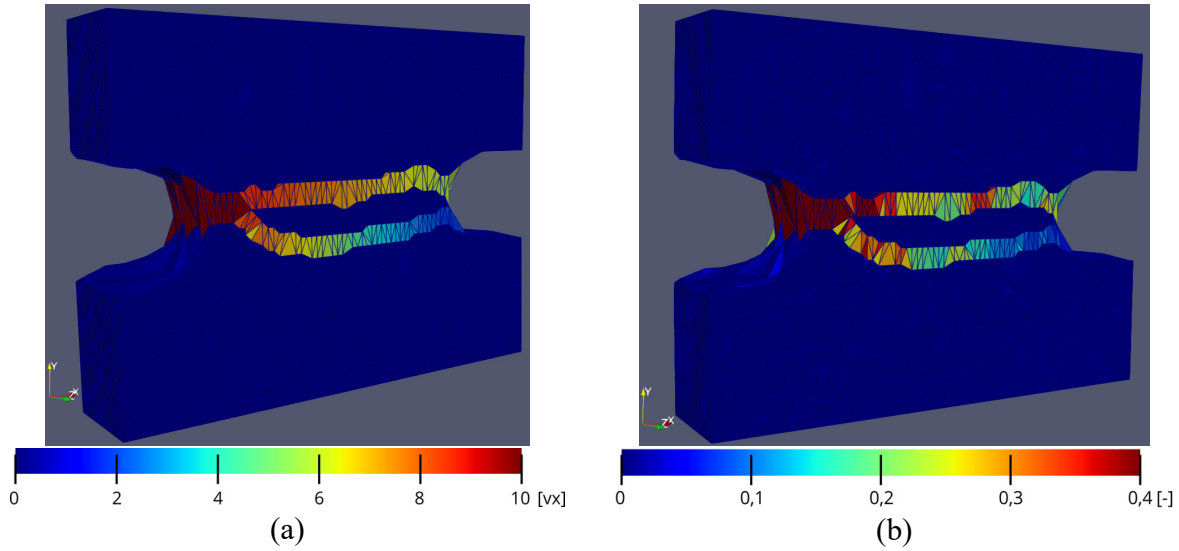


Abb. 8: a) farbkodierte Visualisierung für $\|\vec{t}_{rel}\|$; b) Visualisierung der Hauptdehnungen.

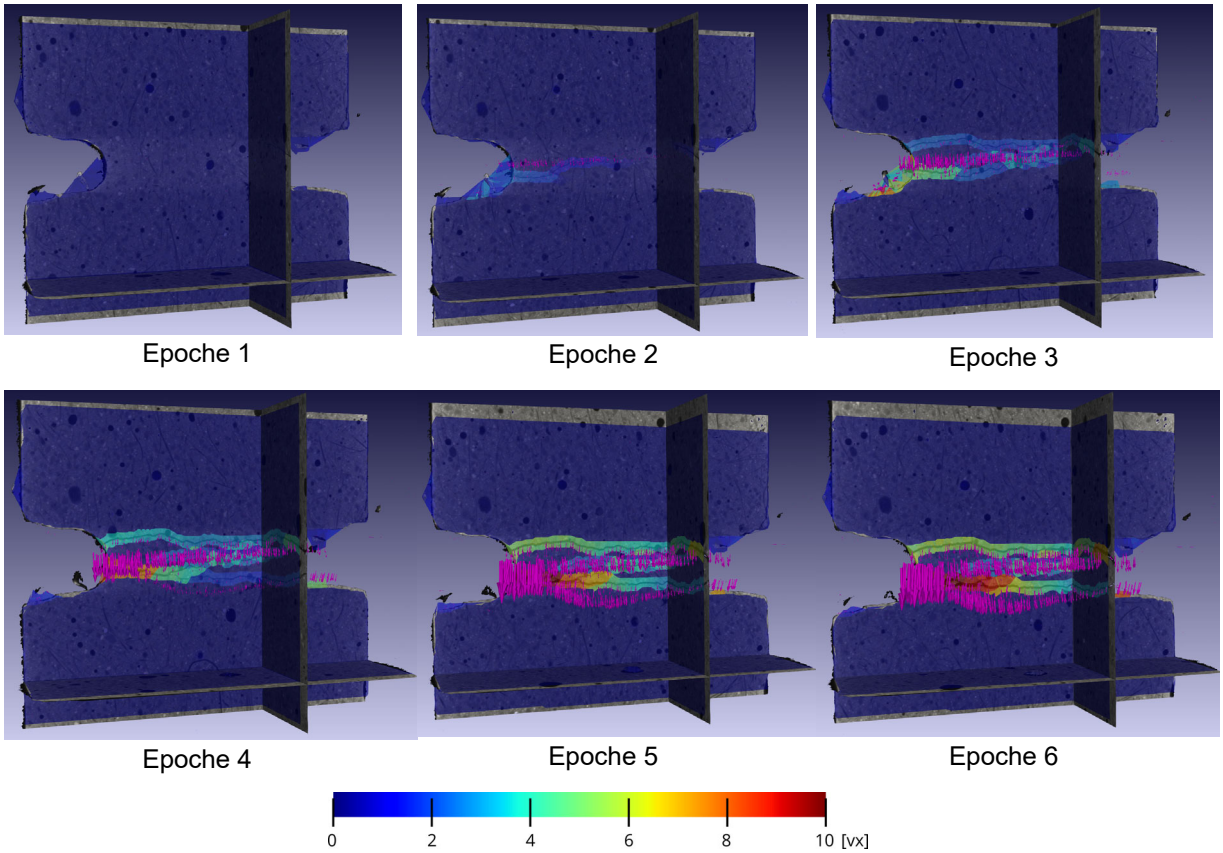


Abb. 9: Vektorfelder der Deformationsvektoren für die sechs deformierten Epochen. Die Längen der Pfeile sind mit dem Faktor 4 skaliert dargestellt.

In Abb. 8b ist eine farbkodierte Visualisierung mit Hauptdehnungen für die Tetraeder zu sehen. Die Berechnungsvorschrift für Dehnungen sind in LIEBOLD et al. (2021) zu finden. Im Unterschied

zu Abb. 8a ist hier zu erkennen, dass diese Dehnungswerte von der Tetraedergröße abhängen, während das für $\|\vec{t}_{rel}\|$ nicht gilt und ein homogenerer Farbverlauf entlang der Risse zu sehen ist. Die Deformationen lassen sich auch als Vektorfelder der Deformationsvektoren darstellen, siehe Abb. 9. Es sind auch jeweils drei Querschnitte durch die Volumendaten und die dazugehörigen transparenten Querschnitte durch die skalaren Deformationsfeld $\|\vec{t}_{rel}\|$ gezeigt.

2.2.4 Detektion von Rissen

Durch Anwendung des Schwellwertverfahren für die Werte $\|\vec{t}_{rel}\|$ können Risstetraeder detektiert werden: Gilt $\|\vec{t}_{rel}\| > \delta_T$, handelt es sich um einen Risskandidaten. Da die Verschiebungen mit Subvoxelgenauigkeit bestimmt werden können, liegt auch der Schwellwert im Subvoxelbereich. Beim vorgestellten Experiment wurde δ_T auf 0,3 vx gesetzt (Abb. 10). Durch eine Konnektivitätsanalyse ist es weiterhin möglich, Zusammenhangskomponenten der Risskandidaten zu bestimmen.

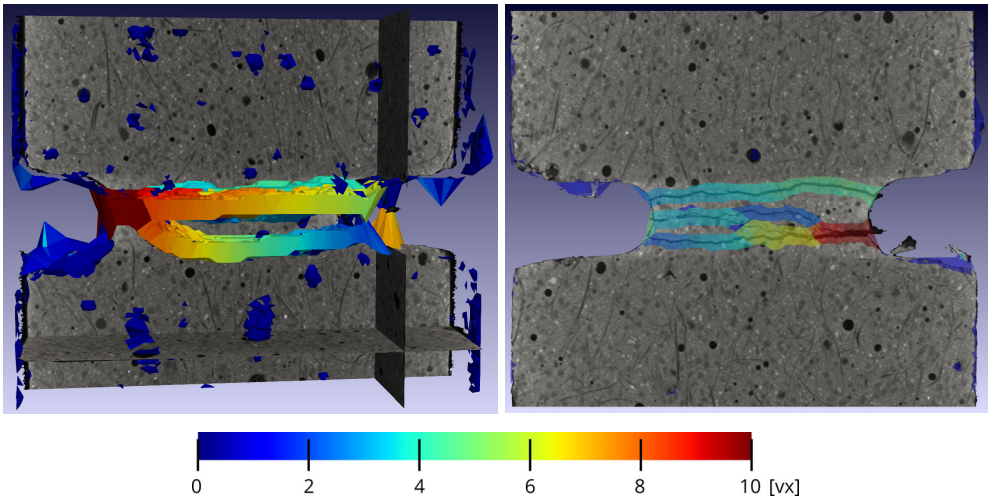


Abb. 10: Links: Farbkodierte Visualisierung für $\|\vec{t}_{rel}\|$ nach dem Schwellwertverfahren für Epoche 5; rechts: Querschnitt von der Rückseitenansicht.

2.2.5 Schätzung der Rissnormalen

Für jeden Risskandidaten ($\|\vec{t}_{rel}\| > \delta_T$) kann eine Normale geschätzt werden. Dazu werden zunächst für alle Risstetraeder deren Mittelpunkte $\vec{m}_{def,j}$ (j: Tetraederindex) berechnet:

$$\vec{m}_{def,j} = \begin{cases} (\vec{p}_{b1} + \vec{p}_{b2} + \vec{p}_{b3} + 3 \cdot \vec{p}_h)/6 \; für \; Fall \; 1 \\ (\vec{p}_{b1} + \vec{p}_{b2} + \vec{p}_{h1} + \vec{p}_{h2})/4 \; für \; Fall \; 2 \end{cases} \quad (3)$$

Dann werden die deformierten Nachbartetraeder des Risskandidaten (deformierter Kandidat) bestimmt. Als Nachbarn 1. Ordnung zählen Tetraeder mit mindestens einem gemeinsamen Eckpunkt. Nutzt man zusätzlich die Nachbarn 2. Ordnung wird die Liste um die deformierten Nachbarn der Nachbarn selbst erweitert. Anschließend wird eine Ausgleichsebene mit den Mittelpunkten der Nachbartetraeder und dem Kandidaten selbst berechnet. Der zugehörige Normalenvektor $\vec{n}$ mit dem Betrag $\|\vec{n}\| = 1$ entspricht dann der Rissnormale. Abb. 11 veranschaulicht das Prinzip der Normalenschätzung: Der Risskandidat ist rot dargestellt, die deformierten Nachbartetraeder

orange und undeformierte Nachbarn blau, die Ausgleichsebene türkis, der Normalenvektor als roter Vektor und die Mittelpunkte als magenta Kugeln.

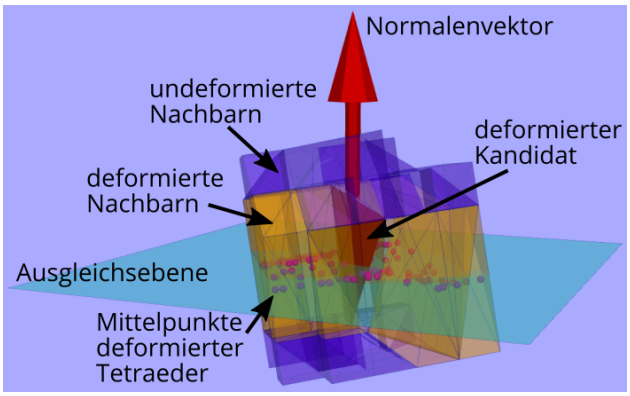


Abb. 11: Prinzip der Normalenschätzung.

Es folgt schließlich noch ein Test, ob die zunächst bestimmte Normale $\vec{n}'$ und der Deformationsvektor $\vec{t}_{rel}$ gegenläufig sind. Im Falle dessen wird die Normale umgekehrt:

$$\text{Falls } \vec{t}_{rel}{}^{T} \cdot \vec{n}' < 0, \text{ dann } \vec{n} = -\vec{n}', \text{ sonst } \vec{n} = \vec{n}' \tag{4}$$

Am Rand von Zusammenhangskomponenten deformierter Tetraeder bzw. bei Kreuzungen von Rissen führt die Schätzung möglicherweise zu systematischen Abweichungen.

2.2.6 Ableitung von Rissbreiten und Scherungsbewegungen

Der Deformationsvektor $\vec{t}_{rel}$ kann mit Kenntnis der Rissnormalenrichtung in einen Teil in diese Richtung $\vec{t}_{rel,n}$ und einen Teil senkrecht dazu $\vec{t}_{rel,s}$ zerlegt werden, siehe Abb. 12.

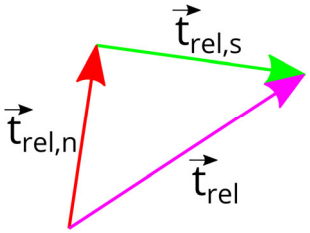


Abb. 12: Zerlegung des Deformationsvektors.

Die Normalenkomponente (in Rissnormalenrichtung) berechnet sich durch Projektion des Vektors $\vec{t}_{rel}$ auf die Rissnormale $\vec{n}$:

$$\vec{t}_{rel,n} = \frac{\vec{t}_{rel}{}^{T} \cdot \vec{n}}{\|\vec{n}\| \cdot \|\vec{t}_{rel}\|} \cdot \|\vec{t}_{rel}\| \cdot \frac{\vec{n}}{\|\vec{n}\|} = \vec{t}_{rel}{}^{T} \cdot \vec{n} \cdot \vec{n} \tag{5}$$

Der Vektor der Scherungsbewegung tangential zum Riss (Scherungskomponente) ergibt sich durch Subtraktion:

$$\vec{t}_{rel,s} = \vec{t}_{rel} - \vec{t}_{rel,n} \tag{6}$$

Die Rissbreite entspricht der Länge der Komponente $\vec{t}_{rel,n}$:

$$r = \left\|\vec{t}_{rel,n}\right\| = \vec{t}_{rel}{}^{T} \cdot \vec{n} \tag{7}$$

Die Rissbreite hängt maßgeblich von der Rissnormalen ab. Dabei können systematische Abweichungen wie oben erwähnt durch eine fehlerhafte Normalenschätzung bei Risskreuzungen bzw. bei Randtetraedern auftreten.
In Abb. 13a sind die Scherungs- als grüne Pfeile bzw. Normalenkomponenten als rote Pfeile dargestellt (Vergrößerungsfaktor: 10). Wo die Zugkraftrichtung senkrecht zum Riss steht, ist eine Übereinstimmung der Normalenkomponenten und der Deformationsvektoren zu erkennen. An der linken Seite der Verjüngung, an der die Risse zusammenlaufen, werden größere Scherungsanteile berechnet, weil der Riss an diesen Stellen auch nicht senkrecht zur Zugrichtung verläuft.

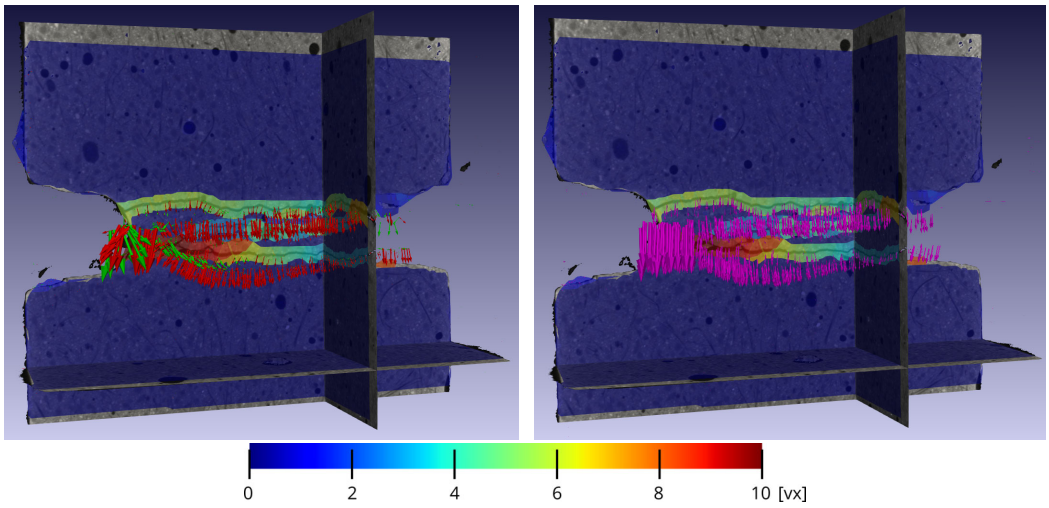


Abb. 13: Links: Normalen- in rot und Scherungskomponenten in grün (Vergrößerungsfaktor 4), dazu drei Volumenquerschnitte mit den dazugehörigen transparenten Querschnitten durch das farbkodierte Deformationsfeld; rechts: Deformationsvektoren für Epoche 6.

3 Fazit & Ausblick

Der Beitrag zeigte am Beispiel eines in-situ-Zugversuchs an einer spröden SHCC-Probe, wie eine Deformations- bzw. Rissanalyse bei 3D-Volumendaten mehrerer Laststufen durchgeführt werden kann. Nach Berechnung eines subvoxel-genauen Verschiebungsvektorfeldes erfolgte eine Vernetzung zu einem Tetraedernetz. Es wurde eine Berechnungsvorschrift für einen subvoxel-genauen Deformationsvektor zur Rissdetektion vorgestellt, von dem sich Rissbreiten und Scherungsbewegungen ableiten lassen. In zukünftige Arbeiten könnte eine Modellerweiterung um eine Relativrotation bzw. eine intensivere Genauigkeitsbetrachtung eine Rolle spielen.

Danksagung

Die gezeigte Forschungsarbeit wurde gefördert durch die Deutsche Forschungsgemeinschaft (DFG – SFB/TRR280, Projektnummer: 417002380). Weiterhin danken wir dem Institut für Baustoffe der Technischen Universität Dresden, dem Department of Chemical and Materials Engineering der Pontificía Universidade Católica do Rio de Janiero und der Bundesanstalt für Materialforschung und -prüfung für die bereitgestellten Datensätze.

4 Literaturverzeichnis

BARAZZETTI, L. & SCAIONI, M., 2010: Development and Implementation of Image-based Algorithms for Measurements of Deformations in Material Testing. Sensors, **10**(8), 7469-7495, https://doi.org/10.3390/s100807469.

BAY, B. K., SMITH, T. S., FYHRIE, D. P. & SAAD, M., 1999: Digital Volume Correlation: Three-dimensional Strain Mapping Using X-ray Tomography. Experimental Mechanics, **39**, 217-226, https://doi.org/10.1007/BF02323555.

BENNING, W., LANGE, J., SCHWERMANN, R., EFFKEMANN, C. & GÖRTZ, S., 2004: Monitoring crack origin and evolution at concrete elements using photogrammetry. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XXXV**-B5, 678-683.

FRASER, C. S. & RIEDEL, B., 2000: Monitoring the thermal deformation of steel beams via vision metrology. ISPRS J. Photogramm. Remote Sens., **55**(4), 268-276, https://doi.org/10.1016/S0924-2716(00)00024-1.

GEERS, M. G. D., DE BORST, R. & BREKELMANS, W. A. M., 1996: Computing strain fields from discrete displacement fields in 2D-solids. International Journal of Solids and Structures, **33**(29), 4293-4307. https://doi.org/10.1016/0020-7683(95)00240-5.

HAMPEL, U. & MAAS, H.-G., 2009: Cascaded image analysis for dynamic crack detection in material testing. ISPRS J. Photogramm. Remote Sens., **64**(4), 345-350. https://doi.org/10.1016/j.isprsjprs.2008.12.006.

KOSCHITZKI, R. & MAAS, H.-G., 2012: Vergleich photogrammetrischer und akustischer Messverfahren zur Rissdetektion bei Belastungsversuchen im Stahlbetonbau. AVN, **119**(1), 3-10.

LANGE, J., BENNING, W. & SIERING, K., 2006: Crack detection at concrete construction units from photogrammetric data using image processing procedures. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **36**(7), 493-496.

LIEBOLD, F., LORENZONI, R., CUROSU, I., LÉONARD, F., MECHTCHERINE, V., PACIORNIK, S. & MAAS, H.-G., 2021: 3D least squares matching applied to micro-tomography data. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XLIII**-B2-2021, 533-539. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-533-2021.

LIEBOLD, F. & MAAS, H.-G., 2016: Advanced spatio-temporal filtering techniques for photogrammetric image sequence analysis in civil engineering material testing. ISPRS J. Photogramm. Remote Sens., **111**, 13-21. https://doi.org/10.1016/j.isprsjprs.2015.10.013.

LIEBOLD, F. & MAAS, H.-G., 2018: Sub-pixel accuracy crack width determination on concrete beams in load tests by triangle mesh geometry analysis. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., **IV**-2, 193-200. https://doi.org/10.5194/isprs-annals-IV-2-193-2018.

LIEBOLD, F. & MAAS, H.-G., 2020: Strategy for Crack Width Measurement of Multiple Crack Patterns in Civil Engineering Material Testing Using a Monocular Image Sequence Analysis. PFG - J. Photogramm. Remote Sens. Geoinf. Sci., **88**(3), 219-238. https://doi.org/10.1007/s41064-020-00103-2.

LIEBOLD, F., MAAS, H.-G. & DEUTSCH, J., 2020: Photogrammetric determination of 3D crack opening vectors from 3D displacement fields. ISPRS J. Photogramm. Remote Sens., **164**, 1-10. https://doi.org/10.1016/j.isprsjprs.2020.03.019.

LIEBOLD, F., MAAS, H.-G. & HERAVI, A. A., 2019: Crack width measurement for non-planar surfaces by triangle mesh analysis in civil engineering material testing. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XLII**-2/W18, 107-113. https://doi.org/10.5194/isprs-archives-XLII-2-W18-107-2019.

LORENZONI, R., CUROSU, I., LÉONARD, F., PACIORNIK, S., MECHTCHERINE, V., SILVA, F. A. & BRUNO, G., 2020: Combined mechanical and 3D-microstructural analysis of strain-hardening cement-based composites (SHCC) by in-situ X-ray microtomography. Cement and Concrete Research, **136**, 106139. https://doi.org/10.1016/j.cemconres.2020.106139.

MAAS, H.-G., STEFANIDIS, A. & GRÜN, A., 1994: From pixels to voxels: tracking volume elements in sequences of 3D digital images. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XXX**-3/2, 539-546. https://doi.org/10.1117/12.182861.

Mowing Detection from Combined Sentinel-1, Sentinel-2, and Landsat 8 Time Series on Fallow Cropland with Transfer Learning

FELIX LOBERT[1,2], NORBERT RÖDER[3], ALEXANDER GOCHT[1], MARCEL SCHWIEDER[1,2] & STEFAN ERASMI[1]

Abstract: With the monitoring of management practices in agriculture we can assess the influence on the environment and evaluate the implementation of policies and guidelines. The timing of mowing or mulching events on fallow cropland is a critical factor in assessing the impact on biodiversity and a core environmental measure of the EU's Common Agricultural Policy (CAP). Dense time series of remote sensing data have already proven to be valuable input for capturing management practices on grassland. Here we present a transfer learning approach using a 1D convolutional neural network trained on grassland management data to predict management practices on fallow cropland. Our results demonstrate the potential of the approach to provide valuable insights for future CAP decision-making.

1 Introduction

The area-wide monitoring of agricultural management practices is an important task that does not only help to assess the impact of agriculture on biodiversity but also to evaluate the implementation of policies and guidelines.

The first management of fallow cropland within a year is an example of a practice whose timing is both related to the agricultural use intensity and increasingly relevant in the context of the growing environmental ambitions of the EU's Common Agricultural Policy (CAP) (EUROPEAN PARLIAMENT 2021). Fallow cropland is usually mown or mulched, whereby after mulching the cut biomass is further shredded and remains on the field. For simplicity, we will use the word mowing in the following.

Dense time series of remote sensing data have already proven to be a valuable data basis for the detection of management practices on the field level, including mowing events on grasslands (e.g., SCHWIEDER et al. 2022; DE VROEY et al. 2021; REINERMANN et al. 2020). The cutting of grass or similar vegetation, which means a removal of standing biomass, leads to abrupt changes in remote sensing signals. This makes them a good basis for the development of automatic detection algorithms. Yet, the various parameters and indices that can be derived from optical and radar imagery both contain complementary and redundant information (HOLTGRAVE et al. 2020). The selection and combination of these indices and parameters, therefore, play an important role and can improve but also impair the performance of mowing detection algorithms. This was recently

[1] Thünen Institute of Farm Economics, Bundesallee 63, D-38116 Braunschweig, E-Mail: [felix.lobert, alexander.gocht, marcel.schwieder, stefan.erasmi]@thuenen.de

[2] Earth Observation Lab, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

[3] Thünen Institute of Rural Studies, Bundesallee 64, D-38116 Braunschweig, E-Mail: norbert.roeder@thuenen.de

demonstrated by LOBERT et al. (2021), who systematically investigated the suitability of different combinations of Sentinel-1 (S1), Sentinel-2 (S2), and Landsat 8 (L8) time series to determine the occurrence, frequency, and date of mowing events.
The objective of our research is to investigate the suitability of remote sensing time series for monitoring the timing of mowing practices on fallow cropland at the parcel level, which has been hardly studied. The actual timing of the first management of fallow cropland is of high importance since fallows serve as an important refuge for many species including ground-nesting birds (VAN DE POEL & ZEHM 2014). Area-wide information can therefore help to evaluate the biodiversity impact of agricultural policy measures.
Reference data for training models on fallow management, however, are not available. Since we expected a similar remote sensing pattern of fallows and grasslands due to mowing, we decided to use a transfer-learning approach. Because fallows are allowed to be mown for the first time even later in the year than grasslands, we expected a remarkable signal here. We, therefore, trained an algorithm to detect mowing events on grassland parcels for which we have reference data available and applied it to fallows afterward. A one-dimensional convolutional neural network (1D-CNN) was the algorithm of choice for our study due to its high performance on time series related tasks (WANG et al. 2017).

2 Study area and data

2.1 Study area and reference data

We examined two different types of land use in our analysis: grassland parcels with reference data to train the model and fallow parcels where we applied it. The data from the used grassland parcels are obtained from the Biodiversity Exploratories project that includes standardized grassland parcels in three parts of Germany with extensive information on management practices provided, including mowing events (FISCHER et al. 2010; VOGT et al. 2019). We used data from 2017 to 2019 which yielded a reference dataset of 257 mowing events on over 60 different parcels.
The study area chosen for the application of the model was the federal state of Brandenburg. Here, the boundaries of all agricultural parcels are contained in the Geospatial Aid Application data together with additional information, including the cultivated crop type and whether a field was declared as fallow in a given year in order to receive subsidies. The data is made openly available by the Ministry of Agriculture, Environment and Climate Protection Brandenburg. We selected the dataset from 2019, containing 8,255 fallows that are reported as ecological focus areas. A negative buffer of 20 m was applied to reduce edge effects. After excluding parcels smaller than 1 ha, 3,107 parcels were left for analysis.

2.2 Remote sensing imagery

2.2.1 SAR imagery

Two SAR parameters were used in this study, the interferometric coherence and the gamma naught ($\gamma 0$) backscatter coefficient. Time series of coherence were observed to show significantly higher values after the cutting of grass on a parcel due to a temporally more stable and coherent scattering of short grass compared to long grass. The reaction of SAR backscatter time series to mowing

events was observed to be indifferent, yet it showed to be of value for mowing detection LOBERT et al. (2021).
The used data were acquired by the C-Band S1 constellation with dual-polarization in VV (vertical transmit and vertical receive) and VH (vertical transmit and horizontal receive) in the interferometric wide swath mode (IW). For the estimation of the interferometric coherence, both amplitude and phase information from the Single Look Complex (SLC) product type are needed. To process the SAR backscatter coefficient, the Ground Range Detected (GRD) product type, containing only amplitude information, is sufficient. Since the S1 constellation consists of two satellites (S1A & S1B), the interval between acquisitions from the same orbit is 6 days. We chose a single relative orbit for each parcel, which ensures a consistent acquisition geometry. For each orbit, we used all scenes for the months March to November from 2017 to 2019 for the grassland sites (414 scenes) and from 2019 for the fallows (180 scenes). We accessed the S1 data through the Copernicus Data and Exploitation Platform - Deutschland (CODE-DE; Benz et al. 2020).
We computed the coherence data using the Graph Processing Tool (GPT) of the Sentinel Application Platform (SNAP). First, we formed image pairs from each SLC acquisition and its predecessor 6 days earlier. Then, we updated the exact orbit position and coregistered the image pair using back-geocoding. Afterward, we estimated the coherence using a window size of 2 pixels in azimuth and 10 in range direction to obtain approximately square pixel size. Finally, we merged the individual bursts of the coherence image followed by a terrain correction to a spatial resolution of 20 m using the Shuttle Radar Topographic Mission (SRTM) 3 arc-second global digital elevation model (DEM; FARR et al. 2007).
To process γ^0, we first applied border and thermal noise removal to the GRD acquisitions. We then calibrated and radiometrically flattened the data to obtain γ^0 backscatter coefficient, which represents the ratio between the incident power and the scattered power for a reference area that is perpendicular to the line of sight from the sensor to an ellipsoidal model of the ground surface (Small 2011). The images were then terrain corrected and resampled to a spatial resolution of 10 m using the SRTM DEM. Finally, we converted γ^0 from linear scale to dB. The GPT was again used for these steps. To exploit the information content of the backscattered signal in both polarizations, we calculated the backscatter cross-ratio (CR):

$$CR = \gamma^0_{VH}[dB] - \gamma^0_{VV}[dB] \quad (1)$$

which was observed to be strongly affected by structural changes in crops like winter cereals (HOLTGRAVE et al. 2020; VREUGDENHIL et al. 2018). Moreover, SCHLUND & ERASMI (2020) reported that the CR stays relatively stable in dense time series over longer periods when agricultural areas are observed, because the effect of terrain and soil properties on the radar signal is similar for both polarizations and, thus, the ratio reduces the impact of these factors on the CR signal.

2.2.2 Optical imagery

Optical vegetation indices (e.g., NDVI) are a common proxy for vegetation vitality and biomass. They were observed to show drops in time series related to mowing events and the resulting removal of green biomass and were therefore frequently used in mowing detection algorithms REINERMANN et al. (2020).

For our analysis, we used all available scenes from S2 and L8 that cover the study area and have a cloud coverage of less than 75%. We then corrected all data for radiometric and geometric effects, clouds, and their shadows using FORCE (FRANTZ 2019) and the Fmask algorithm (FRANTZ et al. 2018; ZHU et al. 2015; ZHU & WOODCOCK 2012). Then, we calculated the NDVI as follows:

$$NDVI = \frac{(\rho NIR - \rho RED)}{(\rho NIR + \rho RED)} \quad (2)$$

where ρNIR is the reflectance in the near-infrared band and ρRED is the measured reflectance in the red band of the respective satellite (TUCKER 1979).

3 Methods

3.1 Time series composition

Our study based on the parcel level as the unit of prediction. We, therefore, derived the median values of all coherence, CR, and NDVI observations for each grassland and fallow parcel using the respective parcel boundaries. An equidistant time series dataset is a technical requirement for the 1D-CNN architecture we used in our study. Since the SAR data already come with an equidistant time interval of 6 days, we matched the NDVI time series with the dates of the SAR acquisitions using linear interpolation. Furthermore, the time series were standardized by parcel and year by subtracting the mean and dividing by the standard deviation.

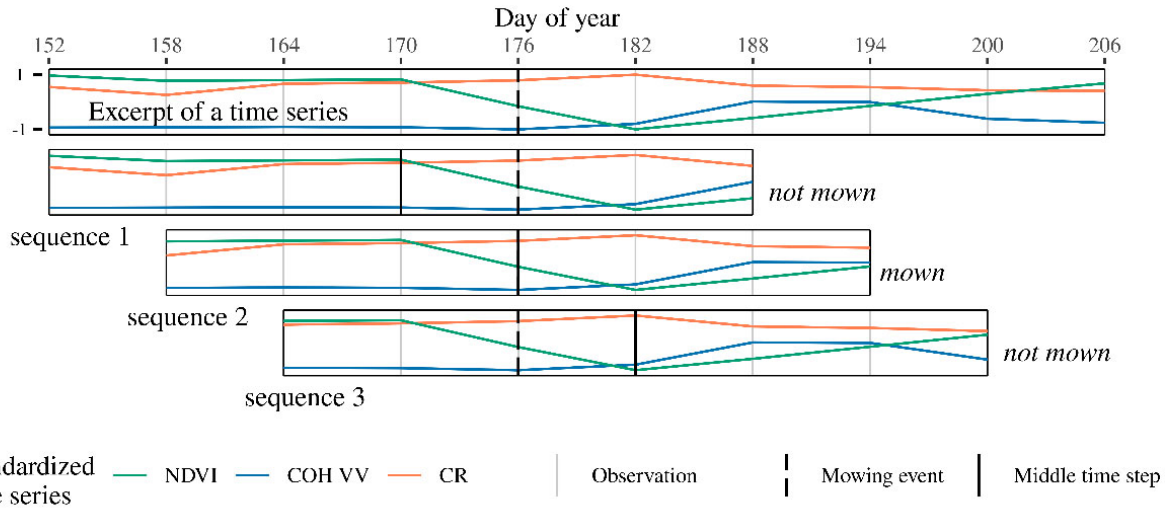


Fig. 1: Schema of the moving window approach applied to an exemplary excerpt of the time series of a grassland parcel in 2019. Depending on whether there was a mowing event on the middle time step of a resulting sequence, it is labeled as mown or not mown.

We then decomposed the time series of each parcel into short overlapping sequences, using a moving window approach which allowed us to use a supervised classification algorithm on our dataset. The approach includes the stepwise moving of a window over the time series with a given size and creating new sequences out of the values that fall into the window at each step.

Subsequently, we labeled the sequences as mown or not mown depending on whether a mowing event occurred at the midpoint of the sequence (Fig. 1). The size of the window was five observations before and after the middle time step, respectively, resulting in a total length of 11 observations. For each new sequence, it was moved by one time step (6 days). This ensured, that every observation is the middle time step of a sequence and can potentially be classified as a mowing event.

3.2 Deep Learning model

Machine learning algorithms are widely used for classification or regression tasks in remote sensing applications. In the last years, especially Artificial Neural Networks (ANN) started to be applied more and more and achieved competitive results. Convolutional Neural Networks (CNN) are a form of ANNs that perform well on sequential data since they convolve the input data in a way, that consecutive values are combined, and higher-level features are extracted. This makes them a frequently applied method in the domain of deep learning (KATTENBORN et al. 2021).

3.2.1 Implementation and training

Here, we used a 1D-CNN proposed by WANG et al. (2017) and adapted by LOBERT et al. (2021) to classify the labeled sequences into mown and not mown. We implemented the 1D-CNN with Keras (CHOLLET & KERAS-TEAM 2015) and TensorFlow (ABADI et al. 2016) as backend using the R interface to Keras (ALLAIRE & CHOLLET 2020). The original model was developed to classify univariate sequences into multiple classes. To match the given classification problem, LOBERT et al. (2021) adapted the model to run on multivariate input data, i.e., the different time series, and give binary classification output in combination with a complexity reduction and architecture optimization.

The model used here consists of two convolutional layers with kernel sizes five and three and filter numbers 128 and 256, respectively. The output of the convolutional layers was zero-padded to remain the same size as the input. Batch-normalization and a rectified linear unit (ReLU) activation function were applied after each layer. The ReLU function converts all negative values to zero, while positive values are preserved. The two convolutional layers were followed by a global average pooling and finally a single densely connected and sigmoid-activated neuron to give the final output (Fig. 2).

LOBERT et al. (2021) conducted an in-depth feature selection and hyperparameter optimization for the presented model architecture, using identical training and reference data. We, therefore, decided to use their proposed best setup and refer the reader to the publication for more detailed information. In our study, we randomly chose 15% of the grassland parcels as test samples and trained our model on the remaining parcels.

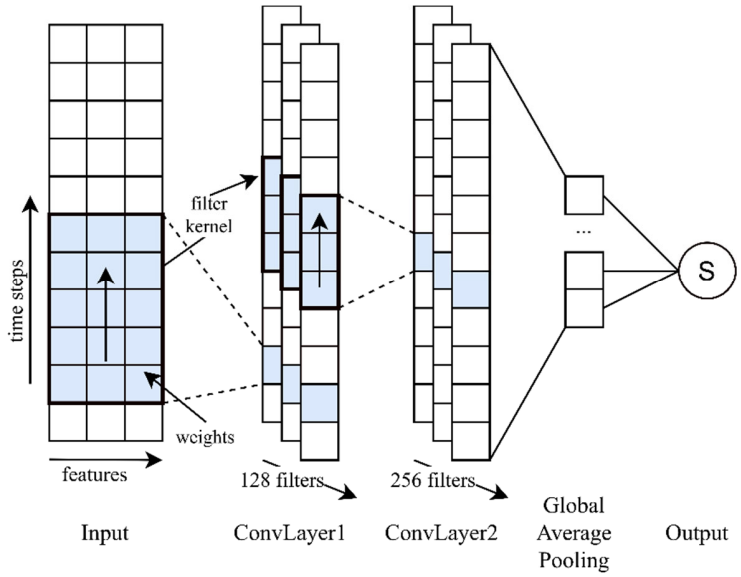


Fig. 2: Schematic architecture of the 1D-CNN by Lobert et al. (2021). The model consists of an input layer with three features of 11 timesteps. The input layer is followed by two convolutional Layers with kernel sizes five and three and filter numbers 128 and 256, respectively. Next is a global pooling layer, summarizing each filter to its average. The last layer is a densely connected and sigmoid-activated output neuron.

3.2.2 Method transfer

Since only reference data for the grassland areas were available, we had to train and validate our model on these data and transfer the model to the fallows afterward. This transfer is based on the assumption that mowing events on grassland and fallows result in similar signals in the remote sensing time series.

Since we were interested in a detection algorithm that would show possible non-compliance with subsidy requirements, the accuracy of our model plays a critical role. We, therefore, measured our accuracy on the grassland test set with two metrics: the recall, stating the share of reference events that our model was able to detect, and the precision, giving the share of the mowing events predicted by our model that are actually correct. The sigmoid activation that is applied to our model output results in a continuous value between 0 and 1 for each sample. This value can be treated as the certainty of the model prediction or, in this case, the probability of a predicted sample to be mown. Usually, a threshold of 0.5 is used to classify the model output into mown (above threshold) and not mown (below threshold). Raising this threshold results in increased precision of the detections, while the recall decreases at the same time, and vice versa. In this trade-off between precision and recall, we decided that precise predictions are more important for our application, even though fewer events may be detected in return.

We, therefore, analyzed several threshold values between 0.5 and 1 regarding the resulting precision and recall on the grassland data. Based on this relationship, we were able to estimate the best threshold to make predictions on the fallows with high precision. With the chosen threshold, we applied the model on the fallows and measured the share of parcels where a mowing event was detected before the 15th of July, which could mean a possible non-compliance with the subsidy requirements for fallow cropland that is declared as ecological focus area in Germany.

4 Results and discussion

Our results show that increasing the threshold for the model output to be classified as mown continuously decreased the recall while improving the precision of the model predictions at the same time (Fig. 3). Above a threshold of 0.93, the precision reaches 1, which means that the model no longer predicted false positives when applied to the grassland test data. Yet, this also comes with a recall rate that has dropped to 0.37, meaning that nearly two-thirds of the mowing events are missed. However, with this threshold, we could expect to make the most precise predictions when applying the model to the fallows and result in the least false positives. With this threshold, 295 from all 3,107 analyzed fallows showed a detected mowing event before the 15th of July, which is about 9.5% of the parcels.

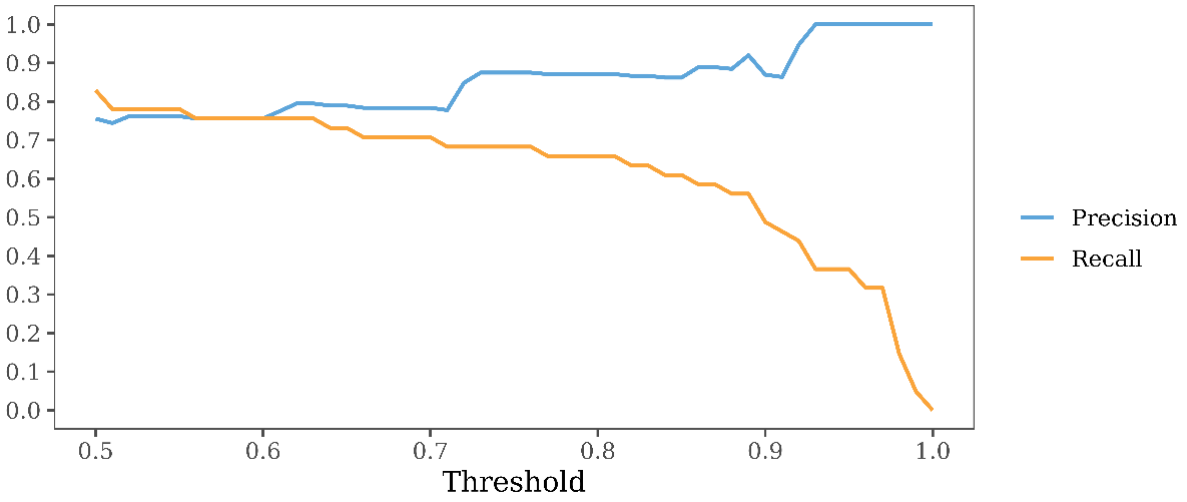


Fig. 3: Precision and recall measured on the grassland test data depending on different threshold values of the sigmoid activated model output to be classified as mown.

The time series and model predictions for four exemplary fallows with predictions before the 15th of July are shown in Fig. 4. All examples show drops in the NDVI at or prior to the detected event. They also show the lowest NDVI value during the observed period with more than one standard deviation below the mean shortly after the detection. This supports the hypothesis that removal of green biomass by mowing has occurred here. However, in the upper left and lower right examples, the NDVI starts to decrease more than one month before the detected cut. This could be induced by withering vegetation. This hypothesis is further supported by very warm and dry conditions in Brandenburg in 2019 and the fact that fallows commonly occur on marginal sites.
The coherence signal can be observed to increase after the detected events. The two fallows on top show high and sharp increases. This could suggest a correct detection on the upper left fallow, despite the early NDVI decline. The lower examples show delayed and weaker signals. These differences could be explained by different practices that partially leave the cut material on the plot like mulching. The CR time series does not show a consistent pattern throughout the four examples.

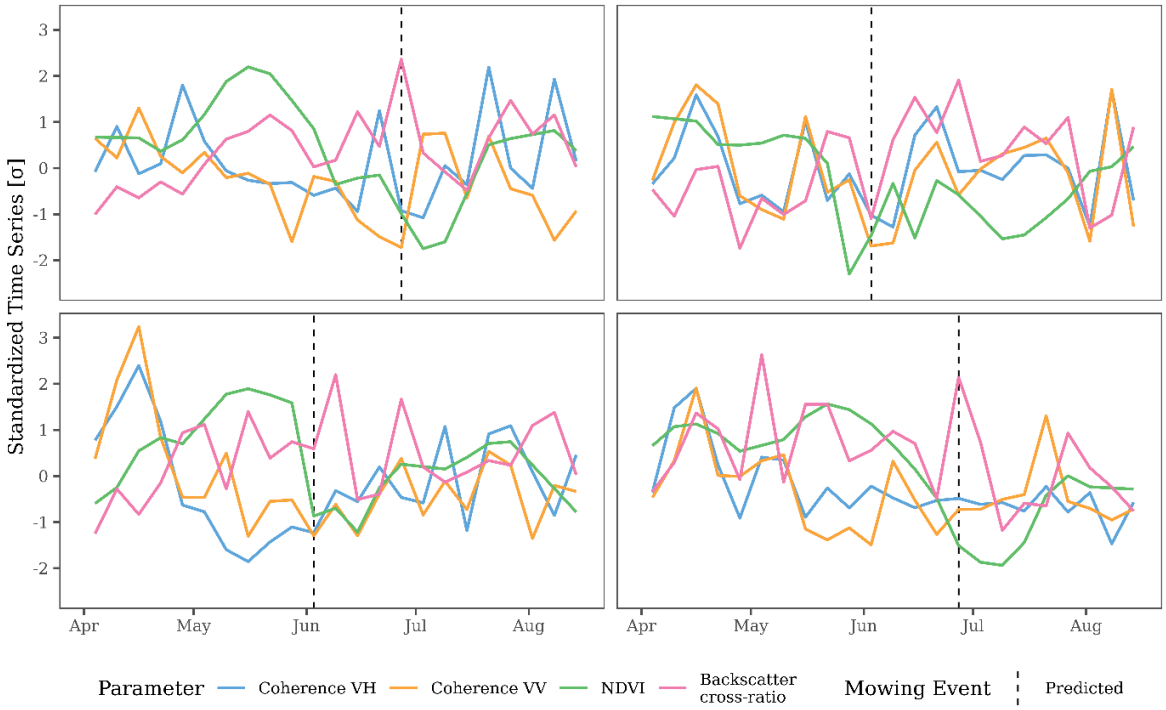


Fig. 4: Time series and predictions made by the model for four exemplary fallows in Brandenburg in 2019 with predicted mowing before the 15th of July with a threshold of 0.93.

These findings of time series patterns related to potential mowing events go in line with the findings of others (e.g., DE VROEY et al. 2021; SCHWIEDER et al. 2022). However, we must be aware that fallow cropland, unlike traditional meadows and pastures, is much more heterogeneous and has a wide variety of vegetation types and species that are either sown or grow spontaneously. Compared to the quite simple phenology of grass in managed meadows, that is mowed several times a year and then grows up again, we are dealing here with a probably much more complex annual course that can contain sudden events that may easily be mistaken for mowing events. Therefore, even if randomly checked detections may seem plausible, the results of the transfer learning should be handled with caution. We must be aware that a model detection does not necessarily mean non-compliance but can be a valuable indicator to plan on-the-spot checks.

5 Conclusion

Our transfer learning approach presented here was able to detect events in the remote sensing time series of fallow cropland that closely resemble mowing events on grasslands. Based on the hypothesis we made at the beginning of this article, these could be actual management measures on the corresponding parcels. Any in-depth interpretation of these results, however, requires thorough validation against reference data on fallow land management and is a pre-requisite for further research.

The results indicate that management practices could have occurred on fallow cropland during periods in which management is prohibited on parcels that were reported as ecological focus areas.

Appropriate reference data and further development of the model could solidify these results and provide a robust overview of the temporal distribution of management practices on fallows. Hereby, this study could have provided the basis for an improved evaluation of CAP measures in the future.

6 References

ABADI, M., BARHAM, P., CHEN, J., CHEN, Z., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S., IRVING, G., ISARD, M., KUDLUR, M., LEVENBERG, J., MONGA, R., MOORE, S., MURRAY, D.G., STEINER, B., TUCKER, P., VASUDEVAN, V., WARDEN, P., WICKE, M., YU, Y. & ZHENG, X., 2016: Tensorflow: A system for large-scale machine learning. 12th USENIX Symposium on Operating Systems Design and Implementation, 265-283.

ALLAIRE, J. & CHOLLET, F., 2020: keras: R Interface to "Keras". R package version 2.3.0.0, https://cran.r-project.org/package=keras.

BENZ, U., BANOVSKY, I., CESARZ, A & SCHMIDT, M., 2020: CODE-DE Portal Handbook, Version 2.0, DLR.

CHOLLET, F. & KERAS-TEAM, 2015: Keras. Available at: https://github.com/fchollet/keras.

DE VROEY, M., RADOUX, J. & DEFOURNY, P., 2021: Grassland mowing detection using Sentinel-1 time series: Potential and limitations. Remote Sensing, **13**, 1-19, https://doi.org/10.3390/rs13030348.

EUROPEAN PARLIAMENT & COUNCIL OF THE EUROPEAN UNION, 2021: Regulation (EU) 2021/2115 of the European Parliament and of the Council. Off. J. Eur. Union 435/1-435/186.

FARR, T.G., ROSEN, P.A., CARO, E., CRIPPEN, R., DUREN, R., HENSLEY, S., KOBRICK, M., PALLER, M., RODRIGUEZ, E., ROTH, L., SEAL, D., SHAFFER, S., SHIMADA, J., UMLAND, J., WERNER, M., OSKIN, M., BURBANK, D. & ALSDORF, D.E., 2007: The shuttle radar topography mission. Rev. Geophys., **45**, https://doi.org/10.1029/2005RG000183.

FISCHER, M., BOSSDORF, O., GOCKEL, S., HÄNSEL, F., HEMP, A., HESSENMÖLLER, D., KORTE, G., NIESCHULZE, J., PFEIFFER, S., PRATI, D., RENNER, S., SCHÖNING, I., SCHUMACHER, U., WELLS, K., BUSCOT, F., KALKO, E.K.V., LINSENMAIR, K.E., SCHULZE, E.D. & WEISSER, W.W., 2010: Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. Basic Appl. Ecol., **11**, 473-485, https://doi.org/10.1016/j.baae.2010.07.009.

FRANTZ, D., 2019: FORCE-Landsat + Sentinel-2 analysis ready data and beyond. Remote Sensing, **11**(9), 1124, https://doi.org/10.3390/rs11091124.

FRANTZ, D., HAß, E., UHL, A., STOFFELS, J. & HILL, J., 2018: Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sens. Environ., **215**, 471-481, https://doi.org/10.1016/j.rse.2018.04.046.

HOLTGRAVE, A.-K., RÖDER, N., ACKERMANN, A., ERASMI, S. & KLEINSCHMIT, B., 2020: Comparing Sentinel-1 and -2 Data and Indices for Agricultural Land Use Monitoring. Remote Sens. 12, 2919, https://doi.org/10.3390/rs12182919.

KATTENBORN, T., LEITLOFF, J., SCHIEFER, F. & HINZ, S., 2021: Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS J. Photogramm. Remote Sensing, **173**, 24-49, https://doi.org/10.1016/j.isprsjprs.2020.12.010.

LOBERT, F., HOLTGRAVE, A.-K., SCHWIEDER, M., PAUSE, M., VOGT, J., GOCHT, A. & ERASMI, S., 2021: Mowing event detection in permanent grasslands: Systematic evaluation of input features from Sentinel-1, Sentinel-2, and Landsat 8 time series. Remote Sens. Environ., **267**, 112751, https://doi.org/10.1016/j.rse.2021.112751.

REINERMANN, S., ASAM, S. & KUENZER, C., 2020: Remote Sensing of Grassland Production and Management - A Review. Remote Sensing, **12**, 1949, https://doi.org/10.3390/rs12121949.

SCHLUND, M. & ERASMI, S., 2020: Sentinel-1 time series data for monitoring the phenology of winter wheat. Remote Sens. Environ., **246**, 111814, https://doi.org/10.1016/j.rse.2020.111814.

SCHWIEDER, M., WESEMEYER, M., FRANTZ, D., PFOCH, K., ERASMI, S., PICKERT, J., NENDEL, C. & HOSTERT, P., 2022: Mapping grassland mowing events across Germany based on combined Sentinel-2 and Landsat 8 time series. Remote Sens. Environ., 112795, https://doi.org/10.1016/j.rse.2021.112795.

SMALL, D., 2011: Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sensing, **49**, 3081-3093, https://doi.org/10.1109/TGRS.2011.2120616.

TUCKER, C.J., 1979: Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ., **8**, 127-150, https://doi.org/10.1016/0034-4257(79)90013-0.

VAN DE POEL, D. & ZEHM, A., 2014: Die Wirkung des Mähens auf die Fauna der Wiesen - Eine Literaturauswertung für den Naturschutz. Handb. Naturschutz und Landschaftspfl., **36**, 1-19, https://doi.org/10.1002/9783527678471.hbnl2015001.

VOGT, J., KLAUS, V., BOTH, S., FÜRSTENAU, C., GOCKEL, S., GOSSNER, M., HEINZE, J., HEMP, A., HÖLZEL, N., JUNG, K., KLEINEBECKER, T., LAUTERBACH, R., LORENZEN, K., OSTROWSKI, A., OTTO, N., PRATI, D., RENNER, S., SCHUMACHER, U., SEIBOLD, S., SIMONS, N., STEITZ, I., TEUSCHER, M., THIELE, J., WEITHMANN, S., WELLS, K., WIESNER, K., AYASSE, M., BLÜTHGEN, N., FISCHER, M. & WEISSER, W., 2019: Eleven years' data of grassland management in Germany. Biodiversity Data Journal, 7: e36387, https://doi.org/10.3897/bdj.7.e36387.

VREUGDENHIL, M., WAGNER, W., BAUER-MARSCHALLINGER, B., PFEIL, I., TEUBNER, I., RÜDIGER, C. & STRAUSS, P., 2018: Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. Remote Sensing, **10**, 1-19, https://doi.org/10.3390/rs10091396.

WANG, Z., YAN, W. & OATES, T., 2017: Time series classification from scratch with deep neural networks: A strong baseline. Proc. Int. Jt. Conf. Neural Networks, 1578-1585, https://doi.org/10.1109/IJCNN.2017.7966039.

ZHU, Z. & WOODCOCK, C.E., 2012: Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ., **118**, 83-94, https://doi.org/10.1016/j.rse.2011.10.028.

ZHU, Z., WANG, S. & WOODCOCK, C.E., 2015: Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. Remote Sens. Environ., **159**, 269-277, https://doi.org/10.1016/j.rse.2014.12.014.

Volumetrisches Full-Waveform Stacking zur Ableitung der Gewässerbodentopographie aus Laserbathymetriedaten – Erste Ergebnisse einer Pilotstudie am Fluss Elbe

DAVID MADER[1], KATJA RICHTER[1], PATRICK WESTFELD[2] & HANS-GERD MAAS[1]

Zusammenfassung: Die Laserbathymetrie ermöglicht eine effiziente und flächendeckende Erfassung der Wasserbodentopographie in Flachwasserbereichen. Die Eindringtiefe dieser Methode in die Wassersäule ist jedoch stark durch die Trübung limitiert, resultierend in einer eingeschränkten Erfassung der Gewässerbodentopographie bei tieferen Gewässern. Dieser Beitrag stellt ein erweitertes Full-Waveform Verarbeitungsverfahren vor. Ziel ist, die auswertbare Eindringtiefe zu erhöhen und eine zuverlässige Extraktion und Erkennung von Bodenpunkten in tieferen Gewässern zu ermöglichen. Die Methodik basiert auf der Analyse von Informationen dicht benachbarter Messungen unter der Annahme, dass die Wassertiefe lokal stetig ist. Die Ergebnisse zeigen eine deutlich höhere auswertbare Eindringtiefe mit einer hohen Zuverlässigkeit der zusätzlich extrahierten Wasserbodenpunkte bei gleichzeitig größerer Abdeckung der Gewässerbodentopographie.

1 Einleitung

Airborne LiDAR Bathymetry (ALB) ist ein aktuelles Verfahren zur effizienten und großflächigen Vermessung von Gewässerbodentopographien in flachen Küstengebieten sowie in Binnengewässern. Das Verfahren basiert auf der simultanen Abtastung von Wasseroberfläche, Wassersäule und Gewässerboden durch bathymetrische Flugzeuglaserscannersysteme, die im grünen Wellenlängenbereich (532 nm) operieren. Die Gewässertrübung bewirkt dabei eine Abschwächung des Signals, welche die Eindringtiefe des grünen Laserstrahls in die Wassersäule limitiert. In Abb. 1 wird das Problem beispielhaft an zwei Full-Waveforms von unterschiedlichen Gewässertiefen illustriert. Bei der Full-Waveform in Abb. 1a kann das Bodenecho in einer Gewässertiefe von 0,78 m eindeutig detektiert werden. Im Beispiel in Abb. 1b ist die Detektion des Bodenechos in einer Gewässertiefe von 1,68 m nicht mehr eindeutig möglich.

In MADER et al. (2021) wurde ein neuartiges Verfahren zur Prozessierung von Laserbathymetriedaten vorgestellt, mit dem die auswertbare Gewässertiefe im konkreten Beispiel um 27% erhöht werden konnte. Grundidee des Verfahrens ist die Kombination von eng benachbarten Full-Waveforms in einem nicht-linearen Full-Waveform Stacking Ansatz (Zusammenfassung von mehreren Full-Waveforms zu einer Gesamtwaveform (stacked Full-Waveform)). Dadurch wird es möglich, Gewässerbodenpunkte zu detektieren, die mit Standardauswertemethoden nicht erkannt werden konnten.

[1] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden, E-Mail: [david.mader, katja.richter1, hans-gerd.maas]@hcu-hamburg.de
[2] Bundesamt für Seeschifffahrt und Hydrographie, Neptunallee 5, D-18057 Rostock, E-Mail: patrick.westfeld@bsh.de

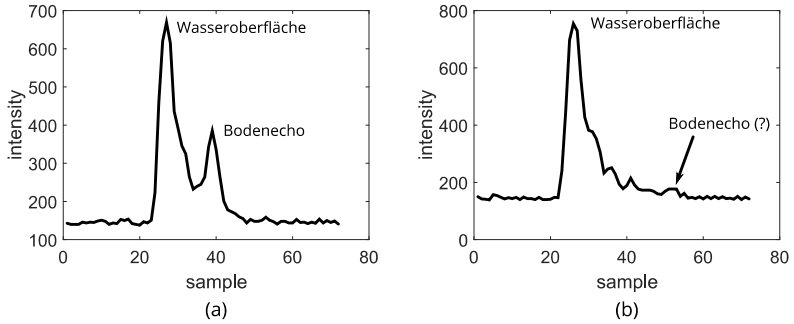


Abb. 1: Full-Waveform Laserbathymetriedaten mit (a) deutlich erkennbarem Gewässerbodenecho und (b) nicht eindeutig detektierbarem Bodenecho (MADER et al. 2021)

In dem Ansatz wurde zur Definition der Nachbarschaft ein regelmäßiges Raster auf der Wasseroberfläche definiert, in dessen Rasterzellen die individuellen Full-Waveforms zusammengefasst und gemeinsam ausgewertet wurden. Durch eine geeignete Wahl der Rasterzellengröße, kann die schräge Richtung des Laserstrahls durch die Wassersäule (z.B. forward-look und backward-look im Palmer-Scanmuster) bei geringen Gewässertiefen von ca. 2,0 m vernachlässigt werden. In tieferen Gewässern, wie z.B. küstennahen Flachwasserbereichen, kann die laterale Nachbarschaft an der Wasseroberfläche jedoch nicht mehr ohne weiteres auf den Gewässerboden übertragen werden. Die Prozessierung derartiger Datensätze erfordert somit eine Erweiterung des entwickelten Full-Waveform Stacking Ansatzes, welche eine geometrisch korrekte Berücksichtigung der Nachbarschaft am Gewässerboden ermöglicht. In diesem Beitrag wird daher ein volumetrischer Full-Waveform Stacking Ansatz zur zuverlässigen und genauen Extraktion des Gewässerbodens aus Laserbathymetriedaten vorgestellt.
Abschnitt 2 gibt zunächst einen Überblick über das Untersuchungsgebiet sowie die Datengrundlage. Anschließend werden in Abschnitt 3 die weiterentwickelten Methoden detailliert behandelt. Im Ergebnisteil (Abschnitt 4) erfolgt die visuelle und quantitative Evaluation und Validierung der erzielten Resultate auf Grundlage von Vergleichen mit standardprozessierten Daten und Echolotmessungen.

2 Untersuchungsgebiet und Datengrundlage

In diesem Beitrag wird ein erweiterter Ansatz zur Prozessierung von Full-Waveform Laserbathymetriedaten mit schwachen Gewässerbodenechos vorgestellt. Die Einschätzung des Potentials dieses Ansatzes erforderte einen ALB Datensatz inklusive Full-Waveform Informationen eines Gewässers und Referenzdaten der Gewässerbodentopographie, um eine aussagekräftige Evaluierung und Validierung der erzielten Ergebnisse gewährleisten zu können. In der hier vorgestellten Pilotstudie kamen ALB-Messdaten eines Abschnittes der Elbe nahe der Stadt Elster zur Anwendung (Abb. 2), die vom Wasserstraßen- und Schifffahrtsamt Elbe (WSA Elbe) und der Bundesanstalt für Gewässerkunde (BfG) bereitgestellt wurden. Diese Daten erfüllen die genannten Anforderungen weitgehend; lediglich die Verfügbarkeit von hochgenauen Referenzdaten der Gewässersohle ist nicht gegeben. Stattdessen kamen Echolotmessungen für die Validierung als Vergleichsdaten zum Einsatz (Genauigkeiten siehe Abschnitt 2.2).

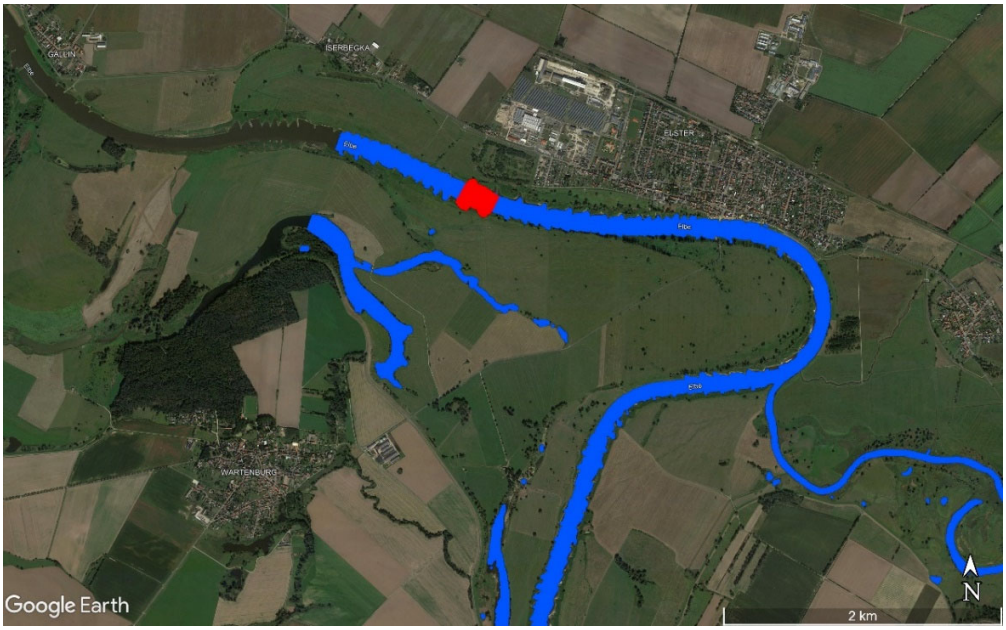


Abb. 2: Karte des Untersuchungsgebiets (blau) mit dem Testgebiet (rot). Nördlich der Elbe liegt die Stadt Elster

Die Ergebnisse der Standardauswertemethoden (online waveform processing - OWP) für die gemessenen ALB-Daten durch den Auftragnehmer bilden den Gewässerboden bis zu einer ungefähren Gewässertiefe von 1,65 m ab. Die Echolotdaten weisen Gewässertiefen über 2,20 m auf. Der Datensatz bietet somit genug Spielraum für die Entwicklung der erweiterten Auswertemethoden und deren Validierung sowie Evaluierung.
Im Folgenden werden das Untersuchungsgebiet, das ALB-Messsystem, die Charakteristik der Messdaten und die als Vergleichsdaten verwendeten Echolotdaten kurz vorgestellt.

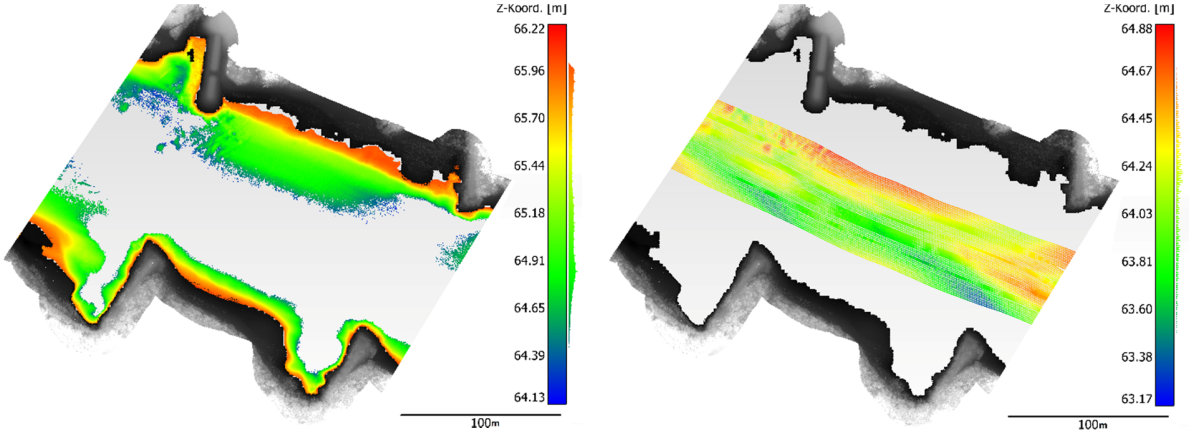
2.1 Untersuchungsgebiet und ALB-Messdaten

Die Datenaufnahme erfolgte im Frühjahr 2015 mit dem topo-bathymetrischen Messsystem RIEGL VQ-880-G in einer ungefähren Flughöhe von 380 m. Die Punktdichte variiert zwischen 11 bis zu 55 Punkten pro m² in Abhängigkeit von Datenlücken, Überlappung der Flugstreifen und heterogener Punktverteilung des Palmer-Scanmuster. Zusätzlich zu den 3D-Punkten wurden die Full-Waveform Daten gespeichert. Eine Full-Waveform besteht aus 60-200 gleichabständigen Abtastwerten (Samples) mit einem konstanten Abtastintervall von 0,575 ns, was unter Wasser einem Intervall von ca. 6,5 cm entspricht. Die Evaluierung der Auswertemethoden erfolgte in einem Testgebiet, welches in Abb. 2 rot dargestellt ist. Abb. 3a zeigt die Ergebnisse der Standardprozessierung im Testgebiet.

2.2 Echolotdaten

Die für die Validierung verwendeten Echolotdaten wurden von einem Messschiff im Herbst 2015 akquiriert. Das Schiff war zu diesem Zeitpunkt mit einem Leica GX1200 RTK SAPOS Rover (für die Positionierung), einem Seatex MRU5 (für die Orientierung) und einem elektroakustischen Peilrahmen ausgestattet. Der Peilrahmen bestand aus 37 Singlebeam-Schwingern, verteilt auf zwei

Seitenauslegern mit einer Spannweite von 12 m. Die Echolote der Firma Dr. Fahrentholz (Kiel) messen mit einer Frequenz von 200 kHz und eine Tiefenauflösung von 1 cm. In WEIß & WIRTH 2015 wird eine 3D-Punktgenauigkeit von 7,5 cm im Böschungsbereichen und 6,4 cm in flachen Flussbettbereichen angegeben. Die Echolotdaten sind in Abb. 3b farbkodiert bzgl. ihrer Höhenkoordinaten dargestellt.



(a) (b)

Abb. 3: Darstellung des Testgebietes mit (a) Ergebnissen der Full-Waveform Standardprozessierung und (b) Echolotmessungen. Die Punktwolken sind farbkodiert bzgl. ihrer Höhenkoordinate. Die grauwertkodierten Bereiche sind Land- und Uferbereiche. Die standardprozessierten Punkte beschränken sich hauptsächlich auf die flacheren Abschnitte des Flusses, während die Echolotpunkte aufgrund der Zugänglichkeit für Schiffe hauptsächlich die tieferen Teile des Flusses abdecken

3 Methoden

Das volumetrische Full-Waveform Stacking verfolgt das Ziel der Detektion von zum Teil sehr schwachen Gewässerbodenechos. Diese sind durch Standardauswerteverfahren nicht mehr detektierbar, da sie in der Regel ohne das Hinzuziehen von zusätzlicher Information nicht vom zufälligen Rauschen unterschieden werden können. Diese Informationen werden aus der gemeinsamen Auswertung benachbarter Full-Waveform Daten gewonnen. Es wird die Annahme zu Grunde gelegt, dass sich durch eine gemeinsame Auswertung und eines lokal stetigen Gewässerbodenverlaufs zufällige Signalanteile (zumeist Rauschsignal) gegenseitig abschwächen oder gar eliminieren. Schwache Objektechos werden hingegen verstärkt und somit einfacher detektierbar.

Für diese Auswertung muss zunächst festgelegt werden, welche Messungen aufgrund ihrer Nachbarschaft zueinander gemeinsam ausgewertet werden sollen. Dazu werden die Messdaten in einen Voxelraum integriert (Details siehe Abschnitt 3.1). Anschließend erfolgt die Auswertung und Analyse räumlich dicht benachbarter Full-Waveform Daten (Abschnitt 3.2), um aus den Ergebnissen einen Suchbereich für das Gewässerbodenecho in den einzelnen Full-Waveforms abzuleiten. Innerhalb dieses Suchbereichs wird das Maximum detektiert, welches mit hoher

Wahrscheinlichkeit dem Gewässerbodenecho entspricht, und der dazugehörige Gewässerbodenpunkt extrahiert (Abschnitt 3.3).

3.1 Integration der Full-Waveform Daten in den Voxelraum

Der volumetrische Full-Waveform Stacking Ansatz basiert auf der Überführung der Laserbathymetriedaten in eine lokale Voxelraumrepräsentation. Es werden also nicht mehr die an der Wasseroberfläche benachbarten individuellen Full-Waveforms durch Akkumulation zu einer sogenannten stacked Full-Waveform zusammengefasst, sondern die individuellen Full-Waveforms geometrisch korrekt in einen Voxelraum eingetragen.

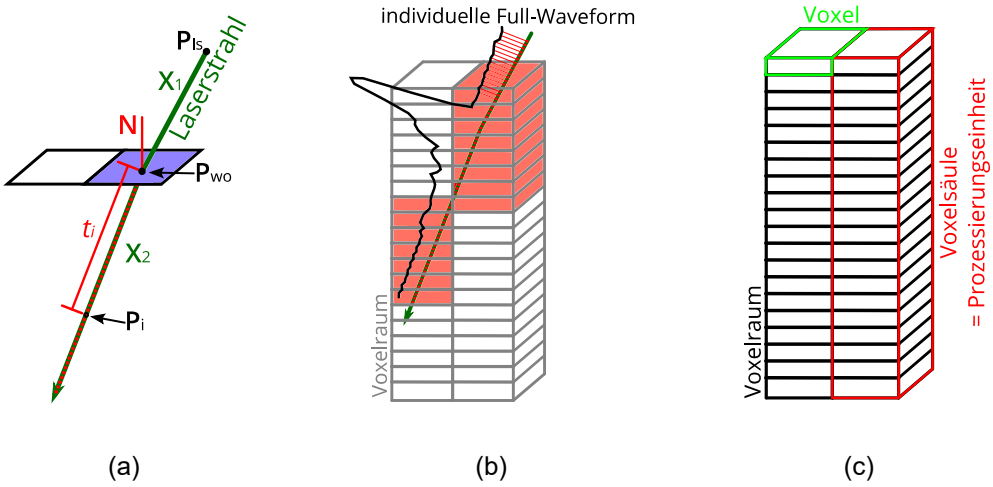


Abb. 4: Schematische Darstellung der Voxelraumrepräsentation. Links: geometrische Darstellung der Sampleintegration, Mitte: Prinzip der Integration von Full-Waveform Samples in den Voxelraum, Rechts: Voxelraum mit einzelnem Voxel (grün) und Voxelsäule bzw. Prozessierungseinheit (rot)

Im Detail bedeutet dies, dass die Intensitäten der georeferenzierten Samples jeder Full-Waveform in den Voxelraum transformiert werden. Dazu wird mit dem Richtungsvektor $\mathbf{X}_1$ (normierter Vektor von Punkt der Pulsemmision $\mathbf{P}_{ls}$ zu Wasseroberflächenpunkt $\mathbf{P}_{wo}$) und der Wasseroberflächennormalen $\mathbf{N}$ der Richtungsvektor $\mathbf{X}_2$ des Laserstrahls in der Wassersäule berechnet (Gl. 1 (GLASSNER 1989); Abb. 4a).

$$\mathbf{X}_2 = \eta \cdot \mathbf{X}_1 + \left(\eta \cdot C - \sqrt{1 + \eta^2 \cdot (C^2 - 1)} \right) \cdot \mathbf{N} \qquad (1)$$

$$C = -\mathbf{N} \cdot \mathbf{X_1}$$

$$\eta = n_{Luft} \; / \, n_{Wasser}$$

mit

n_{Luft}, n_{Wasser}: Brechungsindex Luft, Brechungsindex Wasser

Anhand der Sampleanzahl zwischen Wasseroberflächenpeak und des Samples i sowie des Sampling Intervalls kann für jedes Sample i eine Laufzeit t_i bestimmt werden, woraus der Punkt $\mathbf{P}_i$

berechnet und zusammen mit der dazugehörigem Intensitätswert einem Voxel zugeordnet werden (Gl. 2; Abb. 4b).

$$\mathbf{P}_i = \mathbf{P}_{WO} + t_i \cdot c_{Wasser} \cdot \mathbf{X}_2 \tag{2}$$

mit

c_{Wasser}: Lichtgeschwindigkeit in Wasser (ca. 225000 km/s)

Fallen mehrere Intensitätswerte in einem Voxel zusammen, wird aus den Intensitätswerten das arithmetische Mittel gebildet und dieser Mittelwert dem Voxel zugeordnet. Die Strahlendivergenz des Laserstrahls wird dabei vernachlässigt.
Anschließend erfolgt die Zusammenfassung der vertikal angeordneten Voxel zu einer vertikalen Voxelsäule (= Prozessierungseinheit), was im Prinzip einer Wassersäule entspricht (Abb. 4c). Alle Messdaten, die in eine solche Prozessierungseinheit fallen, werden als dicht benachbarte Messwerte angesehen und in der weiteren Prozessierung gemeinsam ausgewertet.

3.2 Generierung und Analyse der Ortho Full-Waveforms

Nachdem die Messdaten in eine lokale Voxelraumrepräsentation integriert und die Prozessierungseinheiten festgelegt wurden, können diese ausgewertet werden. Dazu wird für jede vertikale Voxelsäule eine sogenannte Ortho Full-Waveform generiert (PAN et al., 2016), indem jedes Voxel der Prozessierungseinheit als ein Sample der Ortho Full-Waveform interpretiert wird (Abb. 5). Die Intensität des Ortho Full-Waveform Samples ergibt sich aus der mittleren Intensität aller Samples im entsprechenden Voxel. Somit kann die Ortho Full-Waveform als repräsentatives Signal der Prozessierungseinheit angesehen werden.

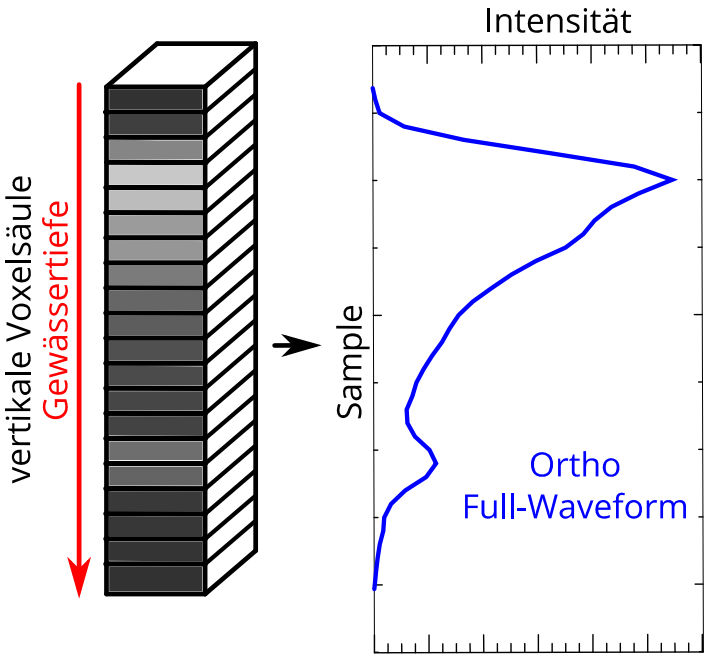


Abb. 5: Generierung einer Ortho Full-Waveform aus einer vertikalen Voxelsäule (RICHTER et al. 2021)

Die Ortho Full-Waveform wird anschließend mit dem Ziel analysiert, einen Suchkorridor zur Detektion der Gewässerbodenechos in den individuellen Full-Waveforms zuverlässig zu bestimmen. Dazu werden alle lokalen Maxima mit den Parametern *Isolation*, *Prominenz* und *Amplitude* charakterisiert (Abb. 6) und daraus die *Signifikanz* der Maxima berechnet.

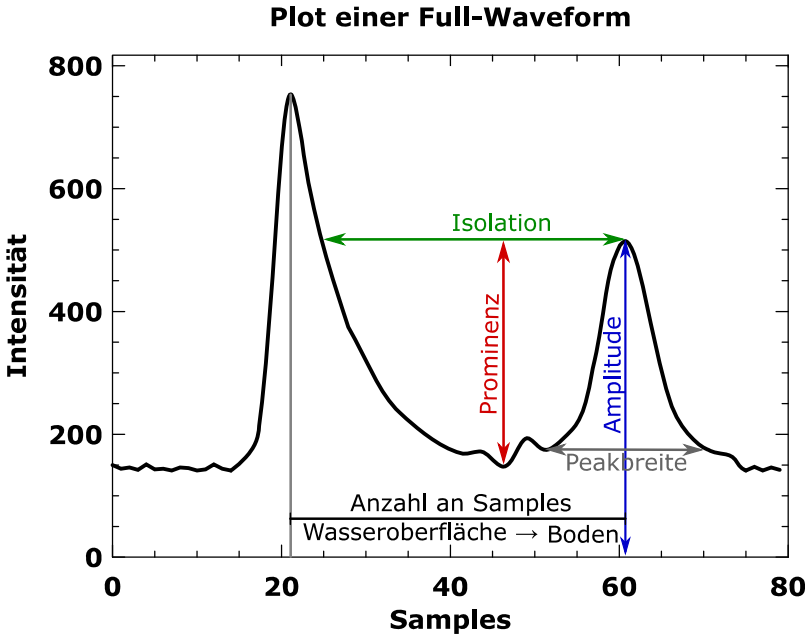


Abb. 6: Schematische Darstellung der Parameter der Full-Waveform Analyse (MADER et al. 2021)

Die Parameter *Isolation* und *Prominenz* sind aus der Arbeit von KIRMSE & DE FERRANTI 2017 abgeleitet, wo sie zur Detektion von signifikanten Berggipfeln aus Höhendaten verwendet wurden. Die *Isolation* gibt die Anzahl an Samples zwischen dem lokalen Maximum und einem Sample mit mindestens dem gleichen Intensitätswert an. Die *Prominenz* ist die Differenz der Intensitäten zwischen dem lokalen Maximum und dem lokalen Minimum (zwischen den beiden Samples der Isolation). Aus dem Produkt von *Isolation*, *Prominenz* und *Amplitude* resultiert die *Signifikanz*, wobei die Anteile bei Bedarf verschieden gewichtet werden können. In SCHWARZ et al. 2019 wurden die *Signifikanz* gemeinsam mit *Isolation*, *Prominenz* und *Amplitude* zur Näherungswertbestimmung von Wasseroberflächen- und Gewässerbodenpeak verwendet.
Anhand der *Signifikanz* der lokalen Maxima können die Echos der Wasseroberfläche und des Gewässerbodens in der Ortho Full-Waveform detektiert werden, unter der Annahme, dass der signifikanteste Peak die Wasseroberfläche und der signifikanteste Peak nach der Wasseroberfläche den Gewässerboden repräsentieren. Nachdem alle Prozessierungseinheiten unabhängig voneinander analysiert wurden, erfolgt eine Filterung, um mögliche Fehldetektionen zu erkennen und zu eliminieren. Die Filterung wird iterativ durchgeführt, indem die ermittelte Wassertiefe jeder Prozessierungseinheit mit dem arithmetischen Mittel der bereits validierten Nachbarschaftsinformationen verglichen wird. Die Einzelheiten zu diesem Filterverfahren sind in MADER et al. 2021 veröffentlicht.

3.3 Detektion und Extraktion von Gewässerbodenpunkte

Die aus der Analyse der Ortho Full-Waveforms resultierende Gewässertiefe dient als Näherungswert für die Detektion und Extraktion der Gewässerbodenechos in den gemessenen Full-Waveforms der Prozessierungseinheit. Für jede gemessene Full-Waveform wird ein Bereich bestimmt, in dem nach Gewässerbodenechos gesucht werden soll. Dazu wird aus dem Näherungswert der Gewässertiefe unter Berücksichtigung der Laserstrahlrichtung die entsprechende Samplenummer in der gemessenen Full-Waveform abgeleitet. Anhand der Charakteristik der Ortho Full-Waveform wird dann ein Toleranzbereich für die Suche festgelegt (MADER et al. 2021). Der Peak im Suchbereich, der dem Näherungswert am nächsten ist, wird als Gewässerbodenecho detektiert. Befindet sich innerhalb des Suchbereichs kein lokales Maximum, wird auch kein Gewässerbodenpunkt für diese Full-Waveform berechnet.
Mit den vorliegenden Informationen erfolgt die Extraktion des Gewässerbodenpunktes, sofern ein Echo detektiert wurde (Abb. 7). Die Koordinaten des Gewässerbodenpunktes $\mathbf{P}_{GB}$ werden mit Gleichung (3) berechnet.

$$\mathbf{P}_{GB} = \mathbf{P}_{WO} + t_2 \cdot c_{Wasser} \cdot \mathbf{X}_2 \tag{3}$$

Der Parameter t_2 ist die Zeit, die der Laserstrahl von der Wasseroberfläche bis zum Gewässerboden benötigt und kann mit der Anzahl der Samples zwischen Wasseroberflächenpeak und Gewässerbodenpeak und dem zeitlichen Abtastintervall berechnet werden.

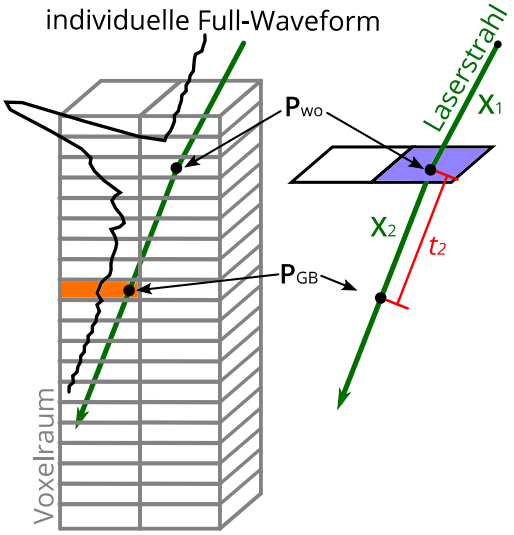


Abb. 7: Schematische Darstellung zur Extraktion eines Gewässerbodenpunktes aus den Ergebnissen des volumetrischen Full-Waveformstacking und der gemessenen Full-Waveform

4 Ergebnisse und Diskussion

Das volumetrische Full-Waveform Stacking (VFWS) wurde für alle Signale im Testgebiet durchgeführt. Für die Prozessierung kam eine Voxelgröße von 2,0 m × 2,0 m × 0,1 m zur Anwendung, welche sich in empirischen Untersuchungen bewährt hat. Aufgrund der Nicht-Linearität des Ansatzes bewirkt die Voxelgröße jedoch keine Reduktion der lateralen Auflösung am Boden.

Als erstes erfolgt eine visuelle Beurteilung der Ergebnisse. Abb. 8 zeigt die aus dem VFWS resultierende Punktwolke, farbkodiert bzgl. der Z-Koordinate. Grundsätzlich ist der in den Vergleichsdaten präsente Höhenverlauf (Abb. 3) hier ebenfalls sehr gut erkennbar, wobei es noch einige wenige falsch detektierte Bereiche (blauen Bereiche in Abb. 8) gibt. Dieser Eindruck wird durch die Profile in Abb. 9 bestätigt. Verglichen mit den OWP Daten (Abb. 3a) konnte eine deutlich höhere Eindringtiefe und Punktabdeckung und damit eine bessere Repräsentation des Gewässerbodens erreicht werden.

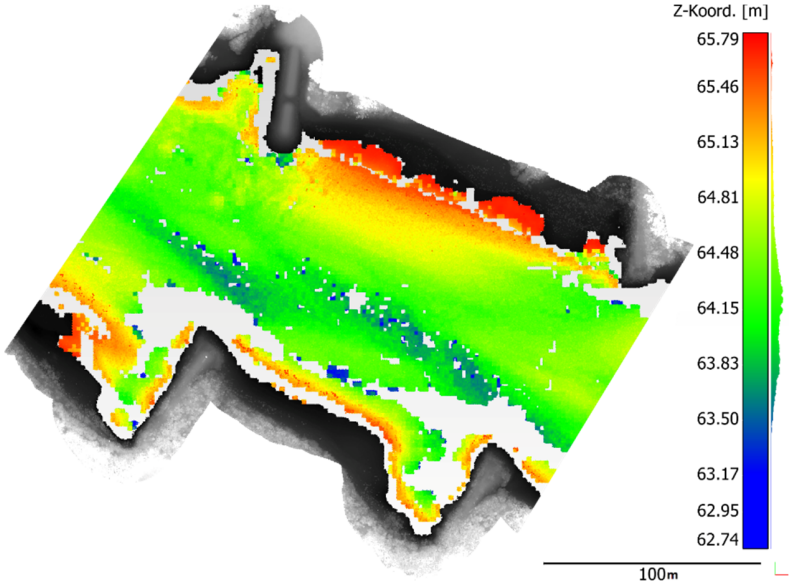


Abb. 8: Ergebnis der volumetrischen Full-Waveform Prozessierung. Die Z-Koordinaten bzw. Höhen sind farbcodiert dargestellt

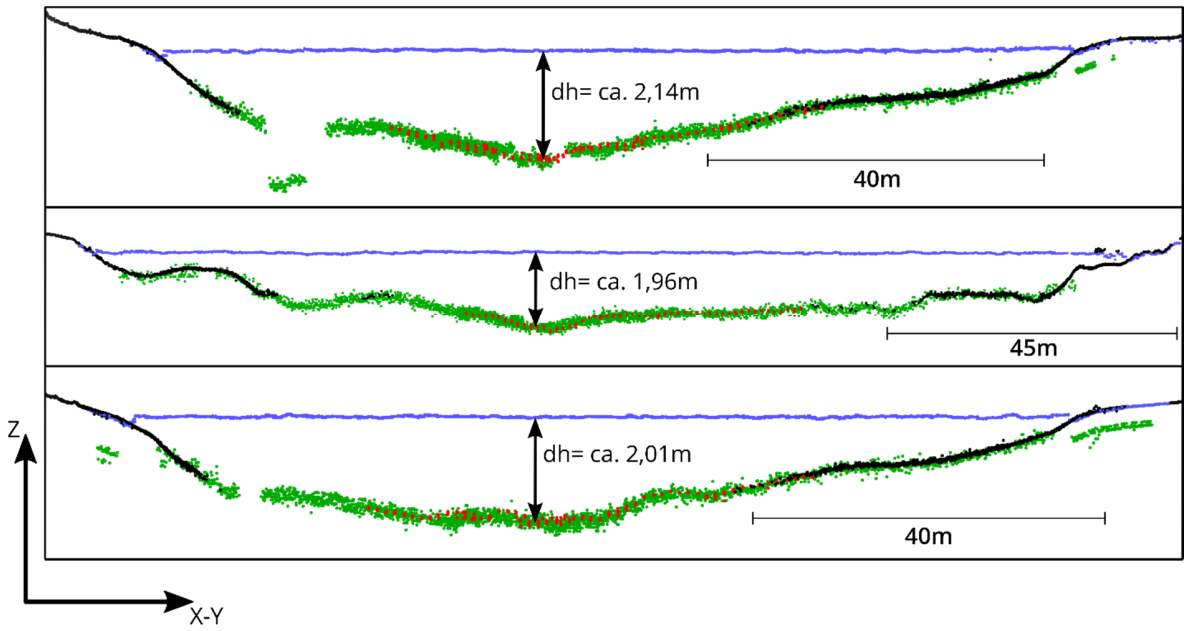
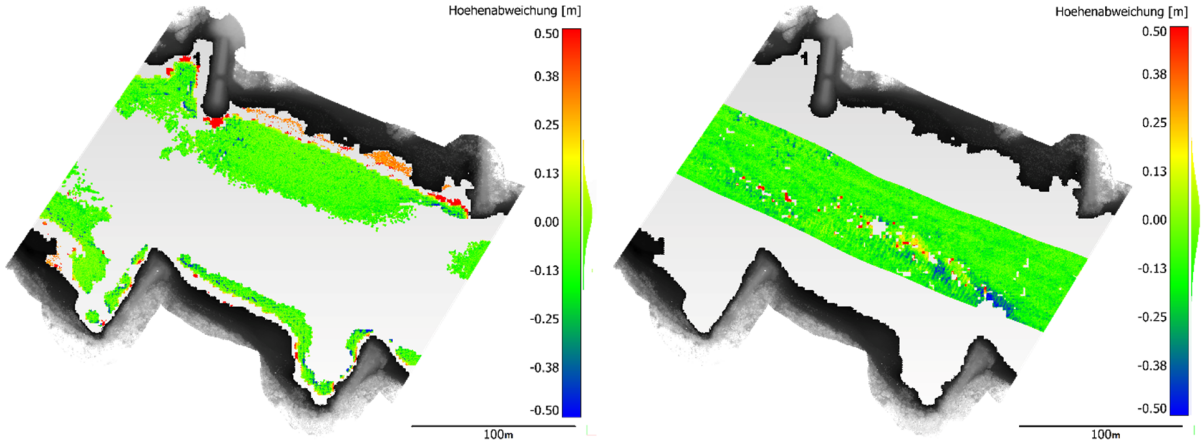


Abb. 9: Profile durch die Elbe. Die Z-Koordinatenkomponente ist 6fach überhöht. Blau: Wasseroberflächenpunkte; Schwarz: OWP Punkte; Rot: ES Punkte; Grün: VFWS Punkte

Zur Evaluation werden die Höhendifferenzen zwischen VFWS Ergebnissen und Vergleichsdaten quantitativ analysiert. Der Fokus liegt dabei auf dem Vergleich mit den Echolotmessungen (ES). Die standardprozessierten Daten (OWP) werden lediglich für eine erste Überprüfung der VFWS Daten herangezogen. Abb. 10 zeigt die Höhendifferenzen als farbcodierte Darstellung.



(a) (b)

Abb. 10: Farbkodierte Darstellung der Höhendifferenzen zwischen neuprozessierten Bodenpunkten und (a) OWP Punkten und (b) ES Punkten

In Tab. 1 sind Root Mean Square (RMS), Mean Absolute Deviation (MAD(mean) / $\sigma_{MAD(mean)}$) und Median Absolute Deviation (MAD(median) / $\sigma_{MAD(median)}$) zusammengefasst. Zur Beurteilung der Zuverlässigkeit wird außerdem die sogenannte Inlier Rate (LITMAN et al. 2015) angegeben, welche den Anteil an VFWS Punkte beschreibt, deren Höhendifferenz zu den Vergleichswerten unter einem bestimmten Grenzwert Δh liegt.

Tab. 1: Statistische Auswertung der Höhendifferenzen zwischen allen neuprozessierten Bodenpunkten und den Vergleichsdaten

Vergleich	**VFWS vs. OWP**	**VFWS vs. ES**
Vergleichspunkte / VFWS Punkte	292093 / 1426958	891358 / 1426958
Mittlere Höhenabweichung / $\sigma_{\text{Mittlere Höhenabw.}}$	0.037 m / 0.15 m	0.012 m / 0.11 m
RMS	0.15 m	0.11 m
MAD(mean) / $\sigma_{MAD(mean)}$	0.073 m / 0.092 m	0.080 m / 0.10 m
MAD(median) / $\sigma_{MAD(median)}$	0.044 m / 0.065 m	0.062 m / 0.092 m
Inlier Rate: $\Delta h \leq 0{,}35$ m	97,90%	99,39%
Inlier Rate: $\Delta h \leq 0{,}25$ m	96,42%	97,43%
Inlier Rate: $\Delta h \leq 0{,}15$ m	90,96%	87,39%

Die statistische Auswertung der Höhendifferenzen bezogen auf die OWP Daten (Tab. 1) und die Visualisierungen (Abb. 8-10) bestätigen die korrekte Funktionsweise des volumetrischen Full-Waveform Stackings. Allerdings sind die oben aufgeführten Werte nicht für jede Gewässertiefe repräsentativ, sodass im Folgenden die Auswertung in Abhängigkeit der Gewässertiefe durchgeführt wurde.

In Abb. 11 sind die Genauigkeitsmaße und Zuverlässigkeitsmaße für den Vergleich von OWP Punkten und VFWS Punkten in Abhängigkeit von der Gewässertiefe dargestellt. Bis zu einer Gewässertiefe von ca. 0,7 m sind die Genauigkeitsmaße relativ schlecht, zeigen danach aber bessere Werte (Abb. 11a). Ein ähnliches Verhalten ist bei den Zuverlässigkeitsmaßen zu beobachten (Abb. 11b). Die Probleme im Gewässertiefenbereich von 0,0 m – 0,7 m sind aus dem Full-Waveform Stacking (MADER et al. 2021) bekannt und wurden unter anderem auf mögliche Probleme mit Mischechos von Wasseroberfläche, Wassersäule und Gewässerboden zurückgeführt. Ab einer Gewässertiefe von etwa 0,7 m liegen die OWP und VFWS Punkte aber dicht beieinander, was die guten Genauigkeitswerte und die hohen Inlier Werte bestätigen. Der Fokus dieses Beitrags und des Full-Waveform Stacking Ansatzes im Allgemeinem liegt auf der Detektion und Extraktion von Gewässerbodenpunkten, welche nicht durch OWP Verfahren erfasst wurden. Die Probleme in Bereichen mit sehr geringer Gewässertiefe spielen für die Beurteilung des Potentials des hier vorgestellten Verfahrens eine untergeordnete Rolle, da dort eine Standardprozessierung problemlos möglich ist.

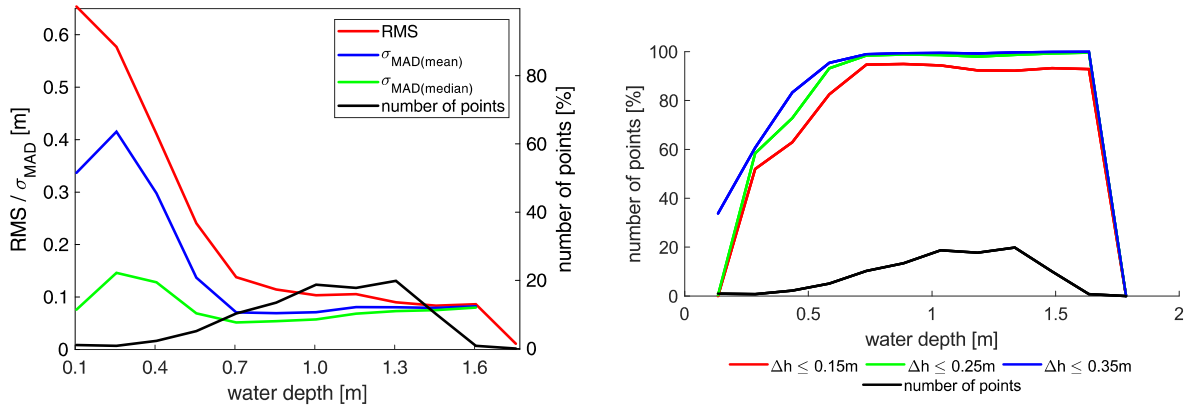


Abb. 11: Ergebnisse des Vergleichs OWP Punkte und VFWS Punkte in Abhängigkeit der Gewässertiefe. (a) Genauigkeitsmaße; (b) Zuverlässigkeitsmaße

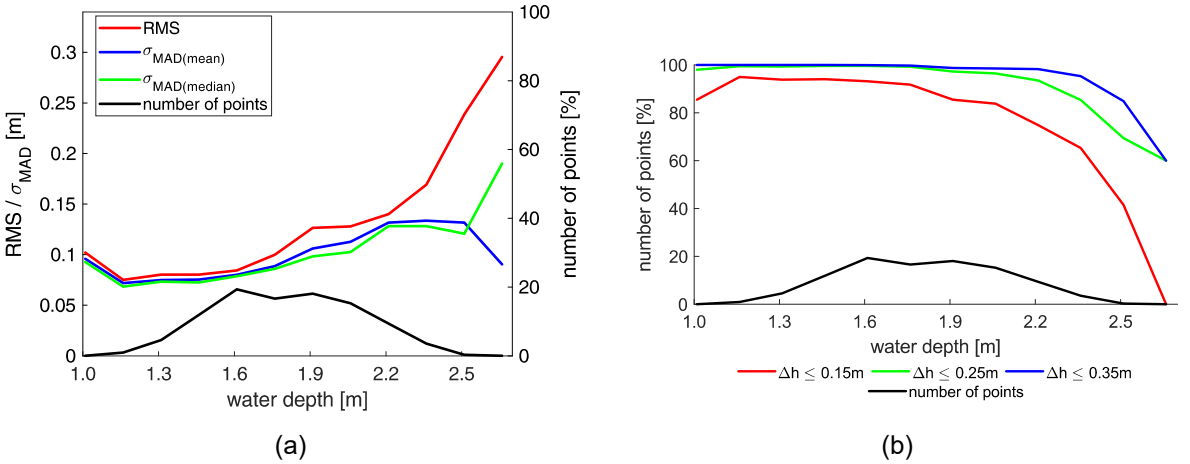


Abb. 12: Ergebnisse des Vergleichs ES Punkte und VFWS Punkte in Abhängigkeit der Gewässertiefe. (a) Genauigkeitswerte; (b) Zuverlässigkeitswerte

Die Ergebnisse des Vergleichs von ES Punkten und VFWS Punkten ist in Abb. 12 dargestellt. Es zeigt sich, dass die Güte der VFWS Punkte durch die ES Punkte validiert wird. Die Genauigkeitswerte in Abb. 12a bewegen sich bis zu einer Gewässertiefe von ca. 2,0 m zwischen 10-15 cm, was von der Inlier Rate in Abb. 12b bestätigt wird.

5 Fazit und Ausblick

Die Eindringtiefe von bathymetrischen Laserscannersystemen wird, aufgrund der starken Abschwächung des ausgesandten Lasersignals in der Wassersäule, stark von der Trübung des Gewässers beeinflusst. Dies führt zu einer Abschwächung des Gewässerbodenechos in den aufgenommen Full-Waveform Daten, sodass ab einer gewissen Gewässertiefe das Bodenecho nicht mehr zuverlässig vom Rauschsignal des Sensors unterscheidbar ist und somit auch nicht mehr detektiert werden kann. Ziel des vorgestellten nicht-linearen volumetrischen Full-Waveform Stacking ist eine gemeinsame Auswertung räumlich dicht benachbarter Full-Waveform Daten mit der Annahme, dass die Anzahl der Samples zwischen Wasserobflächenecho und Bodenecho sehr ähnlich ist. Damit sollen zufällige Signalanteile (wie z.B. das Rauschen) abgeschwächt und redundante Signale (Boden) verstärkt werden. Für die gemeinsame Auswertung von dicht benachbarten Messungen kamen lokale Voxelraumrepräsentationen und daraus resultierende Ortho Full-Waveforms zur Anwendung. Auf Basis der Analyse der Ortho Full-Waveforms konnten Suchbereiche in den gemessenen Full-Waveforms definiert und, insofern vorhanden, Gewässerbodenecho detektiert sowie Gewässerbodenpunkte extrahiert werden.
Die Auswertung hat gezeigt, dass im konkreten Beispiel die auswertbare Eindringtiefe von ca. 1,65 m auf mindestens 2,10 m erhöht werden konnte, was einem Zuwachs um 27% entspricht. Dies führt zu einer deutlich größeren Abdeckung des Gewässerbodens durch die extrahierten Punkte des volumetrischen Full-Waveform Stackings (2,29fache im Vergleich zu der Anzahl an OWP Punkten).
Die quantitative Auswertung ergab gute Genauigkeiten bzgl. der Höhenabweichung. Der RMS Wert für das gesamte Gebiet lag bei 0,11 m. Da der RMS Wert durch Ausreißer oder auch systematische Fehler beeinflusst sein kann, wurden zusätzlich die $\sigma_{MAD(mean)}$ und $\sigma_{MAD(median)}$ Werte für die Beurteilung der Ergebnisse herangezogen. Für $\sigma_{MAD(mean)}$ und $\sigma_{MAD(median)}$ ergaben sich Werte von 0,10 m respektive 0,092 m, was ebenfalls auf eine hohe Güte der erzielten Ergebnisse hindeutet. Die Zuverlässigkeit der Ergebnisse wurde mit der Inlier Rate bestimmt und besagt, dass 97,43% der neuprozessierten Punkte eine Höhendifferenz zu den Vergleichsdaten von weniger als 25 cm aufweisen. Für eine optimierte Repräsentation der Gewässerbodentopographie aus ALB Daten sollten OWP Punkte (vor allem für die ufernahen Bereiche) und VFWS Punkte gemeinsam genutzt werden.
In zukünftigen Arbeiten sollen folgende Aspekte des volumetrischen Full-Waveform Stackings weiterentwickelt werden:

- Verbesserung der Filtermethoden
- Nutzung eines komplexeren Wasseroberflächenmodells nach RICHTER et. al 2021 für eine verbesserte Bestimmung der Wasseroberflächennormalen

- optimierte Bestimmung des Suchbereichs in der gemessenen Full-Waveform durch Einführung eines Gewässerbodenmodells.

Außerdem soll die Methodik auf maritime Datensätze aus küstennahen Flachwasserbereichen übertragen werden. Aufgrund der komplexen Charakteristik dieser Laserbathymetriedaten ist dafür weitere Entwicklungsarbeit erforderlich.

6 Danksagung

Die Arbeiten zur Laserbathymetrie werden von der Deutschen Forschungsgemeinschaft (DFG) gefördert. Wir danken auch der Bundesanstalt für Gewässerkunde (BfG) und dem Wasserstraßen- und Schifffahrtsamt Elbe (WSA Elbe) für die Unterstützung unserer Forschungsarbeiten sowie der Firma Airborne Hydro Mapping GmbH (AHM) für die gute Zusammenarbeit bei der Aufbereitung der Befliegungsdaten.

7 Literaturverzeichnis

GLASSNER, A. S., 1989: An introduction to ray tracing. Academic Press, London.

KIRMSE, A. & DE FERRANTI, J., 2017: Calculating the prominence and isolation of every mountain in the world. Progress in Physical Geography: Earth and Environment, **41**(6), 788-802.

LITMAN, R., KORMAN, S., BRONSTEIN, A. & AVIDAN, S., 2015: Inverting RANSAC: Global Model Detection via Inlier Rate Estimation. Proceedings of the IEEE conference on computer vision and pattern recognition, 5243-5251.

MADER, D., RICHTER, K., WESTFELD, P. & MAAS, H.-G., 2021: Potential of a Non-linear Full-Waveform Stacking Technique in Airborne LiDAR Bathymetry. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **89**, 139-158.

PAN, Z., GLENNIE, C. L., FERNANDEZ-DIAZ, J. C., LEGLEITER, C. J., OVERSTREET, B., 2016: Fusion of LiDAR orthowaveforms and hyperspectral imagery for shallow river bathymetry and turbidity estimation. IEEE Transactions on Geoscience and Remote Sensing, **54**(7), 4165-4177.

RICHTER, K., MADER, D., WESTFELD, P. & MAAS, H.-G., 2021: Water Turbidity Estimation from LiDAR Bathymetry Data by Full-Waveform Analysis Comparison of Two Approaches. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2021, 681-688.

SCHWARZ, R., MANDLBURGER, G., PFENNIGBAUER, M. & PFEIFER N., 2019: Design and evaluation of a full-waveform surface and bottom-detection algorithm for LiDAR bathymetry of very shallow waters. ISPRS Journal of Photogrammetry and Remote Sensing, **150**, 1-10.

WEIß, R. & WIRTH, H., 2015: BfG-1861 Erprobung der Laserbathymetrie im Binnenbereich. Technischer Report, Bundesanstalt für Gewässerkunde, Koblenz.

Semantic UAV Image Segmentation of Mixed Cropping Fields

QUSAI MARASHDEH[1], LUKAS DREES[1] & RIBANA ROSCHER[1]

Abstract: Mixed cropping became an important research topic in the area of sustainable agriculture, aiming on novel insights on how different plants interact with each other in the same agricultural field. Convolutional neural networks show remarkable capabilities to solve tasks like semantic segmentation and have been successfully used for agricultural applications. However, performing such tasks for fields with mixed crops is particularly challenging, especially when only a small number of reference images are available for training and testing neural network models. In this paper, we present a study conducted in the Cluster of Excellence "PhenoRob - Robotics and Phenotyping for Sustainable Crop Production" in which 320 unmanned aerial vehicles (UAVs) images of mixed crop fields from the experimental site Campus Klein-Altendorf are analyzed. Specifically, we perform semantic segmentation using a convolutional neural network with U-Net architecture to distinguish between the mixed crops faba bean and spring wheat. We use two different approaches to create annotation masks which are used to learn our model for segmentation. The goal is to quantify the heterogeneity based on the segmentation results and to analyze the influence of different segmentation masks. We evaluate our results by means of a mean confusion matrix and a visualization of our results comprising the estimated biomass of the two plants and total yield. Preliminary results show an overall accuracy of classified faba beans of 68% and an overall accuracy of classified spring wheat plants of 87%.

1 Introduction

Crop production is an active research area due to the demand of an increased sustainability while ensuring food for a growing population (YADAV et al. 2021). Objectives are focusing on different ways to increase food production considering various aspects such as the limited area, high input cost, saving the soil quality, the effect of pests infestation, and other threats that affect efficient crop production, for instance, climate change. Several studies investigate the statistical analysis of data from intercropping interaction (PEARCE & GRILLIVER 1978). One way is to take advantage of machine learning to analyze observational data to gain insights into the plants' behavior and their interactions. Recently, one successful approach is the use of convolutional neural networks, which have already shown success in related applications such as weed detection or disease detection (KAMILARIS et al. 2018; LOTTES et al. 2018).

In this paper, we present a study conducted within the Cluster of Excellence "PhenoRob - Robotics and Phenotyping for Sustainable Crop Production". In this study, 320 orthophotos were calculated from UAV images acquired over a four-month period from April to July at the experimental field Campus Klein-Altendorf in 2020. The goal of this study is to monitor the mixed plants and derive relevant phenotypic traits such as the yield efficiently at different time points, as shown in Fig. 1. We achieve our goal by performing semantic segmentation using a convolutional neural network

[1] Rheinische Friedrich-Wilhelms-Universität Bonn, Institut für Geodäsie und Geoinformation, Nussallee 15, D-53115 Bonn, E-Mail: qusai.marashdeh94@gmail.com, [ldrees, ribana.roscher]@uni-bonn.de

with U-Net architecture to distinguish between the mixed crops faba bean and spring wheat. We use the segmentation results to estimate the heterogeneity from which insights about different mixing ratios can be derived. We face challenges such as a limited spatial resolution, differences in the texture and spectral signature of mixed plants, and the overlap between the mixed components occurring at later growth stages.

Fig. 1: UAV image patches showing the same area of a mixed cropping field with spring wheat and faba bean at different time points: early growth stage (left), intermediate growth stage with which we train our model (middle), and late growth stage (right).

In our experiments, we show that neural networks are able to distinguish between different plants. We evaluate our results in different ways. First, by estimating the confusion matrix for comparing two annotation methods of mixed crop images. Second, we visualize our obtained results with a specific focus on the heterogeneity with respect to the yield.

2 Data

Data was acquired within the currently running cluster of excellence PhenoRob, specifically in the core project ‘New Field Arrangements’, which aims to estimate and evaluate how different crops evolve in mixed cropping fields.

2.1 Study location and sensor measurements

Our study site is located at Campus Klein-Altendorf close to Rheinbach in North Rhine-Westphalia, Germany. The experimental field consists of 320 plots, each of 15 m^2, arranged in 10 rows and 32 columns. On these plots, 8 different wheat varieties are combined with 2 different arable varieties at varying seed density, resulting in 197 mixture plots, 92 spring wheat and 31 faba bean monoculture reference plots. RGB images were acquired with a UAV at approximately 10 meter flight altitude on a weekly basis during the growth period from April to July 2020. From the images, RGB orthomosaics were computed with 3 mm GSD.

2.2 Patch extraction and data annotation

Each of the 320 plots, with the size of 2700 × 480 px is split into 6 quadratic patches of size 450 × 450 px. We annotated 9 mixed crop plots for training and testing our proposed method, where 5 patches from each mixed crop plot are used for training (in total: 45) and one for testing (in total:

9). For all monoculture plots, we performed the annotation semi-automatically in two different ways:
1- Application of vegetation indices (VI): We estimate the RGB vegetation indices (RGBVI) value for each pixel in the image and set an appropriate threshold to differentiate between the vegetation pixels and the background pixels.
2- Application of a pre-trained Gaussian classifier: We use a pre-trained model, which was trained using a set of plant images. The images were collected using two cameras, a Canon 7D and a high-end mobile phone (Sony XPeria Z3 Compact). Twenty images were collected by each of the cameras (40 images overall), capturing images of a fallow field farm in Gatton, Queensland, and garden beds in Brisbane, Queensland as described in (BAWDEN et al., 2017). The images are divided into three sets, 14 images for training, 14 images for validation, 12 for testing. We apply it to our data to differentiate between vegetation and background.
For images with mixed crops, we annotate in 2 steps: First, we annotate faba beans manually and second, we use the described procedure for monoculture images to annotate spring wheat plants. We combine both annotations by overlaying VI respectively Gaussian derived masks with the manual labeled masks as illustrated in Fig. 2.

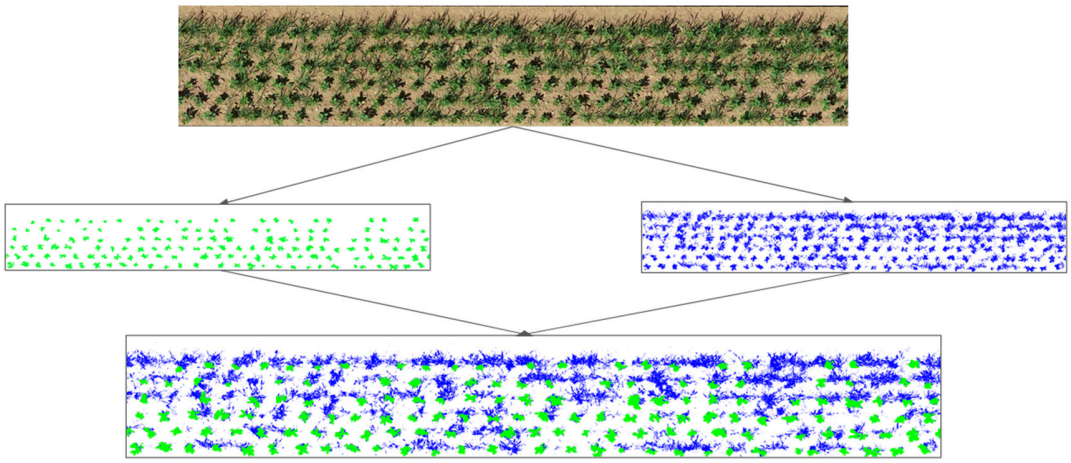


Fig. 2: Steps of producing the label mask. Top: Original image. Middle right: Resulting mask using a pre-trained Gaussian classifier for spring wheat. Middle left: Manually annotated mask for faba bean. Bottom: Combined mask for faba bean and spring wheat.

3 Methodology

We use U-Net architecture (RONNEBERGER et al. 2015) which already showed success for similar applications, and the number of parameters is comparably low.

3.1 U-Net: Convolutional network architecture

The architecture of the U-Net is illustrated in Fig. 3. Our implementation was performed with PyTorch in Python, which is an open-source library based on the Torch library.
The architecture contains two paths: encoding path and decoding path. In the encoder, a lower-dimensional representation of the data is computed to capture the most relevant patterns, while in

the decoder, the representation is mapped back to an intended outcome such as a segmentation mask. The intermediate results, which are obtained after each operation in the network are called feature maps.

1- Encoding path: This part consists of four repeated blocks, each containing two 3×3 convolutions with a rectified linear unit (ReLU) activation applied after each convolutional operation and a 2×2 max pooling operation for downsampling purposes. In each block, we double the number of convolutional filters in comparison to the previous block.

2- Decoding path: The decoder has a symmetric structure to the encoder, however, instead of downsampling operations we use upsampling operations. The final outcome has the same number of rows and columns as the input image. Each block contains an upsampling of the feature map with a 2×2 convolution, where the number of convolutional filters is reduced by 2 in comparison to the previous block. The architecture contains skip-connections, which means that the output of a previous block in the decoder is concatenated with the corresponding feature map in the encoder before applying the operations in each block, which are two 3×3 convolutions with ReLU after each of them. In the final layer, we use a 1×1 convolution layer to project the input channels to the desired output channels according to the class number, and a softmax activation function.

3.2 Evaluation metrics

We evaluate the segmentation results in two different ways. First, by means of a confusion matrix containing the three classes spring wheat, faba bean, and background. The matrix estimates the deviation between the reference mask and the predicted one. Moreover, we present the average confusion matrix. Further, we show the correlation of our estimation of the heterogeneity of the mixed plants by counting the number of pixels of each class and the total yield for each class.

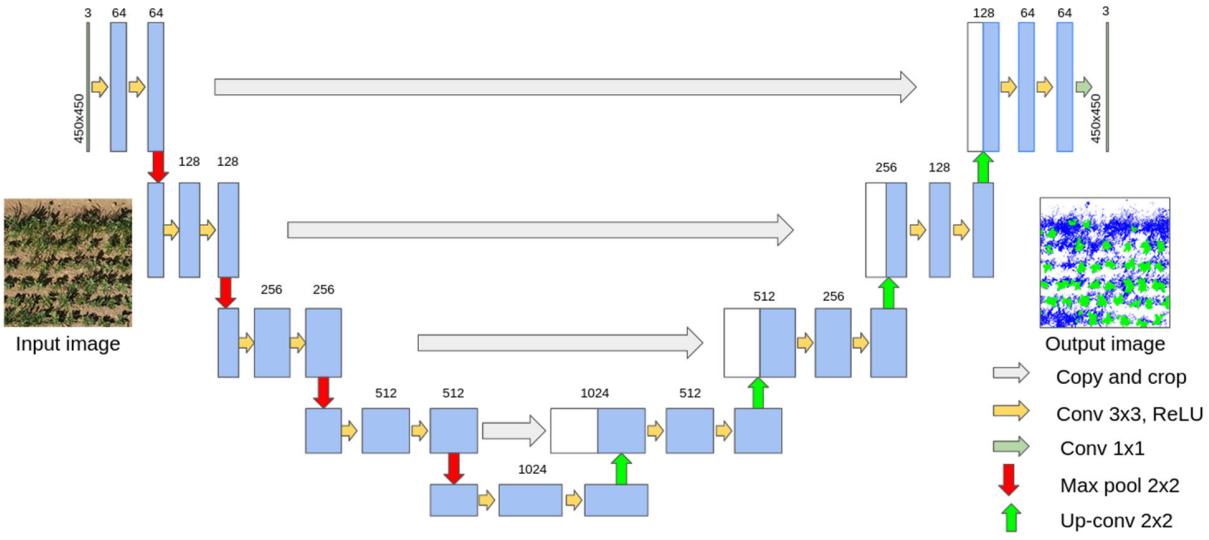


Fig. 3: U-Net architecture with skip connections (gray arrows) used for semantic segmentation. The architecture consists of an encoder (left part of the architecture) and a decoder (right part of the architecture). All arrows represent different mathematical operations.

4 Experiments

4.1 Experimental setup

We perform various experiments to evaluate (1) the influence of different annotation methods, (2) the capability to apply our model to monoculture images of the mixture components, and (3) how close we get with the mixture ratio derived from the segmentation images to in-situ reference measurements performed in the field.

For training the U-Net, we use cross-entropy loss with an Adaptive Moment Estimation (Adam) optimizer and a batch size of 1. Using one-side padding allows us to get an output image size equal to the input size. As data augmentation methods, random horizontal and vertical flipping as well as cropping are used, which help to avoid overfitting. Using all 45 mixed image patches of the training set, the training starts to converge after 200 epochs and thus takes about 2.5 hours. Since there are many monoculture images available that are easy to annotate, in initial experiments, we used an extended training dataset that includes both annotated monocultures and mixed cropping images. However, it has been found that the convergence time takes longer without improving the model, so the monocultures are only used for evaluation.

4.2 Results and Discussions

In Fig. 4, we first compare the confusion matrices of the two annotation methods using VI and a pre-trained Gaussian classifier, where the matrices are split into the classes background, faba bean (FB), and spring wheat (SW). The highest score in both cases is for the class background, which is reliably segmented in more than 95%. This is followed by SW, which is captured slightly better by the model trained using VI (92% compared to 87%). In return, the Gaussian trained model is 8 percentage points better for FB, although at a lower level than SW (68% to 58%). For the Gaussian results, we notice that most of the incorrectly predicted SW pixels are labeled as FB (11%), and vice versa most of the incorrectly predicted FB pixels are labeled as SW (20%). For VI results, the incorrectly predicted pixels of the SW class are equally often confused with FB and background. Noticeably, the incorrectly predicted FB pixels are not equally distributed among the other classes, but mainly belong to SW (34%).

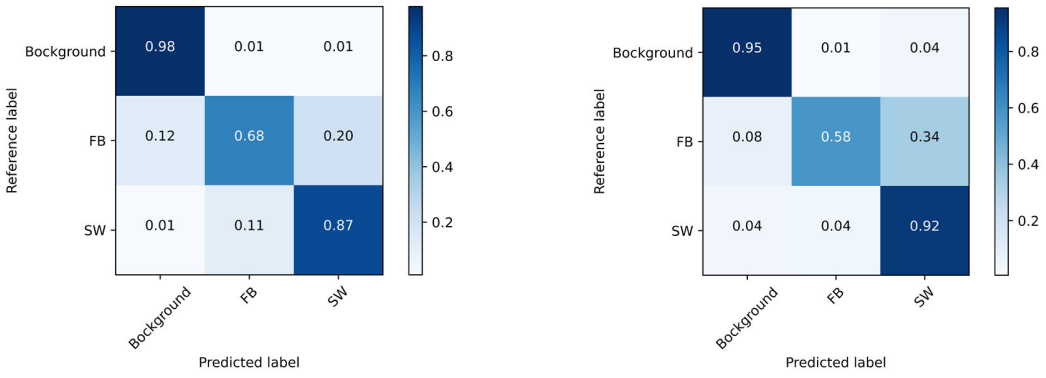


Fig. 4: Confusion matrix of semantic segmentation using the result of Gaussian annotation results (left) and VI annotation results (right) of training only 45 mixed patches to train our U-Net model (%).

Fig. 5 visualizes the differences of the annotation method on an exemplary mixture test patch. It is noticeable that the amount of annotated SW pixels varies significantly. While the result with the pre-trained Gaussian classifier contains less SW overall, the VI result contains more. Accordingly, this trend can also be seen in the segmentation. Although VI masks are controllable via selected threshold value, it is difficult to avoid annotating shadow areas: In the case of a too low threshold value, shadow areas are avoided, but a large number of holes on the plants are created. As a consequence, we have chosen a rather large threshold value in order to capture the plants completely, resulting in some false vegetation pixels in soil and shady areas than with the Gaussian classifier.

We assume that the high threshold value contributes to SW having a higher accuracy with VI annotation. In contrast, one can see in the images that FB annotations are less suppressed in the Gaussian result, resulting in thicker segmentation for this approach. Remarkably, with both annotation methods, our model succeeds in detecting FB plants that were missed in the manual annotation or are heavily overgrown with wheat (top field-row in the images).

To assess the stability of our model, we also applied the model to the monoculture images without training on them before. Fig. 6 shows an example of the results of the VI trained model on both types of monocultures, SW and FB. The results look promising: There are only minor outliers in the FB monoculture, while there are no FB outliers in the SW. It is worth noting that the model successfully segments contiguous FB plants, although the annotation is based on free-standing FB plants in the mixtures only.

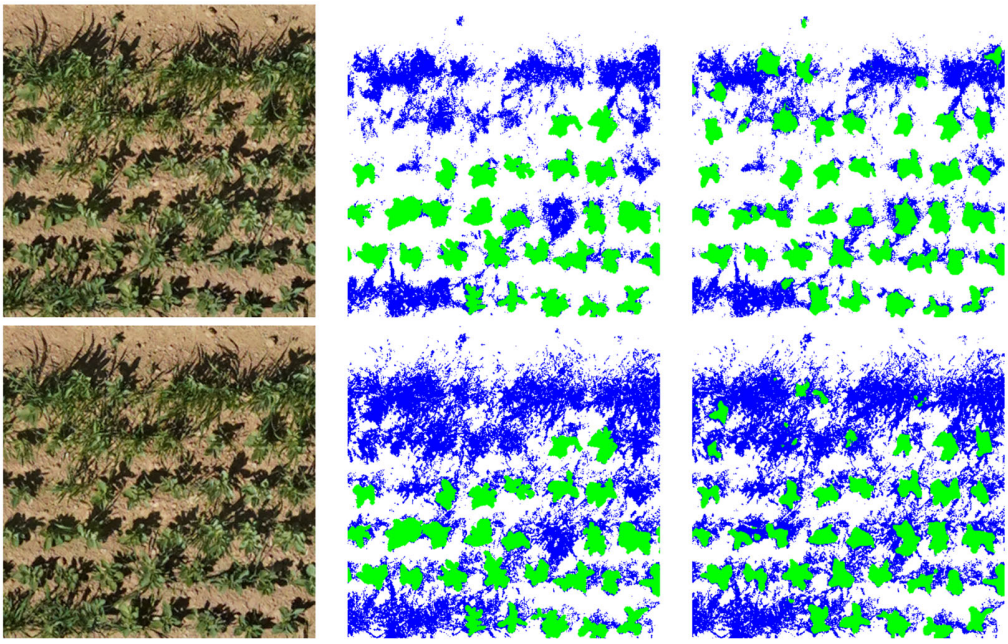


Fig. 5: Segmentation results of a mixture test patch. The first row shows the segmentation result using the annotation obtained from the pre-trained Gaussian classifier. The second row shows the segmentation result using VI for training. From left to right: Image patches, reference annotation masks, and predicted segmentation masks.

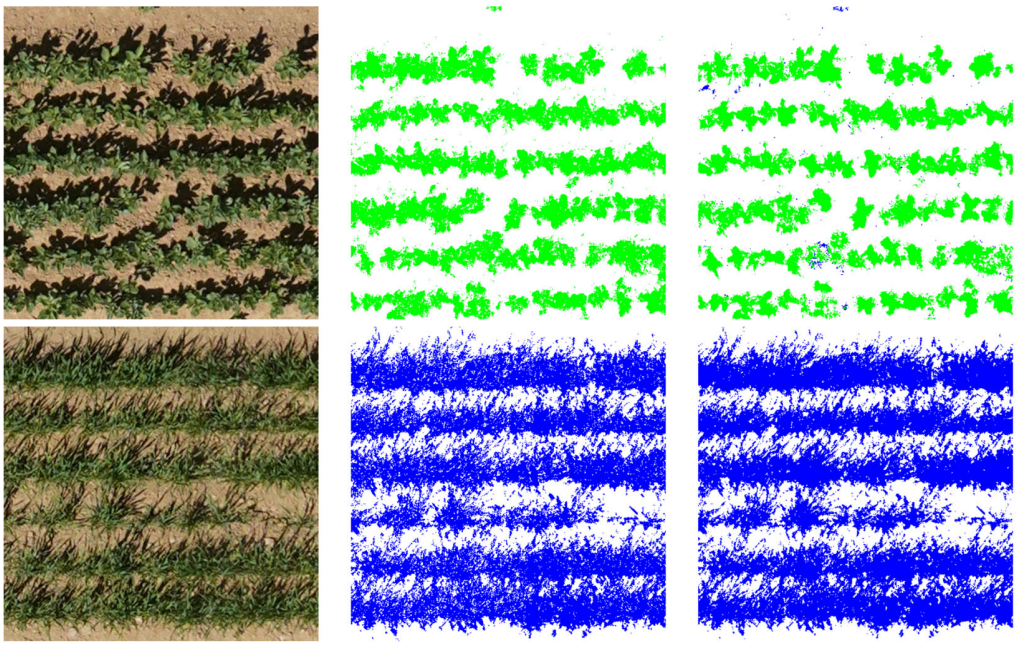


Fig. 6: Segmentation results of faba bean (top) and spring wheat (bottom) monocultures using VI-based U-Net training. From left to right: Image patches, reference annotation masks, and predicted segmentation masks.

To highlight the practical utility of the segmentations, we continue to use the VI model and intend to compare the results with in-situ measurements from the field. Approximately four weeks after the date the images were taken, the yield of the mixture components was manually measured in the field. With this, we assume a correlation between yield in the mixture and the mixing ratio, which we derive from the pixel ratio of the classes SW and FB in the segmentation image.
Fig. 7 shows the R^2 scatter plots in which the in-situ yield in gram (y-axis) of all mixture test images is compared to the respective FB and SW ratios derived from the images (x-axis). Thus each dot represents an image, and the R^2 value (maximum 1) indicates how well the regression line is determined by the dots. For both FB and SW, the R^2 values are rather low, but there is a clear trend indicating that a higher ratio also results in a correspondingly higher yield. Besides, the R^2 plot can only be an approximation, on which two factors have a major influence. First, the temporal difference between the in-situ data and our model estimates, because the mixing ratio may still have developed differently after the image-taking date. Second, the perspective, as the 2D image only considers the canopy surface, whereas 3D dimensional measurements can also count yields of beans and wheat grains that are located between the soil and the upper canopy. Nevertheless, the result indicates that our model is well suited to perform semantic segmentation in complex mixed cropping systems and can provide reasonable approximations for important phenotypic traits, such as the mixture ratio and yield.

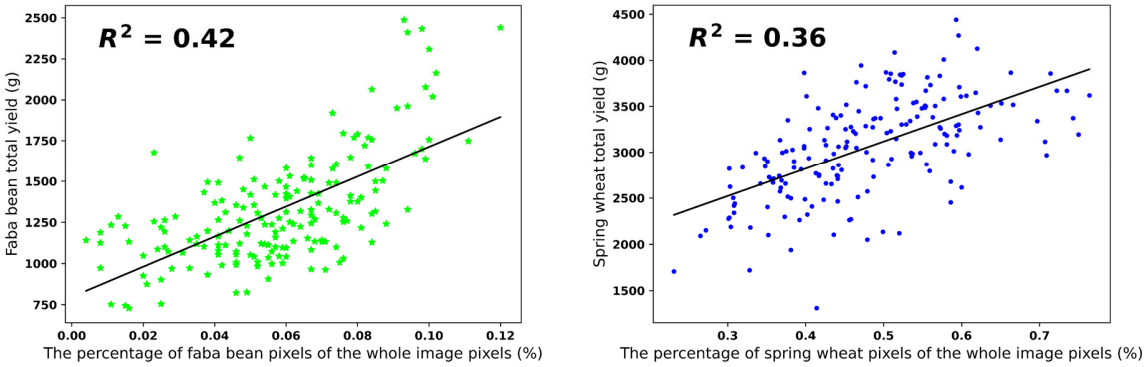


Fig. 7: The correlation between in-situ measured yield in gram (y-axis) and from segmentation images derived mixing ratio in percentage (x-axis) for faba bean on the left and spring wheat on the right. Segmentation images were obtained using VI-based U-Net training.

5 Conclusion

In this paper, we demonstrate a novel application of semantic RGB image segmentation using deep learning to highly complex mixed cropping environments. For this purpose, we use a neural network with the U-Net architecture, which is widely used in the field of semantic segmentation. Since it is time-consuming and difficult to perform manual annotations on such finely-structured and overlapping spring wheat and faba bean plants of a mixed cropping image, we compare two approaches for annotation generation, namely vegetation index-based annotation and annotation by Gaussian classifier. This allows us to manually annotate only the beans from the composite of mixed cropping partners, while wheat is labeled automatically.

Our experiments show that at a medium growth stage we are able to segment wheat in mixtures with up to 95% and beans with up to 70%. It turns out that the choice of both annotation methods has benefits and drawbacks. The VI-based model over-interprets the vegetation because a rather high threshold is needed to close holes in the canopy, but thus shaded areas are also annotated. Hence, spring wheat is well segmented, while beans are often incorrectly segmented as wheat. Meanwhile, the Gaussian classifier provides a stronger separation between vegetation and background, resulting in a better segmentation result for beans, but less correctly segmented wheat. We further show that our model is able to segment monoculture images of mixing components with satisfactory results. Comparison of the mixing ratio derived from the image with in-situ measurements taken later in the field shows trends that automatized image segmentation using deep neural networks is useful in areas of mixed cropping to approximate agricultural yield.

6 Acknowledgment

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2070 – 390732324.

7 References

BAWDEN, O., KULK, J., RUSSELL, R., MCCOOL, C., ENGLISH, A., DAYOUB, F., LEHNERT, F. & PEREZ, T., 2017: Robot for weed species plant-specific management. Journal of Field Robotics, **34**(6), 1179-1199, https://doi.org/10.1002/rob.21727.

KAMILARIS, A. & PRENAFETA-BOLDÚ, F.X., 2018: Deep learning in agriculture: A survey. Computers and electronics in agriculture, **147**, 70-90.

LOTTES, P., BEHLEY, J., MILIOTO, A. & STACHNISS, C., 2018: Fully convolutional networks with sequential information for robust crop and weed detection in precision farming. IEEE Robotics and Automation Letters, **3**(4), 2870-2877, https://10.1109/LRA.2018.2846289.

PEARCE, S.C. & GILLIVER, B., 1978: The statistical analysis of data from intercropping experiments. The Journal of Agricultural Science, **91**(3), 625-632.

RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. CoRR, abs/1505.04597, http://arxiv.org/abs/1505.04597.

YADAV, P., JAISWAL, D.K. & SINHA, R.K., 2021: Climate change: Impact on agricultural production and sustainable mitigation. S. Singh, P. Singh, S. Rangabhashiyam, K. Srivastava (eds), Global Climate Change, Elsevier, 151-174.

Direct Co-registration of As-built and As-designed Data in Digital Fabrication

KARAM MAWAS[1], MEHDI MABOUDI[1] & MARKUS GERKE[1]

Abstract: Additive Manufacturing in Construction (AMC) provides the possibility to print objects in unprecedented and novel ways, pushing the boundaries of what was previously possible in construction. Nonetheless, this requires an adequate and accurate co-registration of the digital representation of the printed object (as-built) with its digital model (as-designed) in order to achieve a robust quality control (QC). Most common automatic approaches are based on ICP-like distance minimization algorithm and target-based direct transformation. ICP suffers from the local minima problem and might take relatively large time to converge. Accordingly, we describe an experimental test specifically designed to underscore problems posed by objects with ambiguous surface structure which renders ICP unsuccessful. In order to overcome these shortcomings, we established a reference surveying network and transformed it to the robot (model) coordinate system. In this way, all future measurements could be transformed to the same coordinate as our as-designed model which is a pre-requisite for further inspections. The overall registration quality resulted in an RMSE of 3 mm.

1 Introduction

Digital fabrication is a design and manufacturing workflow where digital data directly drives manufacturing equipment to print parts and components of varying geometrical complexity on demand (BUCHLI et al. 2018). Compared to other current industries, however, construction is still the least digitized (AGARWAL et al. 2016; PLACZEK et al. 2021). Additive Manufacturing in Construction (AMC), however, streamlines traditional construction processes while allowing even more complex objects to be created more quickly. The use of digital construction workflow planning that relies on bi-directional workflow data, based on the so-called digital twin (EL-SAYEGH et al. 2020), is also far more efficient than analog planning processes. Construction components manufactured through AMC have nearly no limitations in terms of their complexity or shape, and require no special tools for their manufacture except for the printer itself. In doing so it decouples the cost of specialized or complex building components from their previous costs-originally such components would have to be specially ordered or manufactured offsite. Finally, by employing 3D printing, which has consistently proven to be more efficient in its direct use of materials and resources, digital fabrication, reduces waste (EL-SAYEGH et al. 2020).
Digital fabrication cannot be understood separately from the digital workflow process that drives it, which includes how a constructed space is digitally visualized. In order to better understand how digital fabrication in construction works, it is necessary to understand the concept of a digital twin. A Digital Model is a representation of the physical space and does not have any interaction

[1] Technische Universität Braunschweig, Institute of Geodesy and Photogrammetry, Bienroder Weg 81, D-38106 Braunschweig, E-mail: [k.mawas, m.maboudi, m.gerke]@tu-braunschweig.de

with the digital space. A Digital Twin employs a bi-directional data flow between physical objects and digital spaces that allows the physical objects to be modified (EL JAZZAR et al. 2020).
AMC necessarily relies on smooth bi-directional data flow between physical construction spaces and components on site and the virtual environment that models and drives their production. This bi-directional data flow is the heart of digital twin, and is required throughout the entire length of the construction process. In order to accurately print construction components, continuous scanning and quality control is needed in at every step of the multi-step manufacturing process. Before the next step of a sequential workflow can proceed, active and layered scanning data must be instantly available to ensure quality control. The object being manufactured must be scanned after pre-defined production steps and compared with its digital twin to ensure that it conforms accurately with predefined tolerances of the production standards (PLACZEK et al. 2021).
Quality control (QC) can be broken down into two categories: QC without a previously alignment with the as-designed model, and QC based alignment with the as-designed model. The first category of quality control is concerned with inspection of the geometric features of a physical object, such as its edges or relative dimensions. The second, however, is concerned with finding deviations between the printed model and the designed model. This step requires that both models be co-registered accurately, as that the digital model holds the ground truth of the object (GUO et al. 2020). In other words, one approach is required to continuously scan the physical object and determine its size and shape, while another is required to compare the manufactured object with its as designed model. To fully capture the object through a data capturing technique, such as Terrestrial laser scanning (TLS) in our case, however, demands that different scanning stations be aligned into one coordinate system as well. To this end, scanning data from various scanning stations on site must be co-registered together and then accurately co-registered to the as-designed digital model for further inspections.
There are a variety of approaches for co-registration of a point cloud to a model or another point cloud (CHENG et al. 2018). The most common algorithms are distance minimization approaches, such as ICP-like algorithms (ICP: iterative closest point (BESL & MCKAY 1992). Direct co-registration, in contrary, is of advantage since it is independent of the actual object geometry and problems imposed by an iterative approach are not expected. When direct co-registration uses artificial targets on the object, it is required that those are seen in both, the captured data and the model. This requirement is often not practical since during object production the placing of markers is not possible. ICP-based algorithms can only work in principle if the shape and size of objects to be co-registered are similar. If objects have different size or vary in geometric details, for instance in the problem at hand, the model has another size than the actual realization the algorithm will tend to average and distribute the error between the two objects as shown in Fig. 1. Fig. 1 shows an as-designed model (wax material) registered with the as-printed model created from the captured point cloud using TLS. Fig. 1 (left) indicates that deviations are averaged out, due to the ICP registration between the two models. This is obvious on the top surface, ICP method indicates that the we need more material, while Fig. 1 (right), where a direct co-registration is used, a subtraction of the material is needed where we didn't average out the error. The used direct co-registration approach will be explained in Section Experiment Setup & Results3. Also, ICP is known to be susceptible to local minima if surfaces are ambiguous or the initial solution is not

close enough to the optimum. Fig. 2 shows the registration result of the two identical point clouds. However, the ICP-based registration algorithm ended where still a bias between the models in the z-direction exists.

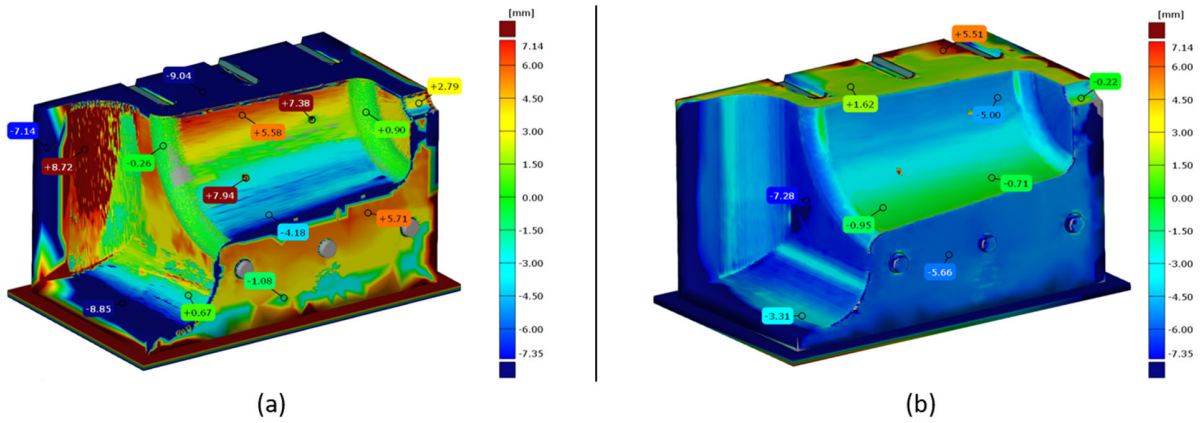


Fig. 1: As-printed model vs. As-designed model. (a) registered with ICP algorithm, (b) Target-based registration

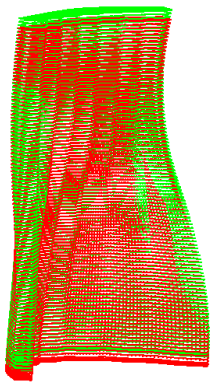


Fig. 2: ICP-based registration of two synthetic data from a wall model

Furthermore, ICP relies on rough co-registration of models, a process that can be time-consuming especially where time is an important production factor. In general, in order to enable proper co-registration and convergence to the correct optimum, the surface geometry in the overlapping area needs to resemble at least three major normal directions which are not parallel.
This paper will introduce a direct registration method between the TLS-captured point-cloud of printed object with the as-designed model. In the next section we will discuss related work surrounding registration. Our own experimental, direct co-registration setup and results will be presented in Section 3. Lastly, our conclusions and future ideas will be discussed in Section 4.

2 Related Work

Co-registration is a fundamental task that brings two or more data sets into one common coordinate system. Due to availability of various sensors which could capture different data from the objects

of interest, the need to coherently combine these data sets is crucial. One of the most common approach in surveying for registration is artificial target-based registration. Targets must be placed in suitable locations in the capture area for co-registration. In the usual case of a three-dimensional object-space at least three non-colinear targets must be placed between adjacent scanning stations. The other common method for registration is ICP, in typical registration pipelines proposed by (RUSINKIEWICZ & LEVOY 2001), the registration of two-point clouds can be split into the following steps: i) Selection: The sampling of the input point clouds. ii) Matching: Estimating the correspondences between the points in the subsampled point clouds. iii) Rejection: Filtering the correspondences to reduce the number of outliers. iv) Alignment: Assigning an error metric, and minimizing it to find the optimal transformation.

Firstly, in the selection method of (BESL & MCKAY 1992) all the available points are used, which is time consuming. On the other hand, the normal-space sampling strategy is good for incised mesh surfaces that guaranties enough samples on the groove to bring the surfaces into alignment (RUSINKIEWICZ & LEVOY 2001). Secondly, there are a variety of matching strategies such as the shortest Euclidian distance (Closest point) from aligned entity to the reference entity data (BESL & MCKAY 1992). On the other hand, other researchers computed the distance between the two aligned entities in the direction of the source point's normal (CHEN & MEDIONI 1991). Nonetheless, filtering strategies are important since they are used to reduce the noise in the data for registration to exclude their effect. One method is to reject pairs with distances more than 2.5 times the standard deviation (MASUDA et al. 1999). Another method is based on rejecting the points on the boundaries (TURK & LEVOY 1994). Excluding pairs, that include points on mesh boundaries, is especially useful for avoiding erroneous pairings, that cause a systematic bias in the estimated transformation in scenarios where the overlap between scans is not completed. Lastly, error metrics are used to measure the error between the two entities as point-to-point that utilized the squared of sum distances (BESL & MCKAY 1992). However, other methods such as point-to-plane are used to find the optimal transformation. This method minimizes the sum of the squared distances between a point and the tangent plane at its correspondence point (CHEN & MEDIONI 1991). Each iteration generally is slower than the point-to-point version, however, often at significantly better convergence rates.

The output from TLS is a discrete point cloud which is a sampling point of the captured object and does not reflect all of its details, in particular the boundaries. In addition, the point cloud usually contains noise resulting from uncertain sensor measurements, environmental influences or surface material properties. If ICP should be used to co-register two TLS-based point clouds or a TLS-based point cloud and other data, it requires an initial registration of the datasets. Thus, many researchers provided an improvement to the algorithm (YANG et al. 2016; ZHOU et al. 2016). A global optimal solution for registration (Go-ICP) using the scheme based on a branch-and-bound (BnB) that searches the entire 3D motion space proposed by (YANG et al. 2016). Go-ICP, does not require an initial alignment between the two data sets, however, the algorithm is considered to be time consuming. While, Fast Global Registration (FGR) proposed by ZHOU et al. (2016), considered to be efficient. Nonetheless, FGR does not guarantee an optimal solution and might fail if the noise is too high.

Numerous researchers utilized TLS for QC of construction elements (NGUYEN & CHOI 2018; WANG et al. 2018; BUSWELL et al. 2020; MABOUDI et al. 2020; TANG et al. 2022). However, different scans are still required to be registered to further proceed QC. WANG et al. (2018) extracted the geometric features like edges from the point cloud in order to compare them with their corresponding features in the CAD model. Hence, no co-registration needed between the as-printed object with its as-designed model. Nevertheless, the authors used ICP algorithm to register the different scan stations. On the other hand, other researchers performed a distance-based deviation analysis between the printed object and its digital model (MABOUDI et al. 2020; NGUYEN & CHOI 2018). Point-to-plane ICP-based registration was employed for fine registration between CAD model and the point cloud (NGUYEN & CHOI 2018).
In our case, during the production process, after each deposition layer of the material, the printed object has to be scanned and finally checked with its digital model as a means of quality control. To insure an automatic, streamlined and fast quality inspection procedure, our aim is to register the scanned data directly to the as-designed model.

3 Experiment Setup & Results

Our experiments are performed in the Digital Building Fabrication Laboratory (DBFL) at the Technical University of Braunschweig. The fabrication room contains two robots (Fig. 3): one for shotcrete 3D printing (SC3DP) (a printing robot) and one for subtracting the material from the printed object (a milling robot). Each robot has its own coordinate system and they are already coregistered together. In addition, this fabrication coordinate system is co-registered to the model space, that is, the coordinate system where models are designed.

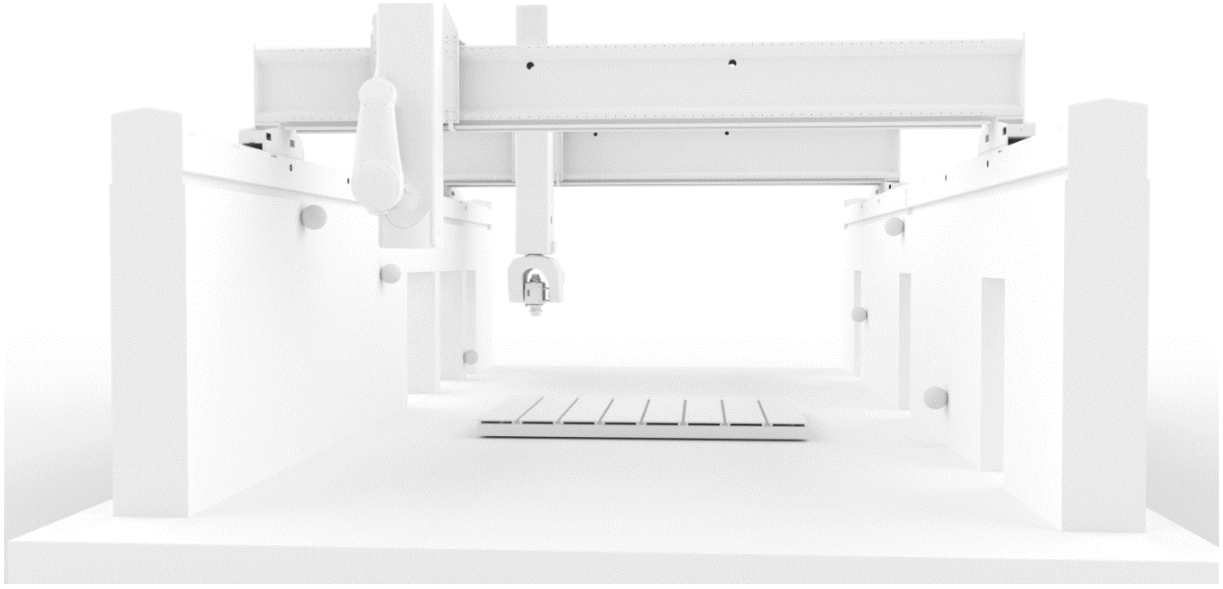


Fig. 3: Digital Building Fabrication Laboratory

Co-registration is thus required between different scan stations, and is also needed for comparison of the manufactured object and the as-designed model. Accordingly, our co-registration workflow

is divided into two main steps: i) Co-registration of scanning stations, and ii) co-registration with the as-designed model as conducted by the milling robot in our fabrication coordinate system.
A terrestrial 3D laser scanner Z+F IMAGER 5010X is utilized for scanning the printed object. For scanning, a closed loop of five scanning stations were placed in the fabrication room. To solve our registration task, the network planning in the fabrication lab for the points' measurement was done, as shown in Fig. 4.

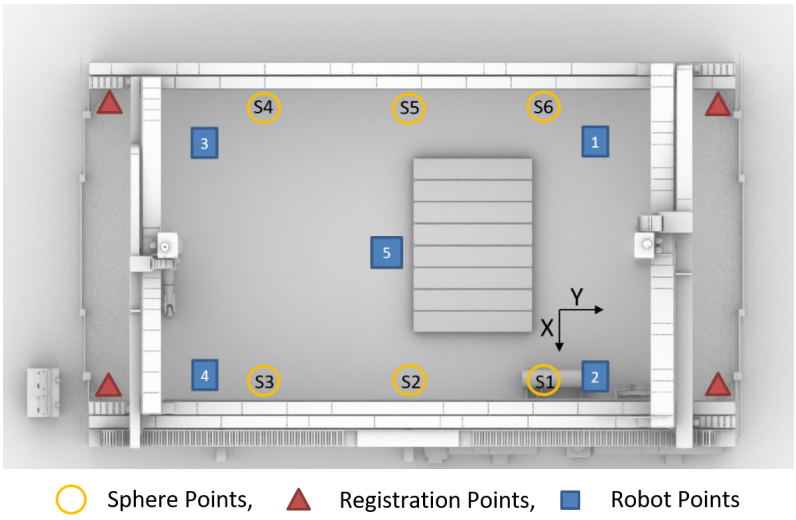


Fig. 4: Surveying network planning in the Fabrication room

In our experiment, we customized and combined a classical surveying network measurement and TLS measurement in order to bringing the as-built data to the as-designed model coordiante system.
We have three coordinate systems:

1- Surveying coordinate system: established using a Leica MS50 multistation.
2- Milling robot coordinate system (fabrication coordinate system)
3- TLS coordinate system.

The idea is that using surveying measurements, we establish a high quality framework which is connected to the fabrication coordinate system and could be used for any further data capturing in this fabrication room. To this end, first the surveying coordinate system was established. The measured points are the four registration points as well as the six sphere points using MS50, accurately. Best practices for surveying measurements (such as reading on the both faces of MS50 and considering the enviromental parameters) are employed to achieve a high quality network. The measurements of the spheres by MS50 were realized by measuring a prism that has the same height of the sphere radius. After the network adjustment, we achieved 1 mm for the maximum standard deviation of the points coordinates. Furthermore, we moved the robot to 5 points in the fabrication room (robot points in Fig. 4) and measured the position of the endpoint of the robot's using MS50 in our reference system. The maximum and minimum standard deviation of the

adjusted network are 1.4 mm and 1mm, accordingly. A calibrated prism is attached to the robot to obtain its position on each of these five points as is illustrated in Fig. 5.

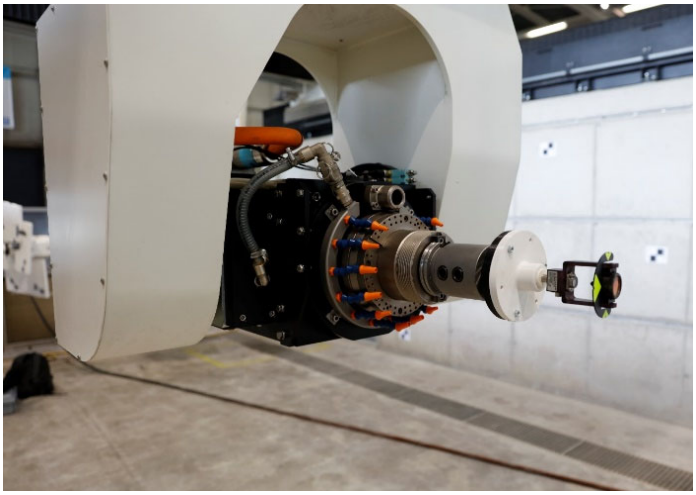


Fig. 5: Prism installed on the robot end effector

To summarize, we measured three sets of the points (Fig. 4) in our reference surveying coordinate system:

1- Registration points: Our surveying network's points to establish a reference coordinate system. There is no need to measure these points in the future experiments.
2- Robot points: To solve the transformation parameters between robot coordinate system and surveying coordinate system. Here also, there is no need to measure these points in the future experiments.
3- Sphere points: To localize TLS in all current and future experiments.

Therefore, in all TLS data capturing sessions in this fabrication room, we can localize our TLS using sphere points and transform the point cloud to the fabrication system directly and hence also a transformation into model space is possible. The positioning precision of the robots themselves has been determined to be better than a millimeter. The quality of our transformation, in fact, is in 0.5 mm range as it is shown in Tab. 1. The coordinates of five common points (milling robot points) are used to compute the parameters of a 3D Helmert transformation between the surveying reference system and robot system. The residuals of these points are listed in Tab. 1.

Tab. 1: MS50 (Surveying Coord. Sys.) to milling robot transformation coordinate system

	Residuals [mm]			
Robot point	3D	X	Y	Z
1	0.41	0.33	-0.24	-0.04
2	0.52	-0.25	0.45	0.06
3	0.57	0.17	-0.49	0.23
4	0.74	-0.61	0.00	-0.42
5	0.49	0.36	0.28	0.17
RMSE	0.56	0.38	0.34	0.23

The coordinates of the six spheres are calculated in the fabrication coordinate system by MS50 as well. These coordinates were used to transform the different scans to the room system. The overall registration of the point cloud with the surveying network resulted in a maximum deviation of 2.7 mm. The results of the registration are given in Tab. 2.

Tab. 2: Distances of spheres in coregistered point cloud to MS50 measurements

Sphere point	mm
S1	2.7
S2	1.6
S3	2.6
S4	2.1
S5	1.5
S6	2.2
Average Deviation	2.1
Standard Deviation	0.5

As it is shown in Tab. 2, the average 3D distance of the six sphere points (Fig. 4) to the corresponding points in our surveying coordinate system is 2.1 mm which is acceptable for most of the inspection processes in AMC.
We believe that the main source of these residuals is the remaining error in our surveying network (1.4 mm) accumulated by the co-registration error of our TLS stations (0.5-1.2 mm). However, the proposed strategy makes the inspection of AMC products in the fabrication room more straightforward and reliable against limitations of ICP-based approaches. The registration between the as-printed point cloud captured with TLS & the wax as-designed model is more reliable using the direct registration (Fig. 1).

4 Conclusion

This paper introduced co-registration of the printed object (as-built) and its as-designed model in context of additive manufacturing in construction. The most common methods for co-registering different entities are ICP-based methods. ICP, however, is time-consuming and often fails to achieve registration without coarse registration and is also susceptible to get stuck in local minima. In addition it assumes that the shape and size of reference and target objects are similar.
This paper expressed our experiments for employing and customizing the classic surveying and TLS measurement to coregister the as-built and as-designed model in additive manufacturing. After establishing an accurate surveying network, we used a 3D transformation to connect the fabrication (model) coordinate system to our TLS measurements. Therefore, for each of future measurements, we can transform the point clouds to the model coordinate system for quality inspection and deviation monitoring of the printed objects.

The analyses of the transformation between the fabrication coordinate system and the robot coordinate system demonstrated 0.5 mm RMSE. However, the transformation quality between scanned data and the fabrication room was in range of 3 mm. This is mainly due to remaining error in our surveying network, and also coregistration errors of TLS stations.
The obtained results are a step forward in increasing the reliability and automation of co-registration between the as-design model and its printed counterpart.
Further experiments are required to improve the accuracy and efficiency of the quality control process. One possible direction could be to employ scan planning before data capturing. This is especially very effective in our problem while the as-designed model of each object is already provided for printing process.
In order to increase the automation and skip the error prone steps of our coregistration process, we also conceptualized an idea to directly mount the TLS on the robot (Fig. 6). This provides the possibility to freely move the TLS in the printing environment, but also to directly transform the derived point clouds to the model space.

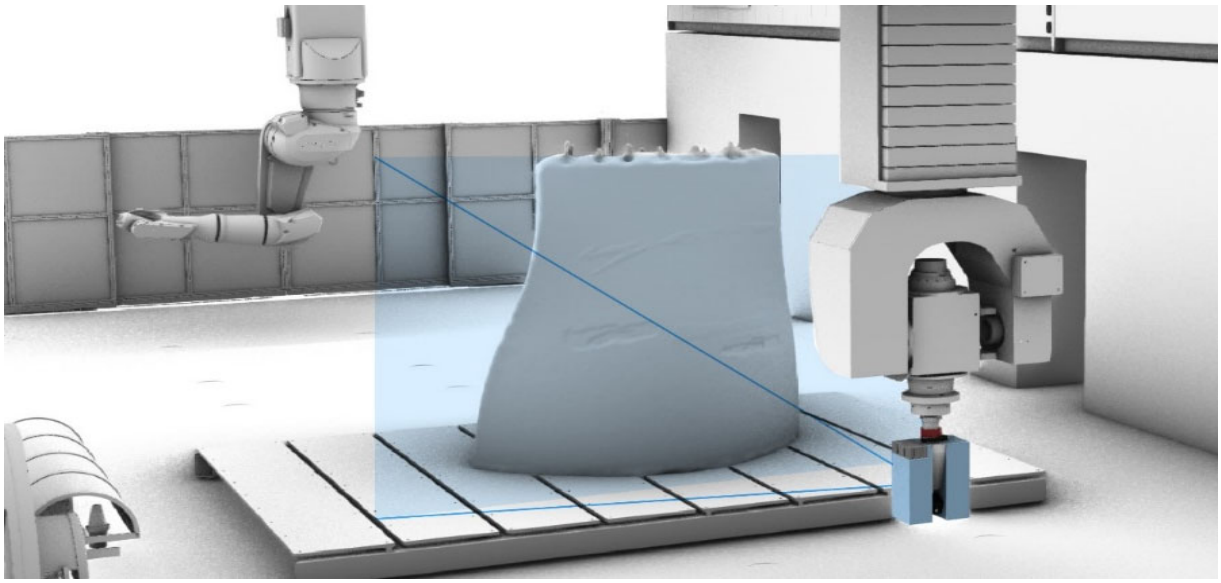


Fig. 6: Conceptual representation of mounting the TLS on the robot

Moving the TLS freely in the printing room, provides the possibility to scan those parts of the objects which cannot be captured through classical TLS setup (tripod on the ground). Moreover, 3D scan planning would be more applicable to increase the efficiency of data capturing.

5 Acknowledgments

This research is funded by the German Research Foundation (DFG) - Project number 414265976 – TRR 277. The authors would like to thank the DFG for the support within the SFB/Transregio 277- Additive manufacturing in construction. (Subproject C06).

6 Bibliography

AGARWAL, R., CHANDRASEKARAN, S. & SRIDHAR, M., 2016: Imagining Construction's Digital Future. Mckinsey & Company Global Insititute. https://www.mckinsey.com/business-functions/operations/our-insights/imagining-constructions-digital-future, last Accessed on 20.01.2022.

BESL, P. J. & MCKAY, N. D., 1992: A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, **14**(2), 239-256, https://doi.org/10.1109/34.121791.

BUCHLI, J., GIFTTHALER, M., KUMAR, N., LUSSI, M., SANDY, T., DÖRFLER, K. & HACK, N., 2018: Digital in Situ Fabrication - Challenges and Opportunities for Robotic in Situ Fabrication in Architecture, Construction, and Beyond. Cement and Concrete Research, **112**, 66-75, https://doi.org/10.1016/j.cemconres.2018.05.013.

BUSWELL, R. A., KINNELL, P., XU, J., HACK, N., KLOFT, H., MABOUDI, M., GERKE, M., MASSIN, P., GRASSER, G., WOLFS, R. & BOS, F., 2020: Inspection Methods for 3D Concrete Printing. Digital Concrete, 1-14, https://doi.org/10.1007/978-3-030-49916-7_78.

CHEN, Y. & MEDIONI, G., 1991: Object Modeling by Registration of Multiple Range Images. Proceedings. 1991 IEEE International Conference on Robotics and Automation, 2724-2729, https://doi.org/10.1109/ROBOT.1991.132043.

CHENG, L., CHEN, S., LIU, X., XU, H., WU, Y., LI, M. & CHEN, Y., 2018: Registration of Laser Scanning Point Clouds: A Review. Sensors, **18**(5), 1641, https://doi.org/10.3390/s18051641.

EL-SAYEGH, S., ROMDHANE, L. & MANJIKIAN, S., 2020: A Critical Review of 3D Printing in Construction: Benefits, Challenges, and Risks. Archives of Civil and Mechanical Engineering, **20**(2), 34, https://doi.org/10.1007/s43452-020-00038-w.

EL JAZZAR, M., PISKERNIK, M. & NASSEREDDINE, H., 2020: Digital Twin in Construction : An Empirical Analysis. EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering, Berlin, 501-510.

GUO, J., WANG, Q. & PARK, J. H., 2020: Geometric Quality Inspection of Prefabricated MEP Modules with 3D Laser Scanning. Automation in Construction, **111**, https://doi.org/10.1016/j.autcon.2019.103053.

MABOUDI, M., GERKE, M., HACK, N., BROHMANN, L., SCHWERDTNER, P. & PLACZEK, G., 2020: Current Surveying Methods for the Integration of Additive Manufacturing in the Construction Process. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **43**(B4), 763-768, https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-763-2020.

MASUDA, T., SAKAUE, K. & YOKOYA, N., 1996: Registration and Integration of Multiple Range Images for 3-D model Construction. 13th International Conference on Pattern Recognition, **1**, 879-883, https://doi.org/10.1109/ICPR.1996.546150.

NGUYEN, C. H. P. & CHOI, Y., 2018: Comparison of Point Cloud Data and 3D CAD Data for on-Site Dimensional Inspection of Industrial Plant Piping Systems. Automation in Construction, **91**, 44-52, https://doi.org/10.1016/j.autcon.2018.03.008.

PLACZEK, G., BROHMANN, L., MAWAS, K., SCHWERDTNER, P., HACK, N., MABOUDI, M. & GERKE, M., 2021: A Lean-based Production Approach for Shotcrete 3D Printed Concrete Components. 38th International Symposium on Automation and Robotics in Construction (ISARC 2021), 811-818, https://doi.org/10.22260/ISARC2021/0110.

RUSINKIEWICZ, S. & LEVOY, M., 2001: Efficient Variants of the ICP Algorithm. Third International Conference on 3-D Digital Imaging and Modeling, 145-152, https://doi.org/10.1109/IM.2001.924423.

TANG, X., WANG, M., WANG, Q., GUO, J. & ZHANG, J., 2022: Benefits of Terrestrial Laser Scanning for Construction QA/QC: A Time and Cost Analysis. Journal of Management in Engineering, **38**(2), https://doi.org/10.1061/(ASCE)ME.1943-5479.0001012.

TURK, G. & LEVOY, M., 1994: Zippered Polygon Meshes from Range Images. Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '94, 311-318, https://doi.org/10.1145/192161.192241.

WANG, Q., SOHN, H. & CHENG, J. C. P., 2018: Automatic As-Built BIM Creation of Precast Concrete Bridge Deck Panels Using Laser Scan Data. Journal of Computing in Civil Engineering, **32**(3), https://doi.org/10.1061/(ASCE)CP.1943-5487.0000754.

YANG, J., LI, H., CAMPBELL, D. & JIA, Y., 2016: Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration," in IEEE Transactions on Pattern Analysis and Machine Intelligence, **38**(11), 2241-2254, https://doi.org/10.1109/TPAMI.2015.2513405.

ZHOU, Q. Y., PARK, J. & KOLTUN, V., 2016: Fast Global Registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9906 LNCS, 766-782, https://doi.org/10.1007/978-3-319-46475-6_47.

Fernerkundungs-Aktivitäten zu großräumigen hydrologischen Extremereignissen im Kontext des Nationalen Hochwasserschutzprogramms (NHWSP)

SILKE MECHERNICH[1] & ROBERT WEIß[1]

Zusammenfassung: Im Fall großräumiger hydrologischer Extremereignisse wie Hoch- oder Niedrigwasser, sind Vorbereitungen erforderlich um ein gesichertes Monitoring zu gewährleisten. Dieser Beitrag betrachtet die übergeordneten Anforderungen zur Vorbereitung eines Monitorings. Entsprechend wird die Notwendigkeit der Koordination mit allen beteiligten Institutionen ebenso benannt, wie die Notwenigkeit homogener und aktueller Geobasisdaten rund um relevante Gewässer. Des Weiteren wird beispielhaft auf die Ableitung von Wasser-Land-Grenzen aus satellitengestützten Radardaten (Sentinel-1, unabhängig von Tageslicht und Wolken) eingegangen und erste Ergebnisse des Projektes Sat-Land-Fluss vorgestellt.

1 Einleitung und Motivation

Während der materielle Schaden bei Hochwasserereignissen offensichtlich und die Bevölkerung stark betroffen ist, liegen die Schäden bei Niedrigwasserereignissen v.a. im wirtschaftlichen und ökologischen Bereich. So konnten beim Niedrigwasser am Rhein 2018 monatelang die Schiffe nur mit geringer Beladung fahren, was u.a. zu Lieferengpässen von Produkten für die chemische Industrie, Treibstoff und pharmazeutischen Waren führte. Der wirtschaftliche Schaden war so groß, dass 2019 seitens des damaligen Bundesministeriums für Verkehr und Digitale Infrastruktur (BMVI, heute Bundesministerium für Digitales und Verkehr, BMDV) der Aktionsplan Niedrigwasser Rhein (8-Punkte Plan) ins Leben gerufen wurde. Auf der anderen Seite treten immer wieder Hochwasserereignisse auf. Die Hochwasserkatastrophe im Bereich Ahr im Juli 2021 mit ihren unfassbaren Schäden, Verlusten von mindestens 220 Menschenleben sowie unzähligen Verletzten und Obdachlosen hat das Jahr 2021 stark überschattet.

Um in solchen Extremereignisfällen agieren und reagieren zu können ist es zwingend erforderlich den jeweiligen Extremereignisverlauf vorher, währenddessen und im Nachgang nachvollziehen zu können. Hierzu ist eine Fülle von quantitativen Daten gefordert. In Anbetracht der erforderlichen Schnelligkeit der Daten- und Produktverfügbarkeit im Ereignisfall, bietet die Fernerkundung ein großes Potential. Um eine effiziente Datenbereitstellung im Ereignisfall zu erreichen, ist es erforderlich, dass alle Beteiligten und betroffene Verwaltungen die Interessen und Anforderungen an die Daten und Produkte möglichst vorab koordinieren und konkrete Maßnahmenpakete vorbereiten. Dementsprechend arbeiten Bundes- und Landesbehörden u.a. im Rahmen des Nationalen Hochwasserschutzprogramms (NHWSP) zusammen. Hierbei müssen mehrere Handlungsfelder berücksichtigt werden (s.a. BfG, BBK, BKG 2019), zu denen in der BfG aktuell u.a. Projekte mit den folgenden Zielen und Schwerpunkten verfolgt werden:

[1] Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, D-56068 Koblenz, Deutschland, E-Mail [mechernich, weiss]@bafg.de

- Koordination und Vorbereitung für den Ereignisfall
- Erforderliche Geobasisdaten
- Forschung und Entwicklung (F&E) zur Unabhängigkeit von Wetterbedingungen

Im Folgenden geben wir für jedes dieser Punkte je ein aktuelles Projektbeispiel.

2 Koordination und Vorbereitung für den Ereignisfall

Um im Ereignisfall vorbereitet zu sein und alle erforderlichen Daten fristgerecht, wirtschaftlich und gesichert in der erforderlichen Qualität zu erfassen, ist ein koordiniertes und strukturiertes Vorgehen erforderlich. Hierzu wurde beispielhaft mit den Konzeptarbeiten auf Bund-Länder-Ebene zur Erstellung eines Messprogramms „Quantitative hydrologische Extremereignisse – Elbe und wichtige Nebenflüsse“ (Messprogramm Extreme Elbe) begonnen.

2.1 Das Projekt Messprogramm Extreme Elbe (Quantität)

Dieses Messprogramm umfasst die Themen Wasser-Land-Grenzen, Abflussmessungen und Wasserspiegelfixierung. Wesentlicher Kernpunkt des Messprogramms ist die Erstellung einer exakten Handlungskette im Ereignisfall, sowie aller nötigen Unterlagen für die Bearbeitung. Dies beinhaltet v.a. folgende Klärungen:

- der Warn- und Auslösestufen,
- des Bearbeitungsgebietes (Gewässerstrecke, Vorschlag in Abb. 1; im Fall der Wasser-Land-Grenzen Bestimmung inkl. Breite des Messbereichs),
- der Verantwortlichkeiten der Koordination zur Erfassung der Daten und Erstellung der Produkte,
- einer eindeutigen fachlichen Beschreibung der erforderlichen Produkte, ggf. in Form einer oder mehrerer Leistungsbeschreibungen für Vergaben (inkl. Messhäufigkeit und Produktarten, inkl. Qualitätsbeschreibung),
- ein Marktüberblick zu in Frage kommenden Firmen für Vergaben,
- Produktbereitstellung (wo, wie schnell, usw.) und die
- Finanzierung.

Der aktuelle Vorschlag für das Monitoring von Wasser-Land-Grenzen ist multidisziplinär und berücksichtigt folgende Standbeine:

- die Nutzung von Satellitendatenauswertungen des Copernicus-Dienstes des Katastrophen- und Krisenmanagements (CEMS) „Rapid Mapping“ zur schnellen Einschätzung der Lage (Beispiel in Abb. 2),
- Bildflüge zum Zeitpunkt des Hochwasser-Scheitels bzw. Niedrigstwasserstandes für eine qualitativ hochwertige und lückenlose Erfassung (Beispiel in Abb. 3),
- ggf. auch lokale Messungen von Hochwassermarken bzw. Geschwemmsellinien im Ereignisfall.

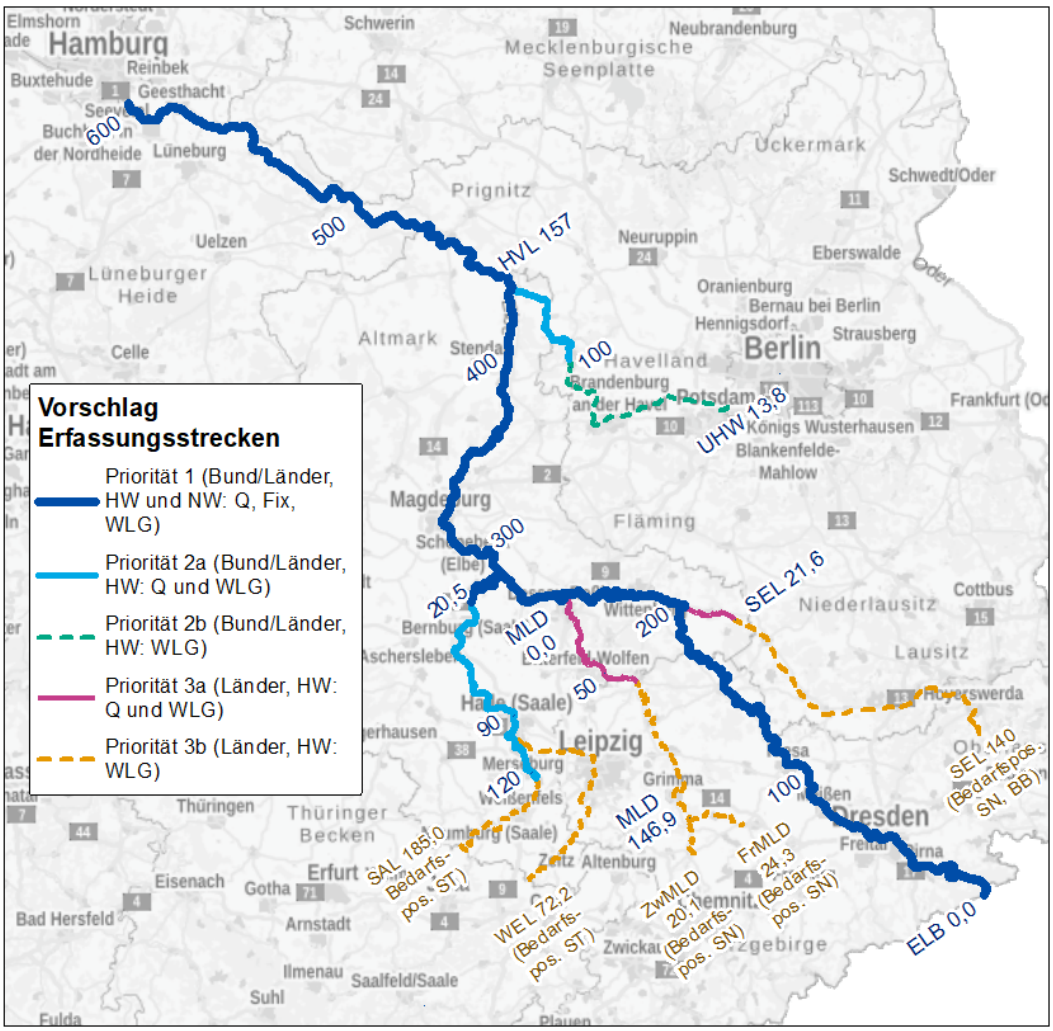


Abb. 1: Gegenwärtiger Vorschlag der zu erfassenden Flusstrecken und deren Priorisierung HW: Hochwasser, NW: Niedrigwasser, Q: Abflussmessungen, Fix: Wasserspiegelfixierungen, WLG: Wasser-Land-Grenzen-Bestimmungen, ELB: Elbe, SEL: Schwarze Elster MLD: (Vereinigte) Mulde, SAL: Saale, UHW: Untere Havelwasserstraße

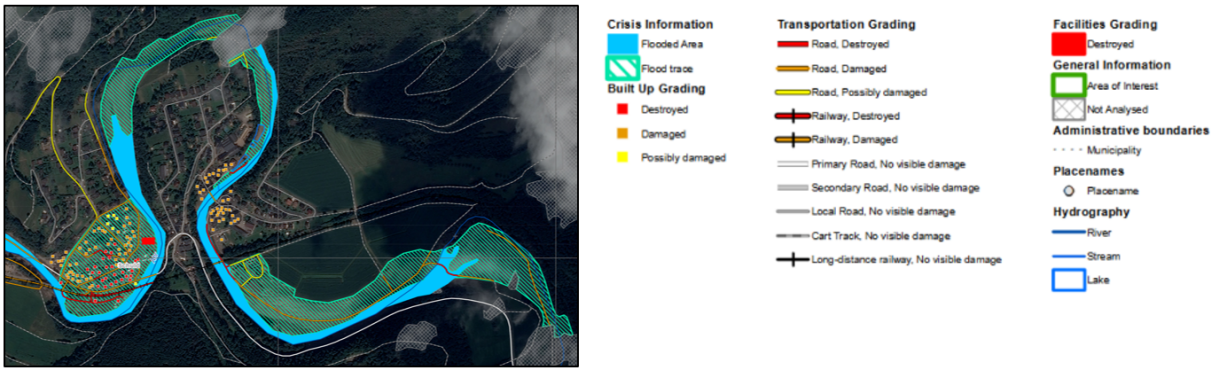


Abb. 2: Ausschnitt eines Produktes des Copernicus-Dienstes für Katastrophen- und Krisenmanagement (CEMS) der Aktivierung EMSR517, Hochwasser 07/21 an der Ahr (COPERNICUS EMERGENCY MANAGEMENT SERVICE 2021)

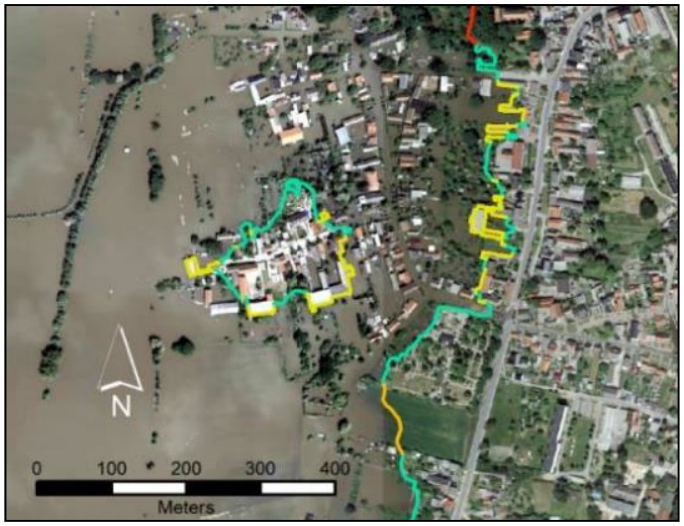


Abb. 3: Digitales Orthophoto des Hochwassers 2013 an der Elbe und die Wasser-Land-Grenzen, attributiert nach ihren Lage- und Höhenunsicherheiten (Karte basierend auf Daten von BfG 2015)

Neben den fernerkundungsbasierten Produkten sind auch die Abflussmessungen und Wasserspiegelfixierungen Gegenstand dieses Projektes. Abflussmessungen dienen der Erfassung des aktuellen Abflusses im Sinne einer Angabe, wieviel m³ Wasser pro Sekunde das Gewässer durchströmen. Diese Messungen werden schiffs- oder USV-basiert mit Acoustic Doppler Current Profiler (ADCP) Systemen durchgeführt. Eine Wasserspiegelfixierung dient der Erfassung der aktuellen Wasseroberfläche entlang der Gewässerachse. In der heutigen Zeit werden entsprechende Fixierungen mit schiffsbasierten Global Navigation Satellite Systems (GNSS) -Messtechniken durchgeführt.

Um die unterschiedlichen Bedarfe v.a. von Bund und Ländern abzudecken, wirken zahlreiche Institutionen der Länder sowie des Bundes mit. Auf die Art ist sichergestellt, dass die Erstellung von Mehrzweck-Produkten für verschiedenste Anwendungen gegeben ist. Die Projektgruppe setzt sich daher aus Bundesvertretern der Bundesanstalt für Gewässerkunde (BfG, Projektleitung) und des Wasserstraßen- und Schifffahrtsamt Elbe (WSA Elbe), sowie aus Vertretern verschiedener Landesdienststellen zusammen. Im Detail wirken seitens der Länder folgende Institutionen mit:

- Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG SN)
- Landestalsperrenverwaltung des Freistaates Sachsen (LTV SN)
- Staatliche Betriebsgesellschaft für Umwelt und Landwirtschaft (BfUL SN)
- Landesamt für Umwelt Brandenburg (LfU BB)
- Landesbetriebes für Hochwasserschutz Sachsen-Anhalt (LHW ST)
- Staatliches Amt für Landwirtschaft und Umwelt Westmecklenburg (StALU WM)
- Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN)
- Landesbetrieb für Küstenschutz, Nationalpark und Meeresschutz Schleswig-Holstein (LKN.SH)

Vor dem Hintergrund der zahlreichen erforderlichen Absprachen und insbesondere der verbindlichen Klärung der Finanzierung soll das Messprogramm 2024 operationell sein. Es wird angestrebt das Projekt als verstetigtes Messprogramm zu etablieren.

3 Erforderliche Geobasisdaten

Für vielfältige Grundlageanwendungen, wie z.B. die Wasserstandsvorhersagen, die Simulation von Hochwasser-Wellenabläufen, oder die Berechnung von HW-Risikokarten, sind aktuelle, homogene und flächendeckende Geobasisdaten erforderlich. In diesem Zusammenhang wird momentan ein aktualisiertes und homogenes Digitales Geländemodell des Wasserlaufs (DGM-W) der Elbe erstellt.

3.1 Das Projekt DGM-W Elbe

Das Bearbeitungsgebiet des Bund-Länder-Projektes umfasst die Binnenelbe von der tschechischen Grenze bis Hamburg, zzgl. der Unteren Havelwasserstraße von Rathenow bis zur Mündung in die Elbe (Abb. 4). Geodätisch erfasst und modelliert werden dabei hochwasserrelevante Gebiete und die Gebiete der Altauen. In Summe werden 4285 km² entlang von 723 Fluss-km bearbeitet.
Die Daten, welche der DGM-W Modellierung zugrunde liegen, entstammen v.a. einer ALS-Datenerfassung die im Rahmen des Projektes erfolgt, aus aktuellen Echolotdaten, sowie erstmalig in einem DGM-W Projekt auch aus flächendeckenden bathymetrischen ALS-Daten (Niedrigwasserbefliegung 2018, KUEHNE 2021). Des Weiteren werden, wie bei jedem DGM-W, Bauwerksdaten der Wehre und Schleusen mit eingebunden. Die Verknüpfung der verschiedenen Datensätze stellt eine Herausforderung dar, da Datensätze unterschiedlicher Sensoren, unterschiedlichen Alters und verschiedenen geometrischen Eigenschaften etc. zu einem homogenen Modell fusioniert werden sollen.
Kernprodukte sind neben dem DGM-W auch die klassifizierten Laserdaten, ein DOM und Querprofile.
Die Projektleitung erfolgt bei diesem Projekt durch die BfG, wobei innerhalb von Projektgruppen alle beteiligten Institutionen des Bundes und der Länder mit eingebunden sind. Im Wesentlichen werden Ingenieurleistungen im Bereich der Datenerfassung und Datenaufbereitung vergeben. Die Kosten hierfür sind anteilig auf Bund und Länder verteilt.
Das Gesamtbearbeitungsgebiet ist in drei Lose aufgeteilt, die separat bearbeitet werden (vgl. Abb. 5 und 6). Die ALS-Befliegung des gesamten Projektgebietes findet im Winter 2021/2022 statt. Neben der Vorgabe zur Vegetations-, Eis- und Schneefreiheit sind weiterführende Rahmenbedingungen, wie etwa definierte Wasserstände bei den Befliegungen einzuhalten. Dies führt dazu, dass die Anzahl der potenziellen Flugtage im Winter relativ gering ist und jedes Flugfenster zu nutzen ist.

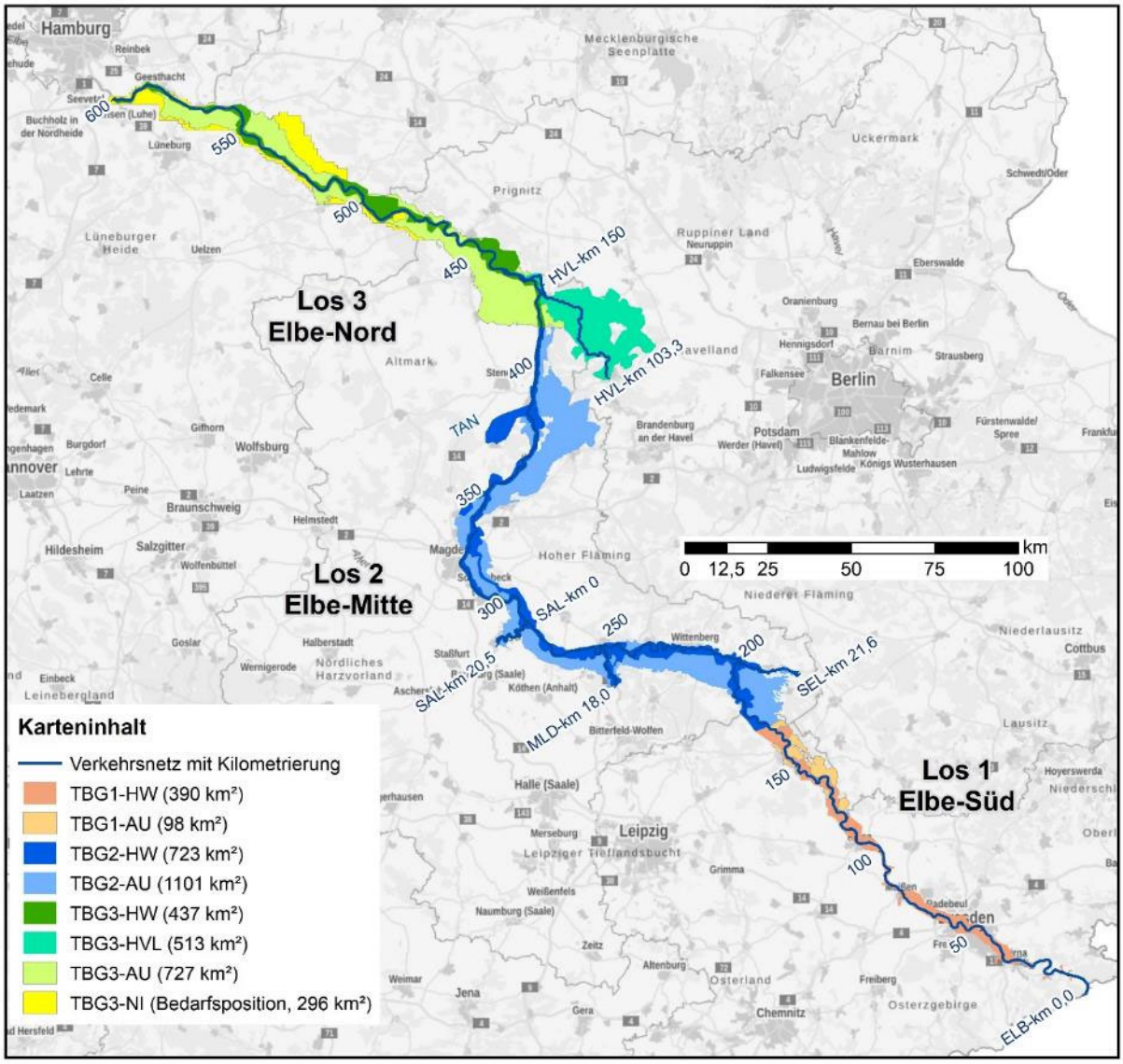


Abb. 4: Bearbeitungsgebiet des DGM-W Elbe inkl. der Gebietsunterteilungen, TBG: Teilbearbeitungsgebiet, HW: Hochwasser, AU: Altauen

Im staugeregelten Bereich der Unteren Havelwasserstraße (UHW) sind die Wasserstände im Winter standardmäßig relativ hoch, so dass auch außerhalb der UWH viele Flächen überflutet sind (Abb. 6). Nach einer detaillierten Abwägung der Interessen aller Beteiligten und betroffenen Institutionen, auch außerhalb der Projektgruppe, wurde ein gemeinsames Vorgehen abstimmt. Hierbei wurden über einen abgestimmten Zeitraum Zielwasserstände in der UHW definiert, die entsprechend der Planung realisiert werden konnten. Die Befliegung dieses kritischen Bereichs von 513 km² konnte im Dezember 2021 abgeschlossen werden. Dies war möglich, nachdem im Interesse aller Beteiligten die Entscheidung getroffen wurde, eine Befliegung innerhalb der Dämmerung durchzuführen, wodurch zwar in einem Teilgebiet keine nutzbaren Luftbilder als Sekundärprodukt erzeugt wurden, aber dafür der Zeitraum der Wasserstandsabsenkung möglichst kurzgehalten wurde.

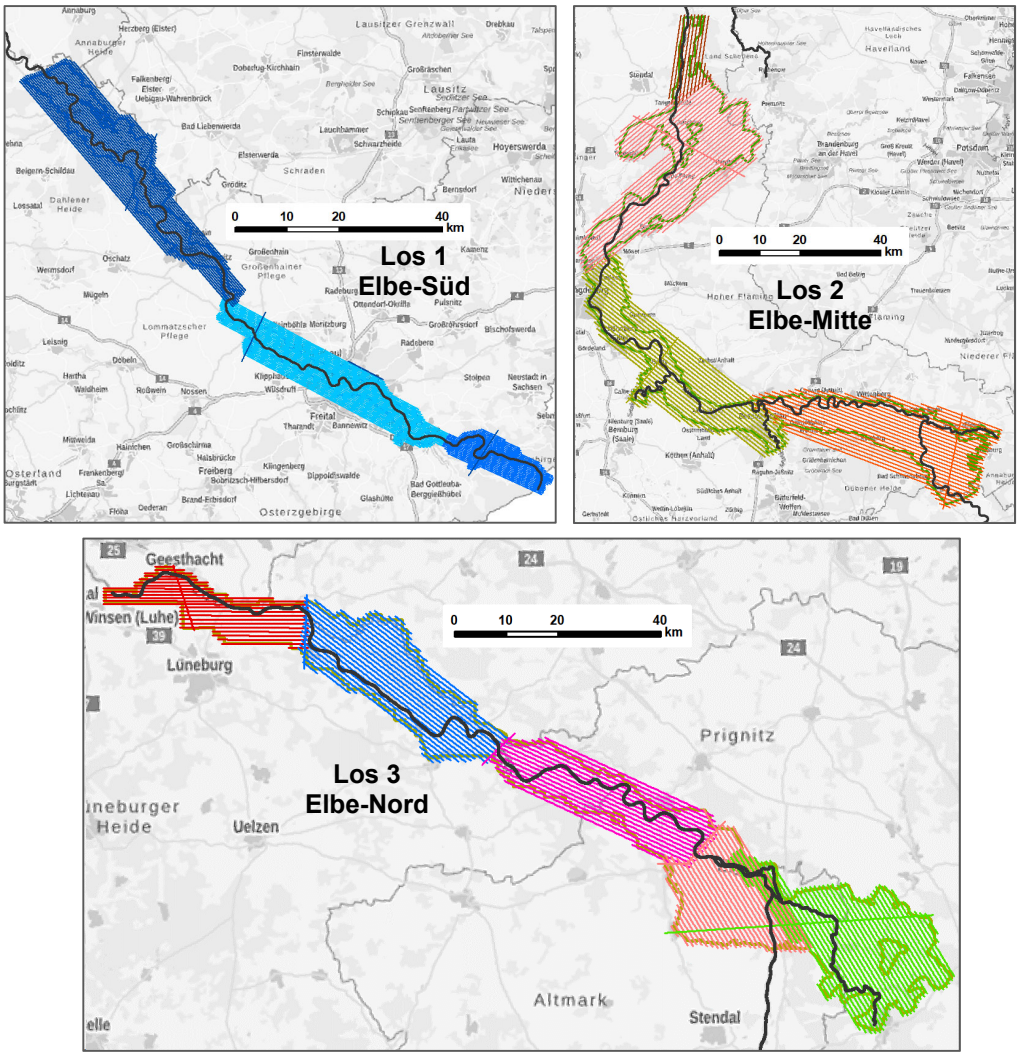


Abb. 5: Blöcke der Befliegung in Los 1 (links), Los 2 (Mitte) und Los 3 (rechts)

Für die freifließende Strecke der Elbe werden täglich die vorhergesagten Wasserstände geprüft und entsprechend die Flugfreigaben erteilt. Bis zum 15.01.2022 wurden ca. 15 % (550 km²) der Elbe-Flächen erfasst.

Neben der aktuellen konventionellen ALS Befliegung fand im Kontext des DGM-W Elbe auch eine flächendeckende Befliegung mit einem bALS System zum Zeitpunkt Niedrigwassers 2018 statt. Diese Befliegung war insofern neu, als dass hierbei erstmalig in Deutschland bei einem so großen Projekt eine flächendeckende Full-Waveform Aufzeichnung durchgeführt wurde und die so erfassten Daten auch einer vollständigen FWF Reprozessierung unterzogen wurden. Insofern stellt diese Befliegung das Maximum dessen dar, was aktuell mit der bALS Technik möglich ist. Aktuell befindet sich das Projekt im Zeitplan und sollte Ende 2024 abgeschlossen sein. Die dann vorliegenden Daten werden frei verfügbar sein.

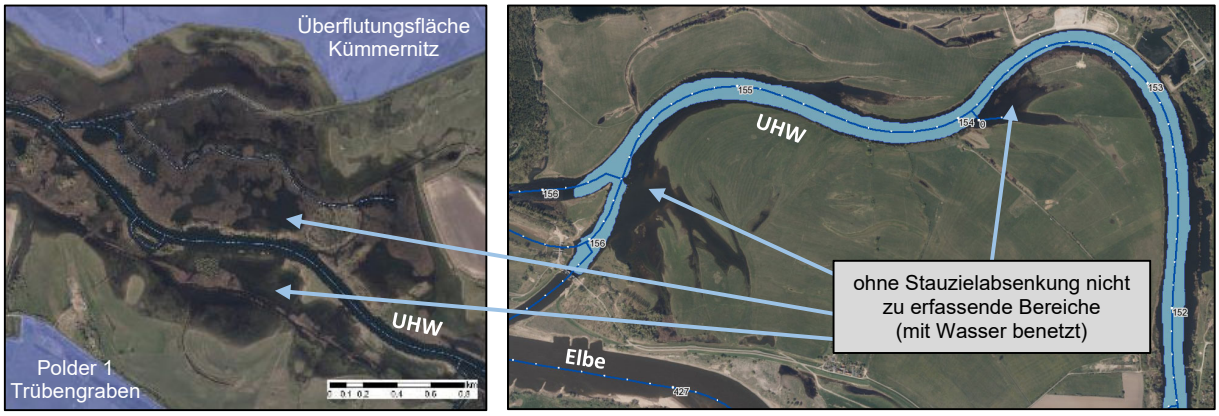


Abb. 6: Gebiet der UHW mit angrenzenden überfluteten Flächen, Hintergrund: DOP-DE ©GeoBasis-DE/BKG (2021)

4 F&E zur Extremereigniserfassung unabhängig von den Wetterbedingungen

Die meisten Messsysteme sind abhängig von guter Sicht und/oder ausreichenden Lichtverhältnissen (ALS, optisch). Synthetic Aperture Radar (SAR) Systeme haben dagegen die Eigenschaft Wolken zu durchdringen und als aktives Messsystem können sie somit jederzeit, unabhängig von Tageslicht und Wetterverhältnissen, angewandt werden. Entsprechend bietet die qualitätsgesicherte Auswertung von Radardaten ein sehr großes Potential um Extremereignisse gesichert erfassen zu können (MARTINIS et al. 2015).

V.a. zur Identifikation von glatten Wasseroberflächen eignen sich die SAR-Daten sehr gut. Meist wird hierbei die Intensität der Rückstreuung genutzt, da eine glatte Wasseroberfläche eine spiegelnde Eigenschaft aufweist, sodass die dazugehörigen Pixel eine sehr geringe Rückstreuung aufzeigen (MARTINIS et al. 2015; Abb. 7). Im Rahmen des europäischen Copernicus-Programms werden mit Sentinel-1A und -1B zwei SAR-Satelliten betrieben, die entsprechende Radaraufnahmen generieren können. Zwar liefern diese SAR-Systeme in regelmäßigen und bekannten Abständen neue Daten, jedoch entstehen aufgrund der relativ niedrigen Satellitenauflösung und anderen Problemen Datenlücken und Unsicherheiten, die in vorhandenen Produkten weitgehend undokumentiert sind.

Aufgrund der schrägen Aufnahmegeometrie kommt es bei SAR-Bildern zu geometrisch bedingten Abbildungseffekten, wie Foreshortening, Layover, Lagefehlern bewegter Objekte, und Abschattungen, die durch unebenes Gelände verursacht werden. Vor diesem Hintergrund ist die Aufarbeitung von Radardaten sehr komplex. Insbesondere die Abschattung (z.B. durch Steilkanten, Vegetation oder Gebäuden) stellt eine Herausforderung dar, da in diesem Bereich ähnlich zu Wasserflächen keine Rückstreuung auftritt. In Gebieten mit starker Vegetation und in urbanen Gebieten kann es zudem zu verstärkten, abgelenkten bzw. Mehrfachreflexionen kommen. Die Rauheit von Wasseroberflächen z.B. durch Windeinwirkung stellt eine weitere Herausforderung dar, da in diesem Fall das Wasser auch eine deutliche Rückstreuung aufweist (MARTINIS & RIEKE 2015).

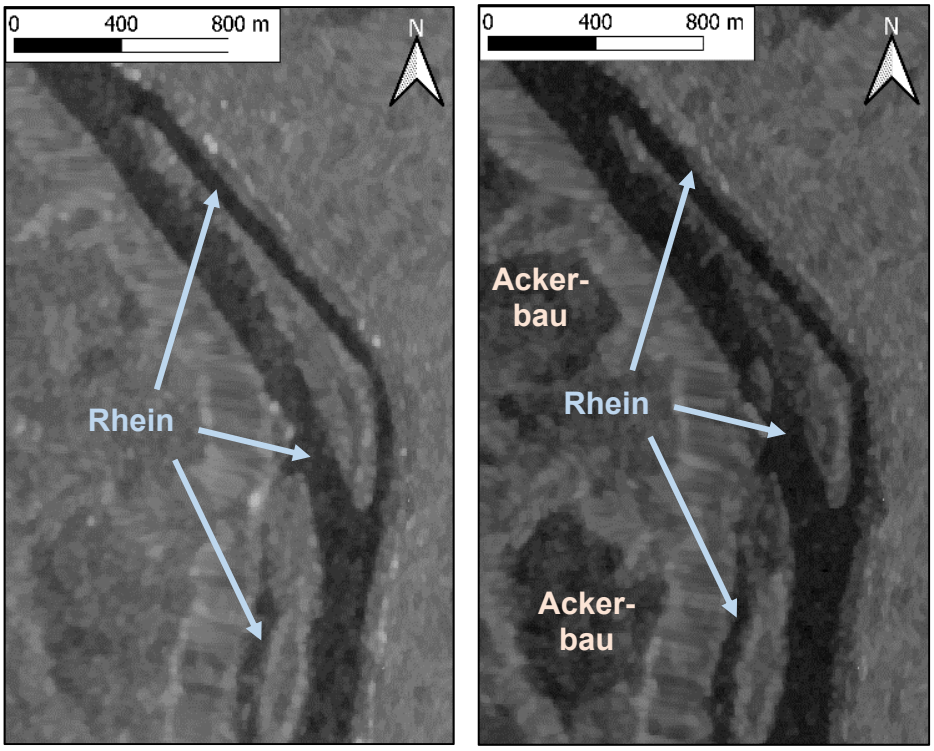


Abb. 7: Rückstreuintensitäten am Mittelrhein von der Sentinel-1B Szene vom 17.11.2018 (descending), der Polarisationen VV (links), VH (rechts); in schwarz/dunkelgrau sind Bereiche niedriger Rückstreuintensität dargestellt, die hier v.a. durch die glatte Wasseroberfläche des Rheins hervorgerufen wird

Es sind verschiedene Produktanbieter für Hochwasserereignisse verfügbar, deren Dienste entsprechend aufbereitete SAR-Daten bereitstellen. Zu nennen wären hierbei z.B. das „Global Flood Monitoring“ des Copernicus-Dienstes für Katastrophen- und Krisenmanagement (CEMS). Hier wird jedoch ein Großteil der Flächen ausgeschlossen (z.B. abgeschattete Gebiete, urbane Gebiete, Wald), was die Nutzung für vielfältige Zwecke stark einschränkt. Des Weiteren erstellt die CEMS-Komponente „Mapping“ unter der Voraussetzung einer Aktivierung schnelle Kartenprodukte aus verfügbaren Satellitendaten. Zusätzlich existieren verschiedene thematisch spezialisierte, oft kommerzielle Dienste. Allen gemeinsam ist, dass die Validierung der Kartenprodukte nicht oder nur sehr begrenzt durchgeführt wurde. Im Kontext dieser Punkte wurde das Projekt Sat-Land-Fluss initiiert.

4.1 Das Projekt Sat-Land-Fluss

In dem laufenden BfG-Projekt „Sat-Land-Fluss“ (Satellitenbasierte Wasser-Land-Grenzen-Bestimmung, DLR-Förderaufruf „Nationale Nutzung Copernicus“) erforscht und entwickelt die BfG das Verbesserungspotential der Satellitendatenauswertung. Hierfür werden Sentinel-1 Daten aufbereitet und mit zusätzlichen Geobasisdaten sowie hydrologischen Daten kombiniert. Ziel ist es, u.a. exemplarisch herauszuarbeiten wie hoch der Mehrgewinn dieser Zusatzdaten ist.
Die Anwendungsthemen des Projektes sind dabei einerseits Extremereignisse im Binnenbereich von Bundeswasserstraßen, aber auch die hochdynamische Situation im Tidebereich.

Methodisch wird in einem ersten Schritt eine geometrische und radiometrische Aufbereitung der Rückstreuintensitäten von S1-Szenen (Modus Interferometric Wide Swath, IW, Level 1 Single Looking Complex, SLC) mittels Sentinel Application Platform (SNAP) durchgeführt. Ein beispielhaftes Ergebnis ist in Abb. 7 dargestellt. Basierend hierauf wird in Anlehnung an das Verfahren von TWELE et al. (2016) automatisiert ein Grenzwert bestimmt. Die Anwendung dieses Grenzwertes ist für den Niedrigwasserfall 2018 am Mittelrhein in Abb. 8a dargestellt. Die reale WLG ist hier sehr gut bekannt, da kurz zuvor bei einem gleichen Wasserstand eine Bildbefliegung (im Auftrag des WSA Oberrhein, Abwicklung durch die BfG) stattgefunden hat. Die stereoskopisch bestimmte WLG ist in Abb. 8 gelb dargestellt und verdeutlicht die Defizite, die die automatisiert abgeleitete Wassermaske der jeweiligen Polarisation ergibt. Entsprechend wird in den nächsten Schritten zur Ausgrenzung von Flächen in Höhenlagen der Algorithmus des „height above nearest drainage“ (HAND, NOBRE et al. 2011) angewandt, wobei hier die DGM5 Daten des Bundesamtes für Kartographie und Geodäsie (BKG) zugrunde gelegt werden. Des Weiteren wird die Lokalität der Fahrrinne genutzt zur Eliminierung von Schiffsrückstreuung sowie zur feingliedrigeren Auslösung im Fall schmaler Gewässerabschnitte. Die deutlichen Verbesserungen sind in Abb. 9b dargestellt.

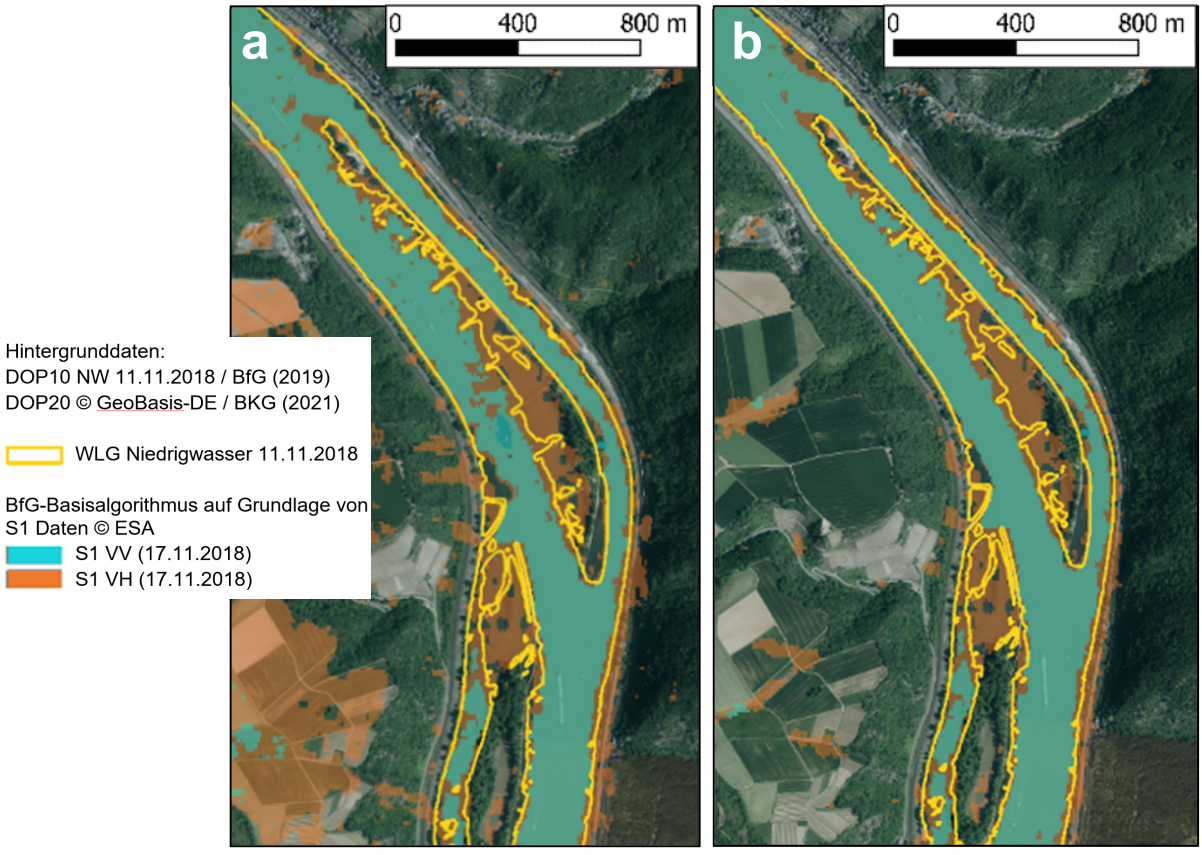


Abb. 8: Beispiel der ersten Schritte der Ableitung einer Wassermaske für das Niedrigwasser 2018 am Mittelrhein mit dem Basisalgorithmus von Sat-Land-Fluss, (a) Wassermaske aus den Polarisationen VV und VH basierend auf der Sentinel-1 (S1) Datenanalyse mit SNAP, (b) verbesserte Wassermaske der Polarisationen VV und VH nach Integration des HAND-Algorithmus und der Lage der Fahrrinne; WLG: Wasser-Land-Grenze, basierend auf stereoskopischer Auswertung eines Bildfluges

Weiterhin bestehen Herausforderungen, v.a. durch

- Wasseroberflächenrauheit oder angelegte Schiffe
- Abschattungen (Gelände, Bebauung, Vegetation)
- Durchfeuchtung oder relativ glatte Oberflächen
- Die niedrige räumliche Auflösung der S1-Daten

Diese Punkte werden aktuell durch Integration weiterer Zusatzdaten bearbeitet, v.a. durch Integration hochaufgelöster Geobasisdaten (u.a. DGM-Ws) sowie durch vorhergesagte bzw. gemessener Wasserstände. Die nachfolgende Validierung der Wassermasken und WLGs hat einen hohen Stellenwert. Das Projekt arbeitet in enger Kooperation mit verschiedenen Verwertungspartnern auf Bundesebene. Zu nennen wären hierbei vor allem das Bundesamt für Bevölkerungsschutz und Katastrophenhilfe (BBK), das Bundesamt für Kartografie und Geodäsie (BKG) mit dem SKD, das Bundesamt für Seeschifffahrt und Hydrografie (BSH) und die Wasserstraßen und Schifffahrtsverwaltung des Bundes (WSV).
Gemeinsam mit den Verwertungspartnern werden weiterführende WLG-Produkte erstellt, mit dem Fokus der Anwenderorientierung und Nutzbarkeit. Denkbar wären hier z.B. Zeitserien für Überflutungsflächen, jährliche Niedrigwasserstände oder die Identifikation Veränderungshotspots. Insgesamt können bei erfolgreichem Projektverlauf WLGs und weitere Produkte relativ kurzfristig und mit einem sehr geringen monetären Aufwand in höherer Qualität abgeleitet werden, als sie aktuell verfügbar sind.

5 Fazit & Ausblick

Insgesamt zeigen diese Projekte beispielhaft den Umfang erforderlicher Vorbereitungen für ein Monitoring hydrologischer Extremereignisse und verdeutlichen, dass zahlreiche Methoden, Kompetenzen und Koordinationen erforderlich sind. Insbesondere vor dem Hintergrund der Wirtschaftlichkeit ist eine optimale Abstimmung aller Beteiligten und aller Bedarfsträger zwingend geboten. Es muss das Ziel sein, alle Bedürfnisse seitens der involvierten Institutionen zu erfüllen und gleichzeitig die technischen Möglichkeiten vollumfänglich auszuschöpfen. Über allem steht dabei die Herausforderung, alle Aktivitäten so zu bündeln und zu fokussieren, dass ein maximaler fachlicher Profit aus den verwendeten Ressourcen generiert wird.
Mit dem Messprogramm Elbe Extrem soll eine mit allen Beteiligten abgestimmte Blaupause geschaffen werden, wie im Fall eines definierten Extremereignisses zu verfahren ist. Es ist das Ziel alle relevanten Daten nur einmal in hochwertiger Qualität zu erfassen und Redundanzen zu vermeiden. Dabei sollen alle Ergebnisse allen Beteiligten zur Verfügung stehen. Dieses Programm steht somit stellvertretend für eine abgestimmte und vorbereitete Vorgehensweise während eines Extremereignisses.
Im Vorfeld ist es die Aufgabe verschiedenster Institutionen auf entsprechende Extremereignisse vorbereitet zu sein. Diese benötigen aktuelle und vor allem hochwertige Geobasisdaten. Auch hier gilt es, die eingesetzten Ressourcen bestmöglich zu nutzen. Im Fall des DGM-W Elbe agieren Bund und Länder gemeinschaftlich und schaffen mit dem DGM-W Elbe ein homogenes Produkt, welches alle Anforderungen erfüllt. Durch die Kooperation zwischen Bund und Ländern können

die bestehenden Ressourcen bestmöglich genutzt werden. Ein weiterer Nebeneffekt ist eine Harmonisierung der Geobasisdaten zwischen Bund und den Ländern.
Die beiden vorgenannten Punkte des Messprogramms und des DGM-W Elbe stellen aus technischer Sicht den aktuellen Status Quo dar. Neue Techniken, insbesondere in der satellitengestützten Radarfernerkundung versprechen ein großes Potenzial, das es zu erschließen und nutzbar zu machen gilt. Im Kontext dieser Aufgabe sind aktuelle F&E Maßnahmen nötig, um die speziellen Bedürfnisse verschiedenster Bedarfsträger innerhalb Deutschlands zu befriedigen und die Möglichkeiten des europäischen Copernicus Programms optimal zu nutzen.
Die hier benannten Projekte der BfG und anderer Institutionen zeigen wegweisende Bausteine auf. Bis zum Erreichen des Ziels sind noch einige Meilensteine und vielfältige Ansätze zu bewältigen.

6 Danksagung

Die Autoren danken für die stets kooperative Mitarbeit der Vertreter der oben genannten Bundes- und Landesbehörden. Besonderer Dank geht dabei an die BfG-Kollegen Marcus Hatz und Jörg Belz für die zahlreichen Arbeiten für das Messprogramm Extreme Elbe, sowie an Bastian Gessler für seine Arbeit an dem Projekt Sat-Land-Fluss.
Des Weiteren danken wir der Geschäftsstelle der FGG-Elbe, insbesondere Wenke Kahrstedt, für die Möglichkeit der Durchführung der Projekte DGM-W Elbe und Messprogramm Extreme Elbe.
Nicht zuletzt gilt unser Dank dem BMDV sowie dem Projektträger des Deutschen Zentrum für Luft- und Raumfahrt (DLR) für die Förderung des Projektes Sat-Land-Fluss (Förderkennzeichen 50EW2015).

7 Literaturverzeichnis

BFG, BBK & BKG, 2019: Ergebnisdokumentation zum Workshop „Großräumiges Hochwassermonitoring - Möglichkeiten, Grenzen, Chancen der Fernerkundung“ am 05./06.12.2018. https://www.bafg.de/DE/05_Wissen/02_Veranst/2018/2018_12_05_ergebnisse.pdf?__blob=publicationFile, letzter Zugriff am 14.01.2022.

BFG, 2015: Flugzeuggestützte Erfassung der Hochwasserereignisse an Elbe und Donau, 2013 (PhgHW-Auswertung Elbe/Donau), Los 7: Fischbeck. Produkt: 3D-Wasser-Land-Grenzen (3DWLGren). Fertigstellung 21.01.2015.

BMVI, 2019: Gemeinsame Erklärung des Bundesministers für Verkehr und digitale Infrastruktur und Vertretern der Stahl -, Chemie- und Mineralölindustrie, der Produzentenmineralischer Massenrohstoffe und des Binnenschifffahrtsgewerbes zur Sicherstellung zuverlässig kalkulierbarer Transportbedingungen am Rhein („8-Punkte-Plan“), https://www.vci.de/themen/logistik-verpackung/logistik/8-punkte-plan-gegen-niedrigwasser-binnenschiffahrt-gemeinsame-erklaerung-bmvi-verbaende.jsp, letzter Zugriff am 14.01.2022.

COPERNICUS EMERGENCY MANAGEMENT SERVICE, © 2021 European Union: [EMSR517] Floods in Western Germany. Bad Neuenahr-Ahrweiler: Grading Product, Monitoring 1, version 3, release 1, RTP Map #02, letzter Zugriff am 14.01.2021.

COPERNICUS EMERGENCY MANAGEMENT SERVICE: Mapping https://emergency.copernicus.eu/mapping/, letzter Zugriff am 14.01.2022.

COPERNICUS EMERGENCY MANAGEMENT SERVICE: Global Flood Monitoring, https://gfm.portal.geoville.com/products, letzter Zugriff am 14.01.2022.

KUEHNE, E., 2021: New Opportunities for Capturing the Topography of the River Elbe by Airborne Hydromapping in a Low Discharge Period 2018. PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **89**, 177-182. https://doi.org/10.1007/s41064-021-00151-2.

MARTINIS, S., KERSTEN, J. & TWELE A., 2015: A fully automated TerraSAR-X based flood service. ISPRS Journal of Photogrammetry and Remote Sensing, **104**, 203-212; https://doi.org/10.3390/rs10040583.

MARTINIS, S. & RIEKE, C., 2015: Backscatter Analysis Using Multi-Temporal and Multi-Frequency SAR Data in the Context of Flood Mapping at River Saale, Germany. Remote Sensing 7, 7732-7752, https://doi.org/10.3390/rs70607732.

NOBRE, A.D., CUARTAS, L.A., HODNETT, M., RENNÓ, C.D., RODRIGUES, G., SILVEIRA, A., WATERLOO, M. & SALESKA, S., 2011: Height Above the Nearest Drainage – a hydrologically relevant new terrain model. Journal of Hydrology, **404**, 13-29, https://doi.org/10.1016/j.jhydrol.2011.03.051.

TWELE, A., CAO, W., PLANK, S. & MARTINIS, S., 2016: Sentinel-1-based flood mapping: a fully automated processing chain. International Journal of Remote Sensing, **37**(13), 2990-3004, http://dx.doi.org/10.1080/01431161.2016.1192304.

Genauigkeitsbetrachtung voxelbasierter Änderungsdetektion im Gebäudeinnenbereich zur automatisierten Baufortschrittsüberwachung

THERESA MEYER[1,2], ANSGAR BRUNN[2] & UWE STILLA[1]

Zusammenfassung: Dieser Beitrag handelt von einer Methode zur Änderungsdetektion im Gebäudeinneren, die auf der Diskretisierung des Objektraums durch Voxel basiert. Die Belegung des Voxelraums wird anhand von Punktwolken aus terrestrischem Laserscanning zu verschiedenen Bauphasen ermittelt. Durch die Aufdeckung von Belegungskonflikten zwischen zwei Messepochen wird ermittelt, ob Änderungen an einer bestimmten Stelle aufgetreten sind. Es wird gezeigt, wie sich der Diskretisierungsfehler als Resultat der Voxelisierung auf die noch erreichbare räumliche Auflösung einer Änderungsaussage auswirkt. Am Beispiel von Abbrucharbeiten bei einem Innenumbauprojekt werden die wechselseitigen Auswirkungen von Aufnahmegeometrie, Scanauflösung und Voxelgröße diskutiert. Die Ergebnisse zeigen, dass die Auflösung von 3D-Punktwolken unter Beachtung einiger Anforderungen an die Aufnahmegeometrie bei Innenraum-Messungen ausreicht, um bei gängigen Genauigkeitsanforderungen auf Voxelebene automatisiert strukturelle Änderungen zu beschreiben.

1 Einleitung

In der Baubranche besteht schon lange ein großer Bedarf nach Methoden zur effizienten Fortschrittsüberwachung von Bau- und Umbauprojekten. Um von Seiten des Baumanagements auf eventuelle Ausführungsfehler frühzeitig reagieren zu können, muss ein regelmäßiger Abgleich von SOLL und IST-Zustand auf der Baustelle erfolgen. Deshalb ist die sog. „as-built" Dokumentation von Bauprojekten ein anspruchsvolles und rasant wachsendes Aufgabenfeld für Geodäten. Änderungsdetektion im Kontext von Baufortschrittsdokumentation ist derzeit von großem Interesse für die gesamte AEC (Architecture, Engineering und Construction) -Branche und BIM (Building Information Modeling). Nicht nur viele positive Erfahrungen in der Baupraxis weisen auf eine starke Korrelation zwischen kontinuierlicher Bauüberwachung und der Einhaltung des Projektzeitplans hin, sondern auch eigens angelegte Untersuchungen und Fallstudien, wie sie beispielsweise in ALIZADEHSALEHI & YITMEN (2019) vorgestellt werden, bestätigen diese positiven Effekte auch aus wissenschaftlicher Sicht. Dennoch stehen bei vielen Bauprojekten nach wie vor kaum Ressourcen für aufwendige metrische Qualitätskontrollen während der Bauphase zur Verfügung, um die Übereinstimmung von Ausführung und Planung zu bestätigen.

Dieser Beitrag handelt von einer Methode zur Änderungsdetektion, die auf 3D-Punktwolken und der Diskretisierung des Objektraums durch Voxel basiert und hier speziell auf die Anwendung im Gebäudeinnenbereich abgestimmt ist. voxelbasierte Änderungsdetektion ist in der Literatur schon

[1] Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstr. 21, D-80333 München, E-Mail: [stilla, theresa.meyer@tum.de]

[2] Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Studienbereich Geo, Labor für Photogrammetrie & Fernerkundung, Röntgenring 8, D-97070 Würzburg, E-Mail: [ansgar.brunn, theresa.meyer]@fhws.de

vielfach thematisiert. HEBEL et al. (2013), detektieren Veränderungen in Airborne Laserscanning Punktwolken. Sie verwenden dabei einen relativ groben Voxelraum für den Vergleich von Punktwolken aus verschiedenen Epochen und verfeinern das Ergebnis der Änderungsdetektion, indem sie einzelne Punkte in vorselektierten Bereichen unter Anwendung der Evidenztheorie eingehender bewerten. Analog analysieren GEHRUNG et al. (2020) Mobile Laserscanning Punktwolken. Andere Autoren wie HUANG et al. (2021) und TUTTAS et al. (2017) demonstrieren voxelbasierte Änderung im Kontext von Baufortschrittsdokumentation mit photogrammetrischen Punktwolken und MEYER et al. (2021) bewerten in diesem Zusammenhang die Genauigkeit von BIM-konformen Gebäudemodellen für die Orientierung von Bildsensoren.

2 Problemstellung

Bislang erfolgt Baufortschrittsdokumentation meist durch Sichtkontrollen und einfache Kontrollmessungen, die von Mitarbeitern durchgeführt werden oder durch sehr zeit- und kostenintensive professionelle Vermessungen mit Totalstationen und Nivelliergeräten. Seit einigen Jahren schon ist das terrestrische 3D Laserscanning (TLS) eine beliebte Alternative für schnelle und hochauflösende Bestandsaufnahmen von Baustellen, jedoch fehlt es dieser Technik noch an gezielten Analyseverfahren, computergestützter Interpretation und Strategien für die automatisierte Evaluation von metrischen Toleranzen in der Ingenieurpraxis. Demnach liegt die Herausforderung beim TLS weniger in der hochauflösenden Bestandserfassung, als vielmehr in der Interpretation dieser Datensätze, mit dem Ziel automatisiert Aussagen über detektierte Änderungen treffen zu können. Insbesondere im Innenausbau gibt es viele detailreiche Zustandsänderungen von Bauteilen. Deshalb sind die Genauigkeitsanforderungen dort besonders hoch. Gleichzeitig herrschen aber im Gebäudeinneren erschwerte Aufnahmebedingungen aufgrund von beengten Bewegungsräumen, eingeschränkten Sichten, GNSS-Abschattung und mehrdeutigen Objektstrukturen, die häufig zu schlechten Aufnahmegeometrien, Verdeckungen, schwacher Georeferenzierung und einer mangelhaften Beobachtungsredundanz führen.

3 Voxelbasierte Änderungsdetektion

Ein Voxel entspricht einer Zelle in einem regelmäßigen 3D-Gitter und mit jedem voxelbasierten Ansatz geht daher von Natur aus die Entscheidung zur Diskretisierung von eigentlich kontinuierlichen Daten einher. Andernfalls wird eine "Voxelisierung" in der Phase der Vorverarbeitung angewandt, um räumliche Abfragen zu beschleunigen und die Datenmenge für anspruchsvolle Analysen in vorselektierten Bereichen zu reduzieren. In diesem Fall dient der Voxelraum dann als 3D-Index (z.B. vom Typ Octree). In den folgenden Abschnitten nutzen wir eine Farbskala (Abb. 1), um klassifizierte Voxel zu kennzeichnen: Aus einer Messepoche werden die Klassen *leer* (Magenta), *unbekannt* (blau) und *belegt* (Cyan) gebildet. Liegen mindestens zwei Messepochen vor, lassen sich Voxel in *verschwunden* (rot), *neu* (gelb) oder *bestätigt* (grün) unterscheiden.

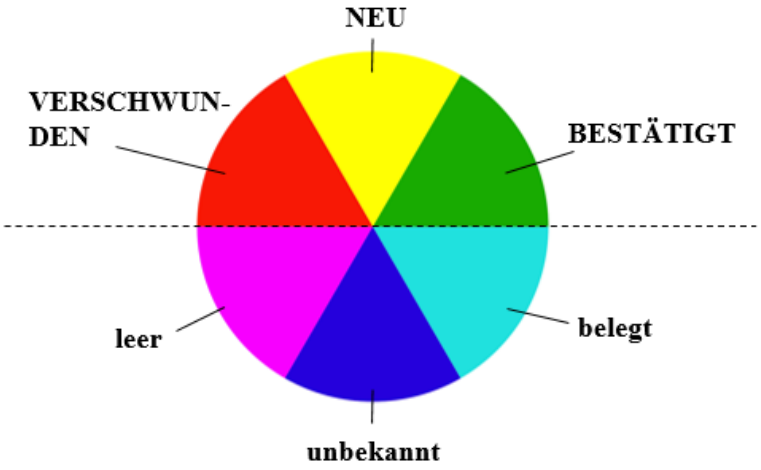


Abb. 1: Farbskala zur Kennzeichnung unterschiedlich klassifizierter Voxel

3.1 Modellierung der Raumbelegung aus einer Messepoche

Eine 3D-Punktwolke wird zu Zwecken der Änderungsdetektion in ein strukturiertes Voxelgitter transformiert, um daraus Erkenntnisse über die Belegung des Raumes abzuleiten. Zu Beginn wird der Voxelraum V initiiert d.h. der gesamte 3D-Raum des Untersuchungsgebietes wird „voxelisiert" bzw. gekachelt und entsprechend der vorgegebenen Voxelgröße diskretisiert. Jeder 3D-Punkt wird seiner entsprechenden Gitterzelle, also seinem Voxel zugeordnet. Die tatsächliche Verteilung und Häufung von 3D-Punkten wird in diesem ersten Schritt nicht berücksichtigt.

Wenn die Anwendung ausschließlich auf Voxelebene arbeiten soll, ist die gewählte Voxelgröße entscheidend für die Genauigkeit und den Detaillierungsgrad. Jeder 3D-Punkt „stimmt" für den Zustand „Belegung" in seinem zugehörigen Voxel. Je mehr 3D-Punkte sich innerhalb eines Voxels befinden, desto größer ist die Annahme von tatsächlicher Belegung des Raumes an dieser Stelle. Für eine endgültige Entscheidung durch Binarisierung muss dann ein geeigneter Schwellenwert gefunden werden, der für jedes Voxel die Entscheidung *leer* (0) oder *belegt* (1) trifft. Dieser Schwellwert hängt stark von der Punktdichte ab. Wenn ein Voxel nur eine begrenzte Anzahl von 3D-Punkten enthält, z. B. nur vereinzelte Punkte, die durch Rauschen entstanden sind, wird es durch den Schwellwert besser als *leer* klassifiziert, anstatt fälschlicherweise als vollständig belegter Raum interpretiert zu werden.

TLS-Punktwolken bergen zusätzliches Wissen über die Belegung des Raums. Ein Laserstrahl traversiert von seinem Ursprung $\boldsymbol{s}$ über den Strahl $\boldsymbol{r}$ zu einem Punkt $\boldsymbol{p}$, der entsteht, wenn der Laserstrahl an einer Objektoberfläche reflektiert wird. Voxel, die weder traversiert werden noch belegt sind verbleiben *unbekannt*. Auf diese Weise können die Zustände der nicht besetzten Voxel in *leer* und *unbekannt* unterschieden werden. Im Zuge der voxelbasierten Modellierung der Raumbelegung aus einer TLS-Punktwolke tritt ein unvermeidlicher Diskretisierungsfehler auf, denn alle Berechnungen und räumlichen Abfragen beziehen sich auf das Voxelzentrum und nicht mehr auf die original gemessenen 3D-Punkte. Das Modell der Raumbelegung ist nur eine Annäherung an die wirkliche Situation und variiert in Abhängigkeit von Voxelgröße, Ursprung und Ausrichtung des Voxelraums (Abb. 2a und 2b). Je nach Voxelgröße und individueller Dichte der Punktwolke fällt dieser Fehler mehr oder weniger ins Gewicht.

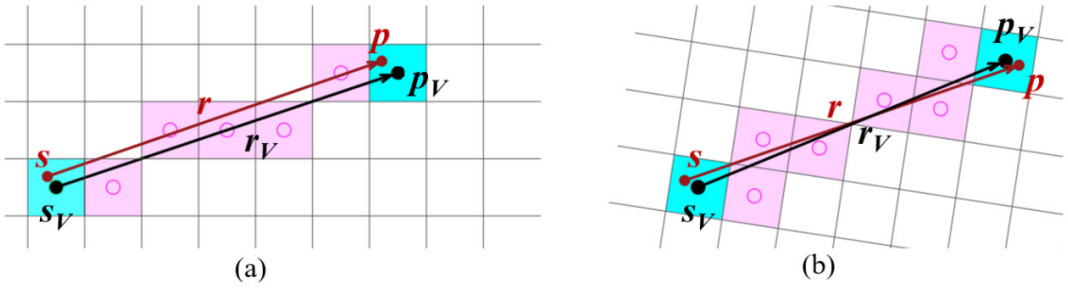


Abb. 2: Auswirkung unterschiedlich angelegter Voxelräume. $\boldsymbol{r_v}$ verbindet $\boldsymbol{s_v}$ und $\boldsymbol{p_v}$ im Voxelraum (schwarz) während $\boldsymbol{r}$ die tatsächliche Lasermessung von $\boldsymbol{s}$ nach $\boldsymbol{p}$ darstellt (rot): a) voxelbasierte Modellierung der Raumbelegung, b) voxelbasierte Modellierung der Raumbelegung unter Verwendung eines anders angelegten Voxelraumes

3.2 Aufdeckung von Belegungskonflikten aus zwei Messepochen

Aus einer einzigen Punktwolke lassen sich bereits viele wertvolle räumliche Informationen über eine Szene ableiten. Eine zweite Punktwolke, die den gleichen räumlichen Bereich abdeckt, ermöglicht eine noch genauere Analyse, da sie entweder die ursprünglichen Annahmen über die Raumbelegung bestätigt oder einen Widerspruch hervorruft. Es gibt zwei Arten von möglichen Konflikten:

I. Ein neuer Strahl $\overrightarrow{\boldsymbol{s_{v2}}\,\boldsymbol{p_{v2}}}$ durchquert ein Voxel, das zuvor als belegt markiert wurde

II. Ein neuer Punkt $\boldsymbol{p_{v2}}$ tritt in einem Voxel auf, das in der ersten Epoche als leer markiert wurde

Die Aufdeckung und Bewertung solcher Konflikte erfolgt durch die Erkennung von veränderten Belegungszuständen (Abb. 3). Verschwundene Punkte des Konflikttyps I deuten auf einen Abriss hin, während solche vom Typ II auf neue Bauelemente oder Störobjekte (z.B. Baustellenmaterial) hinweisen. Das Ergebnis dieses Ansatzes zur Erkennung von Veränderungen ist eine Voxel-Punktwolke, die aus den geometrischen Schwerpunkten der klassifizierten Voxel besteht.

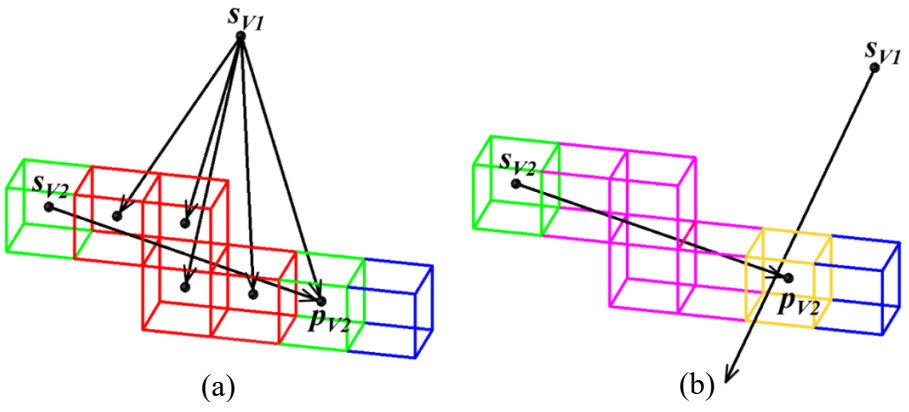


Abb. 3: Mögliche Belegungskonflikte aus zwei Messepochen: a) verschwundene Voxel als Belegungskonflikttyp I, wie z.B. bei einem Abbruch von Bauteilelementen b) neu als „belegt“ hinzu-gekommene Voxel als Belegungskonflikttyp II, wie z.B. beim Aufbau einer neuen Wand oder beim Auftreten von Baustellenmaterial und sonstigen Störobjekten

4 Genauigkeitsanforderungen im Kontext von BIM

Das U.S. Institute of Building Documentation (USIBD 2022) gibt mit der Level of Accuracy (LOA) Spezifikation einen Leitfaden zur Diskussion und Festlegung des geometrischen und räumlichen Genauigkeitsniveaus für professionelle Gebäudedokumentation. Mit der Spezifikation soll ein standardisierter Rahmen geschaffen werden, um konsistente Gebäudebestandserfassung zu ermöglichen. Die LOA des USIBD sind in fünf Zehnerschritten inkrementiert, beginnend mit LOA10 bis LOA50 (Abb. 4). Jede der fünf Stufen kann innerhalb desselben Projekts angewendet werden, denn Genauigkeitsangaben können zwar für einzelne Gebäudeelemente angemessen sein, jedoch nicht unbedingt für das ganze Bauprojekt. Beispielsweise wird gem. USIBD für innenliegende Türen, Wände und Fußböden LOD30 oder LOA40 empfohlen. Die in der LOA-Spezifikation definierten Konfidenzbereiche beziehen sich auf eine Sicherheitswahrscheinlichkeit von 95% (2 Sigma).
Die Angabe eines LOA kann sich entweder auf die Messung (z. B. Laserscanning) oder auf die Modellgenauigkeit beziehen. Letzteres ist besonders wichtig für die Erkennung und Bewertung von Änderungen, wenn es um den Vergleich von Modell (repräsentierte Oberfläche im Referenz-BIM) und erfassten Bestandsdaten (z. B. 3D-Punktwolke) geht.

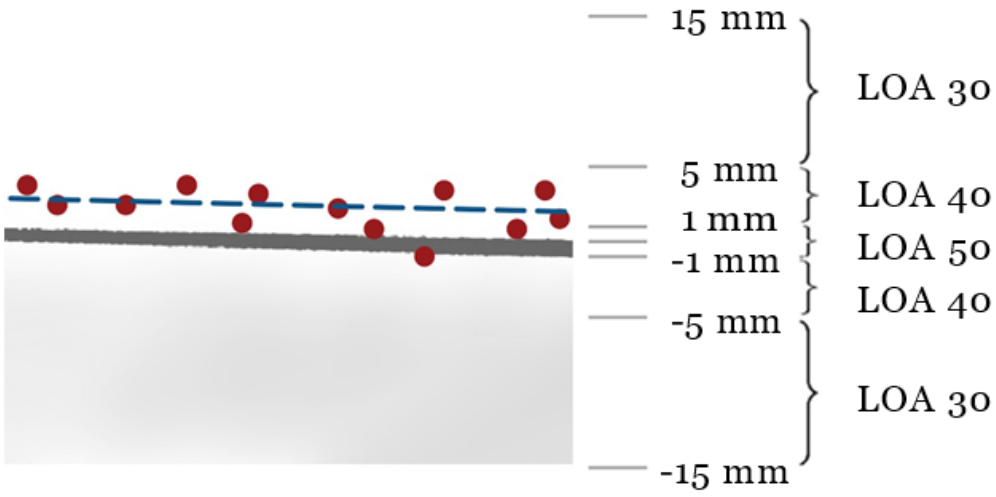
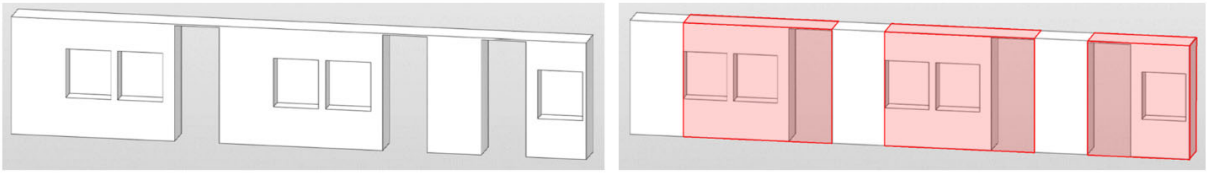


Abb. 4: Darstellung der Level of Accuracy (LOA) gem. USIBD: es gibt eine Diskrepanz zwischen der Bestandsmessung (rote Punkte) der tatsächlichen Objektoberfläche (graue Linie) und der im BIM repräsentierten Oberfläche (blau gestrichelte Linie)

5 Baufortschrittsdokumentation im Gebäudeinnenbereich

Die folgenden Experimente und Ergebnisse beziehen sich auf einen real durchgeführten Umbau im Gebäudeinneren. Das Projekt umfasste im Wesentlichen den Abbruch von nicht tragenden Wandelementen entlang eines langen Flurs, um mehr Platz für einen offenen Seminarraum zu schaffen. Abb. 5 zeigt das entsprechende BIM mit den SOLL-Zuständen vor und nach der Umbauphase.



(a) (b)

Abb. 5: BIM zum Umbauprojekt: a) Modell, das den alten SOLL-Zustand zeigt, b) Modell, das den neu geplanten Zustand zeigt, bei dem nicht-tragende Wandelemente als „abzubrechen“ markiert sind (rot)

(a)

(b)

(c)

Abb. 6: Originale TLS Punktwolken mit RGB-Farben zu den drei Messepochen: a) Epoche 1 (vor Umbau), b) Epoche 2 (während fortschreitender Abbrucharbeiten), c) Epoche 3 (nach Abschluss der Abbrucharbeiten)

5.1 Datenerfassung

Die Abbrucharbeiten erstreckten sich über einen Zeitraum von ca. 3 Wochen im Gebäudeinneren und ein FARO® Focus 3D Laserscanner wurde zur Dokumentation des Projektfortschritts eingesetzt. Diese Dokumentation umfasste 3 Scanepochen: 1) zu Beginn des Projekts, d.h. der alte Zustand, 2) während der laufenden Abbrucharbeiten und 3) nach Abschluss der Abbrucharbeiten,

d.h. der neue Zustand. Jede Messepoche setzt sich aus drei sich überlappenden Punktwolken zusammen. Die relativen Ausrichtungen der Punktwolken innerhalb jeder Epoche wurden mit einem "Plane-to-Plane"-Ansatz realisiert, wobei durchweg gute Registrierungsergebnisse im Submillimeterbereich erzielt wurden. Um die Scandaten mit dem BIM-Referenzsystem abzugleichen, wurden tachymetrisch beobachtete, Schachbrett-Targets verwendet. Die absolute Orientierung jeder Epoche ergab ≈ 1 - 2 mm (die Messungen wurden mit einer Trimble® S6 Totalstation durchgeführt). Die verfügbaren Punktwolken aus den Epochen 1-3 umfassen jeweils mehrere Millionen Punkte und liefern somit sehr hoch aufgelöste räumliche Informationen über die Baustelle. Epoche 1 enthält ≈ 33 Mio. zu verarbeitende Punkte, Epoche 2 ≈ 28 Mio. und Epoche 3 ≈ 35 Mio. Punkte. Jeder Datensatz weist eine recht heterogene Punktdichte auf, was auf die stark variierenden Abstände zwischen Scanner und Objekt in der begrenzten Innenraumumgebung zurückzuführen ist. Im Zielbereich beträgt der durchschnittliche Punkt-zu-Punkt-Abstand ≈ 2 mm.

5.2 Baufortschrittsdokumentation

Die Transformation der Punktwolken in den Voxelraum wurde mit Hilfe der Bibliotheken PCL (Point Cloud Library) und OpenCV in einer C++ Programmierung nach HEBEL et al. (2013) umgesetzt. Die Umrechnung der Weltkoordinaten anhand von Koordinatenoffset und Umrechnungsfaktor (=100 für Weltkoordinaten [m] und 1cm Voxelgröße) in den Voxelraum erfolgte durch Iteration über jedes einzelne Punktwolkenelement. Dabei wurde für jede „Höhenschicht“ ein eigenes zweidimensionales Belegungsgitter ausgegeben. Für das gesamte Untersuchungsgebiet (11,46 × 15,25 × 3,38 m) wurden 338 Belegungsgitter mit je 1146 × 1525 Voxel-Zellen erstellt. Insgesamt entstanden 590.705.700 Voxel bei 1 cm Voxelgröße. Zur Berechnung des Weges des Laserstrahls $\boldsymbol{r_v}$ durch den Voxelraum *V* wurde eine in OpenCV implementierte Variante des Algorithmus von Bresenham (BRESENHAM 1965) verwendet. Der sog. *LineIterator* iteriert über alle Zellen eines Rasterliniensegments, das zwei gegebene Positionen, in dem Fall $\boldsymbol{s_v}$ und $\boldsymbol{p_v}$, miteinander verbindet. Schließlich trägt jede Zelle alle Informationen über ihre Belegung bzw. nicht-Belegung und darüber, ob sie von Messstrahlen traversiert wird. Das Erzeugen des Voxelraums auf einem Intel Core i7 (2.9 GHz CPU) Notebook mit 16 GB RAM hat für jede Messepoche ca. 5-6 Minuten in Anspruch genommen.

Während das Aufbauen der Voxelräume für jede Messepoche autark ablief, wurden zur Änderungsdetektion jeweils zwei Messepochen kombiniert analysiert und verarbeitet. Jedes Voxel einer Epoche wurde mit den entsprechenden Voxeln (und einem kleinen Nachbarschaftsbereich von 2 Voxeln) der anderen Epochen verglichen, um dabei entweder Bestätigung zu finden oder Widersprüche betreffend der vorliegenden Belegungsinformation aufzudecken und zu bewerten. *Bestätigt* belegte Voxel wurden grün gelabelt. Voxel, die in Epoche 1 leer waren, aber in Epoche 2 belegt sind, erscheinen gelb, d.h. etwas ist *neu*. Rote Voxel kennzeichnen solche, die noch in Epoche 1 vorhanden waren aber in Epoche 2 von Messstrahlen traversiert werden und somit offensichtlich leer sein müssen. Diese Voxel zeigen das Verschwinden bzw. den Abbruch von Elementen an. Voxel für die es nur in einer Messepoche, nicht aber in der zu vergleichenden Epoche Informationen gibt, verbleiben im Status *unbekannt* und erscheinen blau. Als Ergebnis der Änderungsdetektion wird eine klassifizierte Voxelpunktwolke ausgegeben (Abb. 7). Die Klassenzugehörigkeit kann entweder als zusätzliches Attribut (Label) oder im RGB-Wert hinterlegt werden.

Der Vergleich im Voxelraum und das Schreiben der Ergebnis-Voxelwolken haben für jede Messepoche ca. 50 Sekunden gedauert.

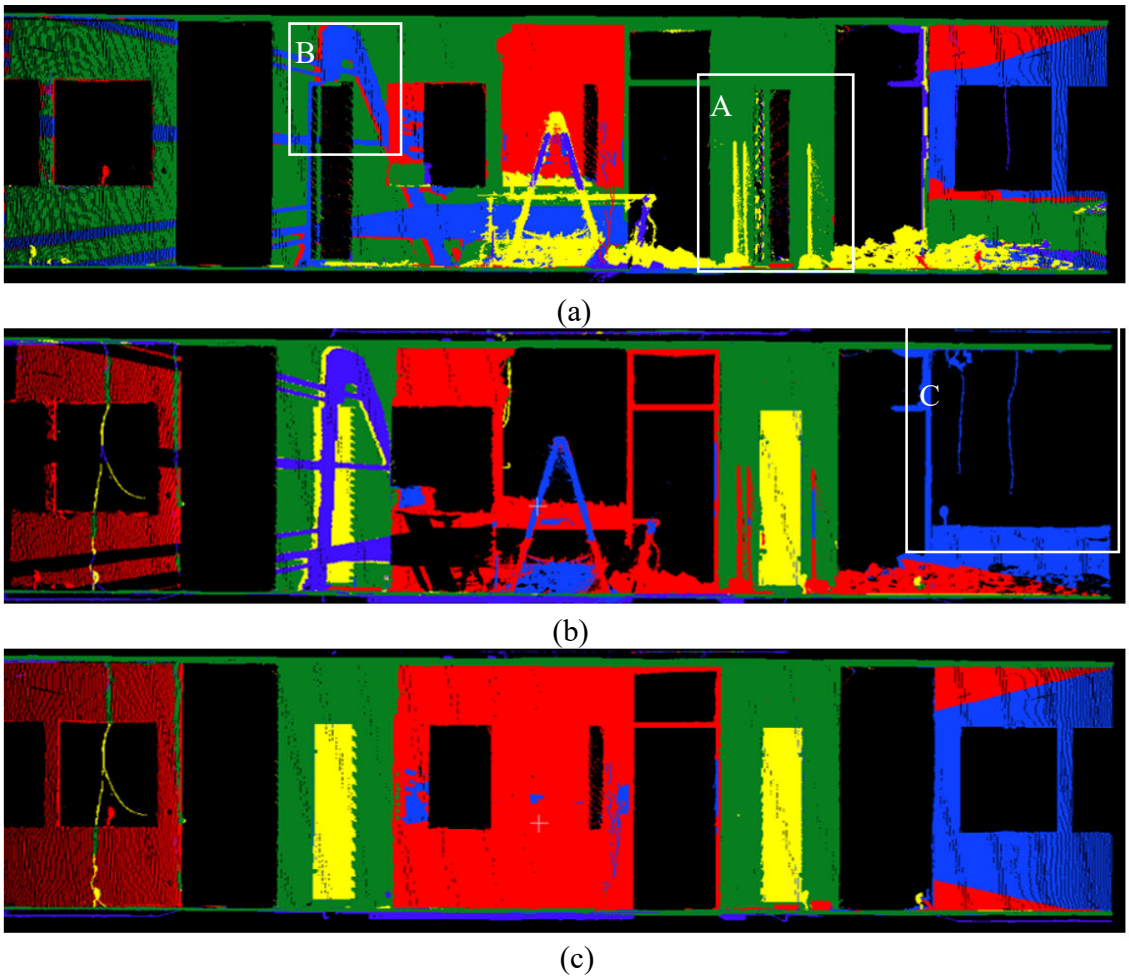
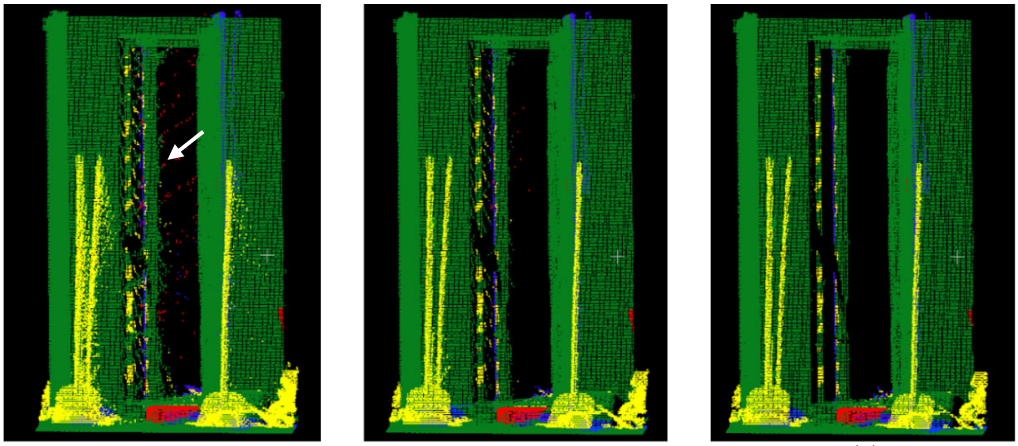


Abb. 7: Ergebnisse der Änderungsdetektion: a) Epoche 2 vs. Epoche 1, b) Epoche 3 vs. Epoche 2, c) Epoche 3 vs. Epoche 1

5.3 Einfluss von Schwellwert und Voxelgröße

Der Schwellwert (*th*) zur Binarisierung regelt die Relevanz von belegten und traversierten Voxeln im Hinblick auf die Gesamtbewertung von Änderungskonflikten. Um die Auswirkungen von Voxelgröße, Punktdichte und Punktrauschen innerhalb der Scans zu erfahren, d. h. Voxel mit unterschiedlich vielen zugehörigen Punkten, haben wir mehrere Durchläufe mit jeweils unterschiedlichem Schwellwert berechnet. Je niedriger *th* ist, desto mehr Voxel werden als relevant angesehen und desto mehr Informationen liefert das Ergebnis. Allerdings nehmen die Auswirkungen von Rauschen und Ausreißern mit abnehmendem *th* zu. Der Schwellwert wirkt sich besonders stark auf die Bereiche aus, die weit von den Scannerstandpunkten entfernt sind und daher nur eine geringe Punktdichte, aber gleichzeitig viel Rauschen aufweisen. Mit $th = 20$ haben wir einen guten Kompromiss und zufriedenstellende Ergebnisse für hohe Qualitätsansprüche erzielt (Abb. 8). Die numerischen Ergebnisse sind in Tab. 1 dargestellt.



(a) (b) (c)

Abb. 8: Veränderung des Schwellwertes *th* zur Reduktion von Punktrauschen für einen Detailausschnitt (vgl. Abb. 7a): a) *th*=3, b) *th*=10, c) *th*=20

Tab. 1: Ergebnisse der voxelbasierten Änderungsdetektion mit 1 cm Voxelgröße und verschiedenen Schwellwerten zur Anpassung der Relevanz von Einzelmessungen

Epoche	**bestätigt** [%]		**neu** [%]		**verschwunden** [%]		**unbekannt** [%]		gesamt [# Punkte]	
	th = 3	th = 10	th = 3	th = 10	th = 3	th = 10	th = 3	th = 10	th=3	th=10
2 vs. 1	53,2	52,2	7,9	12,6	4,3	7,8	34,7	27,3	1.712.632	706.100
3 vs. 2	49,73	42,9	1,47	2,3	8,4	15,62	40,39	39,18	2.327.731	809.194
3 vs. 1	43,64	45,1	1,12	2,3	7,9	16,65	47,37	35,94	2.266.865	803.064

5.4 Auswirkungen der Scan-Geometrie

Die Aufnahmegeometrie beim TLS hat großen Einfluss auf die Qualität einer nachfolgenden voxelbasierten Änderungsdetektion. Bei der Aufnahme vor Ort muss darauf geachtet werden, dass alle Bereiche in beiden Epochen ausreichend vom Scanner erfasst werden. Auch solche, in denen vermeintlich keine neue Information zu finden ist, wie z.B. im Bereich einer bereits abgebrochenen Wand. Es ist wichtig, dass auch leere Bereiche in einer nachfolgenden Messung trotzdem von Messstrahlen durchdrungen werden, damit der Voxelraum an diesen Stellen traversiert und damit auch tatsächlich als *leer* angenommen wird. Nur so können klare Aussagen über Bestätigung oder Widerspruch getroffen werden. Sonst kann es passieren, dass relativ große Anteile des Untersuchungsgebiets im Status *unbekannt* verbleiben (Abb. 7c). Vor diesem Hintergrund unterscheidet sich die Aufnahmeplanung für eine Änderungsdetektion von einer solchen zur reinen Neu-Bestandsmodellierung. Das Scanning vor Ort sollte in dem Bewusstsein durchgeführt werden, dass belegter und nicht-belegter Objektraum gleichwertig wichtig für nachfolgende Analysen sind.

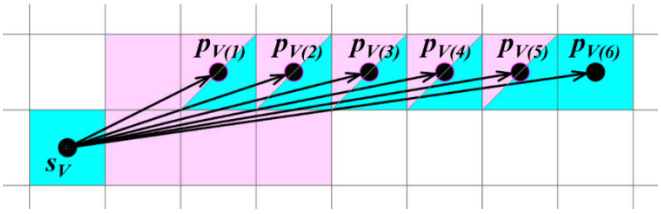


Abb. 9: Auswirkung von schleifenden Schnitten auf die Modellierung der Raumbelegung mit Voxeln

Die architektonische Beschaffenheit im Gebäudeinneren, insbesondere im (Um-) Baubetrieb, ist grundsätzlich eine herausfordernde Umgebung für terrestrisches Laserscanning. Eingeschränkte Bewegungsräume, stark variierende Objektabstände und i.d.R. viele Störobjekte wie Möbel und ggf. Baustellenmaterial verursachen geringe Überlappung von benachbarten Punktwolken, Lücken, Verdeckungen, heterogene Punktdichten und schleifende Schnitte, d.h. kleine Auftreffwinkel zwischen Messstrahl und Objektoberflächennormale. Letzteres tritt auf, wenn der Abstand zwischen Scanner und Objekt zu gering ist (z.B. in einem schmalen, langen Flur) und resultiert nicht nur in einer verminderten Messgenauigkeit, sondern auch fehlerhaft klassifizierten Voxeln zur Änderungsdetektion. Schleifende Schnitte haben zur Folge, dass belegte Voxel gleichzeitig auch traversiert werden und je nach Schwellwert, Punktdichte und Voxelgröße fälschlicherweise als leer klassifiziert werden (siehe Abb. 7b und Abb. 9). Um diesen Effekt zu vermeiden sollte auf eine ausreichende Anzahl von Standpunkten bei der Aufnahme geachtet werden und darauf, die Punktwolken vor der Weiterverarbeitung z.B. mit einem Distanzfilter zu bereinigen. Dieser Mehraufwand zahlt sich zugunsten einer automatisierten voxelbasierten Interpretation wieder aus.

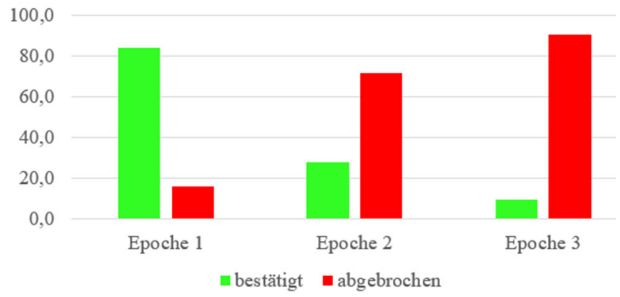


Abb. 10: Ergebnisdiagramm: Anzahl von Voxeln der Klassen *bestätigt* und *abgebrochen* [%]

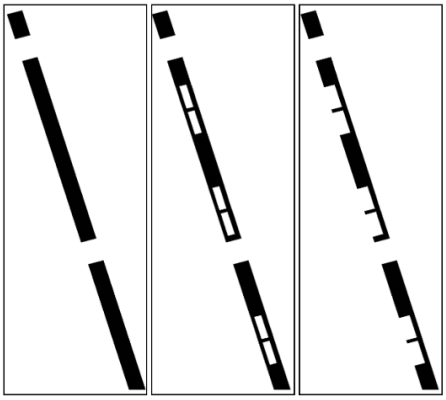
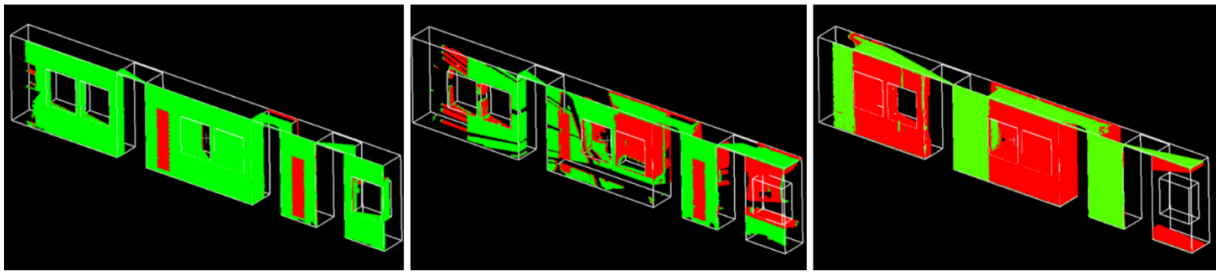


Abb. 11: Aus dem BIM abgeleitete Belegungsgitter für ein Wandstück. Schwarze Rasterzellen zeigen belegten Raum

5.5 Abgleich mit gegebenem BIM

Die Raumbelegung kann nicht nur aus Messdaten modelliert werden, sondern lässt sich auch aus einem gegebenen BIM ableiten. Sie entspricht dann der SOLL-Belegung und kann damit nicht nur zur Änderungsdetektion, sondern auch zur Verifizierung des Modells genutzt werden. Bei einer Voxelgröße von 1 cm wäre das bis einschließlich LOA 30 möglich. Das BIM wird dafür ebenfalls in den Voxelraum überführt und gemäß der Zielauflösung diskretisiert, sodass eine Reihe zweidimensionaler Belegungsgitter entsteht (Abb. 11). Ein BIM liegt üblicherweise im IFC-Format vor. Solche Geometrieobjekte können praktisch mit Autodesk Revit in ein CAD-Format konvertiert und weiterverarbeitet werden. Alternativ empfiehlt sich zur direkten Verarbeitung im Programmcode die open source Implementierung *IFC++* für C++.



(a) (b) (c)

Abb. 12: Ergebnis der Änderungsdetektion BIM (alter Zustand) vs. Punktwolke: a) BIM vs. Epoche 1, b) BIM vs. Epoche 2, c) BIM vs. Epoche 3

Jede Messepoche wurde gegen das BIM (Zustand vor den Abbrucharbeiten) verglichen. Die Ergebnisse spiegeln die fortschreitenden Abbrucharbeiten optisch und numerisch wider (Abb. 10 und Abb. 12). Während Epoche 1 einen hohen Anteil von über 80% bestätigt *belegten* Voxeln hat, kehrt sich das Bild bis zu Epoche 3 mit ca. 90% als *abgebrochen* gekennzeichneten Voxeln um. Der Vergleich von *Punktwolke vs. BIM* hat gegenüber dem Vergleich *Punktwolke vs. Punktwolke* den Vorteil, dass die Raumbelegung aus dem BIM volumetrisch bzw. flächig vollständig beschrieben wird, während die Raumbelegung aus zwei Punktwolken nur in deren Überlappungsbereich und auch nur anhand der Abtastung auf der Objektoberfläche meist unvollständig und lückenhaft vorliegt. Bei einem vollständigen Belegungsmodell ist u. a. auch das Innere eines Objektkörpers, z.B. einer Wand, beschrieben, sodass Messstrahlen, die den Raum traversieren den volumetrisch vollständigen Abbruch anzeigen. Diese Vollständigkeit ist besonders wertvoll, wenn klassifizierte Voxel für räumliche Abfragen und Operationen in einer räumlichen Gebäudedatenbank oder einem 3D-/ Indoor-GIS herangezogen werden (Abb. 13). Bei entsprechender „3D-Fähigkeit“ eines solchen Systems ist es möglich die Modellgeometrie durch 3D-Funktionen wie *3D Difference*, *3D Intersection* und *3D Merge* zu aktualisieren. Um kombinierte Analysen von Modell, Punktwolke und Voxeln auch mit Level of Detail Strukturen zu ermöglichen, bedarf es jedoch der Implementierung von 3D-Indizes und der Möglichkeit zur Abbildung sämtlicher topologischer Beziehungen in 3D, wie sie z.B. in BORRMAN et al. (2006) vorgeschlagen und in MEYER et al. (2019) diskutiert werden.

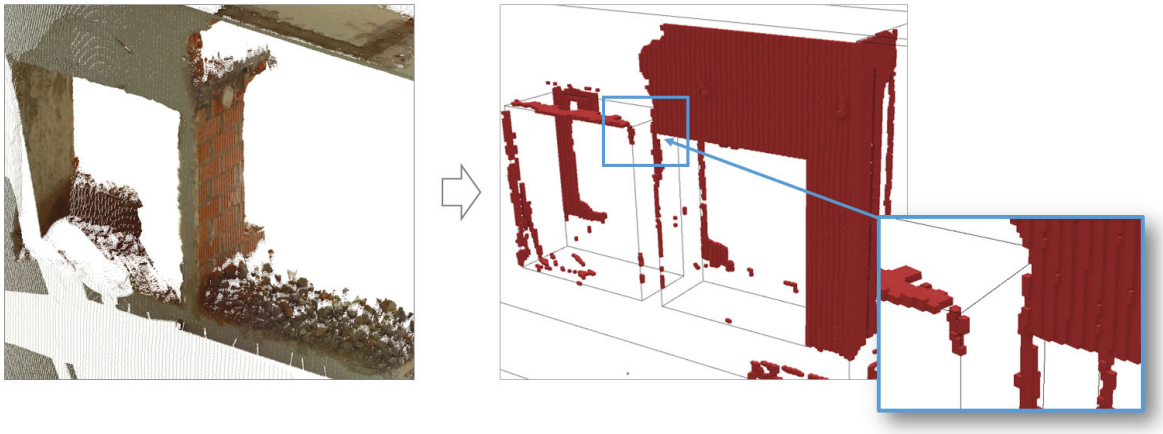


Abb. 13: Originale 3D-Punktwolke und 3D-Gebäudemodell mit "Abbruch-Voxeln", als Ergebnis der Änderungsdetektion

6 Fazit & Ausblick

3D-Punktwolken aus Laserscanning im Nahbereich bergen durch ihre hohe Auflösung ein immenses Informationspotenzial, das nur selten voll ausgeschöpft wird, weil es immer noch häufig an geeigneten anwendungsspezifischen und vor allem automatisierten Auswertestrategien mangelt. In diesem Beitrag haben wir eine Methode zur voxelbasierten Änderungsdetektion im Kontext von automatisierter Baufortschrittsdokumentation im Gebäudeinnenbereich vorgestellt und anhand eines realen Umbauprojekts demonstriert. Bei einer hohen Punktdichte, die einen hohen Schwellwert zur Evaluation der Voxel-Belegung und damit zur Gewährleistung ausreichender Qualität erlaubt, sind Voxelgrößen von nur einem Zentimeter möglich. Das entspricht einer sehr feinen Auflösung mit einem Diskretisierungsfehler, der unvermeidlich vorhanden, aber klein genug ist, um mittleren Genauigkeitsanforderungen gemäß LOA Spezifikation für BIM gerecht zu werden.
Die Ergebnisse haben gezeigt, dass multitemporale Innenraum-Messungen mit einem terrestrischen Laserscanner zu Zwecken der Änderungsdetektion insgesamt jedoch deutlich höheren Ansprüchen vor allem an Qualität, Aufnahmegeometrie, Registrierung und Punktdichte genügen müssen, als es für Messprojekte zur reinen Bestandsmodellierung der Fall ist. Insbesondere bei beengten Aufnahmesituationen (z.B. in langen, schmalen Fluren) muss darauf geachtet werden, dass die Scanauflösung der räumlichen Ziel-Auflösung zur Änderungsdetektion noch gerecht wird und dass der Objektraum auch in nicht-belegten Bereichen möglichst lückenlos erfasst wird. Dabei kann es auch erforderlich sein, Fehlereffekte wie sog. Mixed Pixel, die im Innenbereich häufig entlang von Kanten auftreten zu berücksichtigen, indem für jede Messepoche auf gleiche Messanordnungen bis hin zu identischen Scanstandpunkten geachtet wird, was wiederum eine Vermarkung der Standpunkte notwendig machen würde.
Zur kombinierten Verarbeitung von BIM und 3D-Punktwolken erwarten wir Fortschritte auf dem Gebiet der BIM-GIS-Integration, um klassifizierte Voxel mit erweiterter Funktionalität von 3D- und insbesondere Indoor-GIS zu analysieren. Die voxelbasierte Änderungsdetektion werden wir noch in weiteren Umbauprojekten im Gebäudeinneren verifizieren und erproben. Dabei wollen wir auch größere Datenmengen durch Parallelisierung und Integration der GPU als zusätzlichen

Prozessor beschleunigt verarbeiten. Um höchsten ingenieurgeodätischen Genauigkeitsansprüchen gerecht zu werden, werden wir außerdem das 3D-Voxelraster zum Aufdecken von Belegungskonflikten nutzen, wobei dann ein präziser Direktvergleich nach dem Ansatz von HEBEL et al. (2013) auf Einzelpunkt-Ebene mit dem Modell stattfinden soll.

Danksagungen

Wir danken Herrn Dr. Marcus Hebel von Fraunhofer IOSB für die Unterstützung beim Aufbau der Verarbeitungsroutinen.
Diese Arbeit wurde im Rahmen des Leonhard Obermeyer Center (LOC) an der Technischen Universität München (TUM) durchgeführt (https://www.ed.tum.de/loc/home/).

7 Literaturverzeichnis

BORRMANN, A., VAN TREECK, C. & RANK, E., 2006: Towards a 3D Spatial Query Language. Proc. Joint Int. Conf. of Computing and Decision Making in Civil and Building Engineering (ICCCBE-XI).

BRESENHAM, J. E., 1965: Algorithm for computer control of a digital plotter. IBM Syst. J., **4** (1), 25-30. https://10.1147/sj.41.0025.

GEHRUNG, J., HEBEL, M., ARENS, M. & STILLA, U., 2020: Change detection and deformation analysis based on mobile laser scanning data of urban areas. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. V-2-2020, 703-710. https://10.5194/isprs-annals-V-2-2020-703-2020.

HEBEL, M., ARENS, M. & STILLA, U., 2013: Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. ISPRS Journal of Photogrammetry and Remote Sensing, **86**, 52-64. https://10.1016/j.isprsjprs.2013.09.005.

HUANG, R., XU, Y., HOEGNER, L. & STILLA, U., 2021: Semantics-aided 3D change detection on construction sites using UAV-based photogrammetric point clouds. Automation in Construction, **134**, 104057, https://doi.org/10.1016/j.autcon.2021.104057.

MEYER, T. & BRUNN, A., 2019: 3D Point Clouds in PostgreSQL/PostGIS for Applications in GIS and Geodesy. Proceedings of the 5th International Conference on Geographical Information Systems Theory, Applications and Management, 154-163, https://10.5220/0007840901540163.

MEYER, T., BRUNN, A. & STILLA, U., 2021: Accuracy Investigation on Image-based Change Detection from BIM compliant Indoor Models. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-4-2021, 105-112. https://10.5194/isprs-annals-V-4-2021-105-2021.

TUTTAS, S., BRAUN, A., BORRMANN, A. & STILLA, U., 2017: Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4D BIM. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **85**(1), 3-15, https://doi.org/10.1007/s41064-016-0002-z.

USIBD, 2022: U.S. Institute of Building Documentation. https://usibd.org/, letzter Zugriff am 08.01.2022.

Automatisierte Einbindung und Verarbeitung von Copernicus-Daten in ein Digitales Lager geotopografischer Daten

MARKUS MUERTH[1], MARCO ZULKOWSKI[1], NINA KRÜGER[1], INES KOPPENHAGEN[1] & DANIEL HOLWEG[1]

Zusammenfassung: Das Ziel von Geodateninfrastrukturen (GDI) in Verwaltung und Wirtschaft ist meist die unterbrechungsfreie Bereitstellung und Verarbeitung von Geodaten unterschiedlichster Formate und Quellen. Satellitendaten des EU Copernicus Programm stellen u.a. durch ihre hohe zeitliche Auflösung eine besondere Ergänzung des Datenarchivs dar. Ein Digitales Lager geotopographischer Daten ermöglicht die vollautomatische und regelbasierte Einbindung, Konfektionierung, Datenhaltung, Bereitstellung und Verarbeitung vieler Arten von Geodaten. Neben einer konkreten Umsetzung des Konzepts als Software, zeigen wir zwei typische Anwendungsfälle für die Einbindung von Sentinel-2 Daten auf, die mit Hilfe des Digitalen Lagers wiederkehrend automatisch ausgeführt werden können.

1 Einleitung

Im Rahmen des europäischen Erdbeobachtungsprogramms Copernicus wurden eine Reihe von Möglichkeiten geschaffen, die Nutzern einen leichten und freien Zugang zu Sentinel Satellitendaten und weiteren Copernicus-Diensten über Datenportale (z.B. ESA Science Hub) und Cloud-Plattformen (z.B. CODE-DE, AWS, Creodias) ermöglichen. Besonders seit dem Start von Sentinel-2A im Jahr 2015 entwickelte sich ein großes Interesse an Copernicus in verschiedenen potentiellen Nutzerkreisen. Andererseits entwickelten sich bereits seit den 1990er Jahren, durch den Einzug von Geoinformationssystemen in Wirtschaft und Verwaltung, server-basierte Geodateninfrastrukturen (GDI) zur Datenhaltung und Verarbeitung geotopografischer Daten. Bei diesen werden neue Daten entweder durch die Nutzer oder über vordefinierte Schnittstellen eingebunden. Moderne GDI erlauben darüber hinaus auch die Darbietung der Produkte an Abnehmer, z.B. über Portale und Web Dienste wie WMS oder WFS.

Aufgrund der oft großen Datenmengen und zum Teil komplexen Datenstrukturen von gerasterten Geodaten, gilt es, jetzt bestehende lokale Dateninfrastrukturen (z.B. für geotopographische Daten) mit den neueren, cloudbasierten Infrastrukturen (z.B. für Satellitendaten) zu verbinden. Dies kann durch eine regelbasierte, voll automatisierte Einbindung von Satellitendaten aus der Cloud in eine GDI geschehen. Beispielsweise da große Mengen an Daten schon in der GDI vorgehalten werden bzw. dort entsprechende Rechenkapazitäten und Algorithmen bereitstehen.

Nach einer kurzen Übersicht über die Problemstellung und den Lösungsansatz werden die Vorteile der Einbindung von Satellitendaten in ein Digitales Lager für Geodaten herausgestellt. Schließlich beleuchtet der Beitrag die effiziente Einbindung von Satellitendaten in ein Digitales Lager anhand eines Beispiels. In diesem Beispiel erfolgt die automatisierte regelbasierte Auswahl, Aufbereitung und Bereitstellung von Sentinel-2 Daten für das Digitale Lager durch den Dienst *Sentinel Archiver*.

[1] M.O.S.S. Computer Grafik Systeme GmbH, Hohenbrunner Weg 13, D-82024 Taufkirchen, E-Mail: [mmuerth, dholweg]@moss.de

Das Modul *Sentinel Connector* als Teil der GDI-Komponente *novaFACTORY* registriert die bereitgestellten Daten in der GDI und koordiniert voreingestellte Verarbeitungsschritte. Schließlich werden die Satellitendaten im Digitalen Lager auf Basis der Software *novaFACTORY* konfektioniert und gelagert, sowie Anwendern und Veredelungs-Prozessen bereitgestellt. So können sich die Nutzer ganz auf die produktive Auswertung der Satellitendaten auch im Zusammenspiel mit anderen Fernerkundungsdaten wie z.B. Luftbildern und amtlichen, vektor-basierten Geodaten (wie z.B. ALKIS und ATKIS) fokussieren.

2 Regelbasierte Bereitstellung von Satellitendaten für geotopographische Produktions- und Vertriebsprozesse

2.1 Problemstellung

In verschiedenen öffentlichen und wirtschaftlichen Bereichen werden bereits seit Jahrzehnten Geodatenbestände aufgebaut und verarbeitet, die insbesondere seit dem Start des Copernicus-Programms nun auch die Möglichkeit bieten, unterschiedlichste frei verfügbare Satellitendaten und darauf aufbauende Produkte für ihre Anwendungen zu nutzen. Dazu gehören insbesondere optische Sensordaten und abgeleitete Produkte aus der Sentinel-2 Mission, die z.B. die Landnutzungskartierung (z.B. MALINOWSKI et al. 2020) oder das Monitoring von Wäldern (FASSNACHT et al. 2021) durch eine hohe Dichte der Zeitserie verbessern. Satellitendaten ermöglichen aber auch verschiedene Anwendungen in der Landschaftsplanung, z.B. das Ausloten von Potentialen für die Nutzung Erneuerbarer Energien im Projekt COP4EE (JONAS et al. 2019).

Da moderne Satellitensensoren eine große Menge an Daten bereitstellen, werden diese in zunehmendem Maße über Datenportale und Cloud-Plattformen auch zur Verarbeitung in der Cloud bereitgestellt. Jedoch stellt die wiederkehrende Verwendung von Satellitendaten in operativen Prozessen Anforderungen, die z.T. über den Möglichkeiten dieser Plattformen liegen (siehe u.a. KLIEN et al. 2019).

- Die **Massendatenverarbeitung** von Satellitendaten mit z.T. noch größeren Mengen an lokal vorliegenden Daten (z.B. Luftbildern)
- Der **Datenschutz** von sensitiven Eingangsdaten muss gewährleistet werden.
- Eine „**Datenfusion**“ von Sensoren mit ähnlichen Eigenschaften (z.B. Sentinel-2 und Landsat) oder von Daten aus unterschiedlichen Aufnahmetechniken (z.B. Sentinel-1 SAR, Sentinel-2 MSI oder Laserscanner Punktwolken) erfordert immer wiederkehrende Vorverarbeitung der Satellitendaten.
- Satellitendaten unterschiedlicher Anbieter werden über jeweils **unterschiedliche Repositorien** bzw. Plattformen bereitgestellt.
- **Kommerzielle Anwendungen** oder vom Nutzer selbst erstellte Algorithmen lassen sich oft nur aufwändig oder gar nicht in andere Infrastrukturen übertragen.
- Die automatische Ausführung von **datengetriebenen Workflows** für komplexe Arbeitsabläufe benötigt ein spezialisiertes Workflow Management System.

Jedoch erfüllen manche GDI Software Lösungen diese Anforderungen, z.B. solche die auf dem Konzept des Digitalen Lagers beruhen.

2.2 Lösungsansatz

Technische Basis der vorgestellten Lösung für eine operationelle Einbindung von Satellitendaten in eine GDI ist eine server-basierte Datenverwaltungs- und Prozessierungsumgebung, wie sie bereits vielfach in der Vermessungs- und Umweltverwaltung, sowie in der Planung, Verwaltung und Überwachung von raumwirksamen wirtschaftlichen Infrastrukturen eingesetzt wird. Zusammen mit Frontend-Anwendungen wie z.B. GIS Oberflächen für Desktop und Browser kann ein sogenanntes Digitales Lager für eine Vielzahl von geotopografischen Daten (3D, Punktwolken, Raster, Vektor), und auch räumlich verorteter Dokumente (Geodokumente) entstehen. Beispiele dafür sind Fotographien, Verträge oder amtliche Dokumente wie grafische Nachweise und Zahlennachweise der Vermessung.
Die Integration von Erdbeobachtungsdaten in solch eine GDI sollte durch die Anbindung eines Dienstes erfolgen, der in der Cloud-Umgebung, in der Satellitendaten zur Verfügung gestellt werden, installiert ist. Die Raum-Zeit Komponente und die Art der importierten Daten können so von der lokalen GDI über den Dienst in der Cloud gesteuert werden. Es entsteht bei den Nutzern der lokalen GDI die volle Kontrolle über die verwendeten Daten. Auch ermöglicht es ihnen einen regelbasierten, automatischen Import und die Weiterverarbeitung der Daten mit bekannten Werkzeugen.

2.3 Methodik

2.3.1 Das Digitale Lager

Mit dem Konzept des Digitalen Lager für Geodaten werden Grundgedanken der Lagerhaltung aus der industriellen Produktion auf Geodaten übertragen. Damit soll eine zuverlässige, effiziente und ressourcenschonende Umgebung für die Produktion und Bereitstellung von Geodatenbeständen geschaffen werden. Hierbei liegt der Schwerpunkt auf der bedarfsgerechten Bereitstellung von Daten und der darauf aufbauenden Prozesse, unabhängig davon, ob es sich dabei um interne Verarbeitungs- oder externe Bereitstellungsprozesse finaler Datenprodukte handelt. (siehe auch M.O.S.S., 2019)
Zu den wichtigsten Aufgaben des Digitalen Lagers gehört dabei die **Konfektionierung** der Geodaten über klar definierte Datenformate, Kachelungen, Metadaten und Lagerorte. Wie in einem physischen Lager werden alle Informationen über die Daten in einer zentralen Datenbank gespeichert. Zusätzliche können Datenflüsse und Verarbeitungsschritte so festgelegt werden, dass sie wiederholbar und automatisiert ablaufen können.
Dies ermöglicht eine automatisierte **Veredelung** der Daten, die über die reine, einstufige Satellitendatenverarbeitung (wie z.B. Indexberechnungen) hinausgeht. Dazu gehören:

- Verschneidung z.B. mit stark volumenbildenden Daten wie Luftbilder und Vektordaten, die bei den Nutzern von GDIs meist bereits lokal vorliegen,
- Anbindung von frei definierbaren oder vorhandenen Methoden durch Workflow Management (z.B. kaskadierende und mehrstufige Bearbeitungsverfahren),
- Kundenprozesse die auf Anwender-eigenen oder kommerziellen Algorithmen (z.B. Change Detection mit Künstlicher Intelligenz) beruhen, und die Ergebnisse auf Kunden-eigenen Endpunkten veröffentlichen.

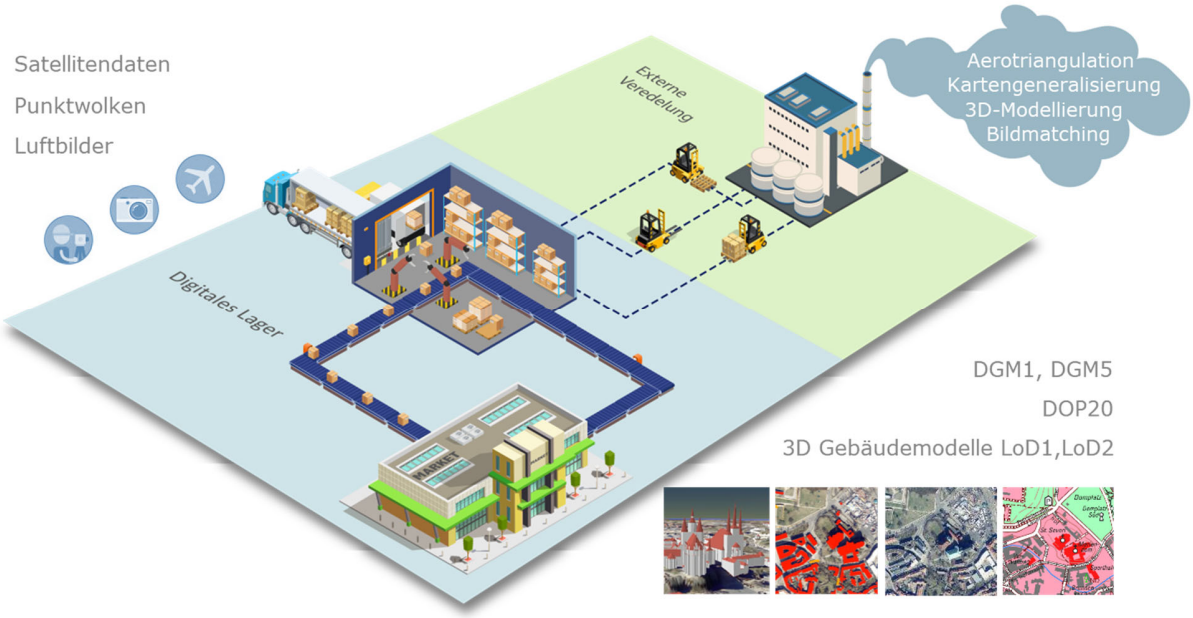


Abb. 1: Konzeptionelle Darstellung eines Digitalen Lagers

Die Vorteile einer automatisierten Veredelung können nur unter bestimmten Voraussetzungen für die Nutzer in Wert gesetzt werden. Dazu gehören:

- Automatische, **regelbasierte Einbindung** aller volumenbildenden bzw. zeitlich variablen Daten in das Digitale Lager
- Bereitstellung von Produkten oder Zwischenprodukten über **regelbasierte Dienste**, insbesondere bei zeitlich variablen Eingangsdaten
- Die Workflows zur Veredelung der Daten über größere Zeiträume oder Gebiete müssen in einer **skalierenden Rechenumgebung** ausführbar sein.

Eine moderne GDI, die das Konzept des Digitalen Lagers aufgreift, stellt genau diese Funktionalitäten bereit. So können z.B. aktuelle Themen in der Fernerkundung wie die automatische Fusion von Daten unterschiedlicher Satellitensensoren oder jede andere Art von Erdbeobachtungsdaten über eine GDI abgedeckt werden. Schließlich sollte ein Digitales Lager auch erlauben, diese Datenprodukte als „Pakete“ (Dateien), webbasierte Dienste (z.B. WMS etc.) und über Portale nach außen bereitzustellen.

2.3.2 Regelbasierter Zugriff auf Satellitendaten

Um Satellitendaten (beispielsweise von Copernicus) in einem Digitalen Lager bereitzustellen, muss die integrierte **Bereitstellungsfunktion** mehrere Bedingungen erfüllen:

- Die Satellitendaten müssen über voreingestellte oder dynamische Raum- und Zeitparameter automatisch ausgewählt werden.
- Idealerweise erlaubt die Schnittstelle zwischen Digitalem Lager und Datenquelle die Auswahl bestimmter Kanäle und vorverarbeiteter Satellitendatenprodukte eines Sensors.
- Die Daten müssen in einheitliche Datenformate übertragen werden und ihre Metadaten auf eine klar strukturierte Weise im Digitalen Lager zur Verfügung gestellt werden.

- Wenn nötig, muss eine Vorverarbeitung der Daten (z.B. Transformation oder zeitliche Aggregation) automatisiert ohne Interaktion mit dem Nutzer durchgeführt werden.

Werden diese Voraussetzungen erfüllt, kann eine Vielzahl von Satellitendaten in automatisierte Workflows eingebunden werden. Dazu können auch Verarbeitungsabläufe gehören, die bei einem bestimmten Auslöser auf das eingebundene Digitale Lager zugreifen, um in kürzester Zeit beispielsweise Kartierungen von Hochwasser- oder Sturmschäden zu erzeugen.
Die von der GDI gesteuerte Auswahl und Aufbereitung einzelner Satellitendaten und ihrer Metadaten sollte jedoch direkt in der Umgebung ausgeführt werden in der diese archiviert sind. Dazu bietet sich an, einen Dienst in der Cloud zu installieren, da dieser dann performant auf das gesamte Satellitendaten-Archiv (z.B. in einer Cloud Plattform) zugreifen kann.

2.4 Beispielhafte Umsetzung des Digitalen Lagers

Eine Möglichkeit der Umsetzung des Konzepts Digitales Lager für geotopographische Daten bietet die modulare und erweiterbare GDI-Komponente *novaFACTORY* (M.O.S.S. 2019). *novaFACTORY* erlaubt die Einbindung, Verwaltung und Bereitstellung einer Vielzahl von Geodaten unterschiedlichster Formate und Quellen und wird hierfür bereits bei einer großen Anzahl an öffentlichen Institutionen und Betrieben genutzt. Zudem ist eine weitere Hauptanwendung von *novaFACTORY* die wiederholbare, vollautomatische Steuerung von Workflows für den Import, den Export und die Verarbeitung von Geodaten.
Die Satellitendaten für die regelbasierte Einbindung in *novaFACTORY* stellt hierbei der cloudbasierte Dienst *Sentinel Archiver* bereit, der die automatisierte regelbasierte Auswahl, Aufbereitung und Bereitstellung von Sentinel-2 Daten für das Digitale Lager durchführt. Er stellt die nötigen Metadaten bereit und bereinigt nicht mehr benötigte oder aktualisierte Daten automatisch. Anwender profitieren von der hohen Frequenz der Datenaktualisierung ohne regelmäßig Kapazitäten dafür zu binden.

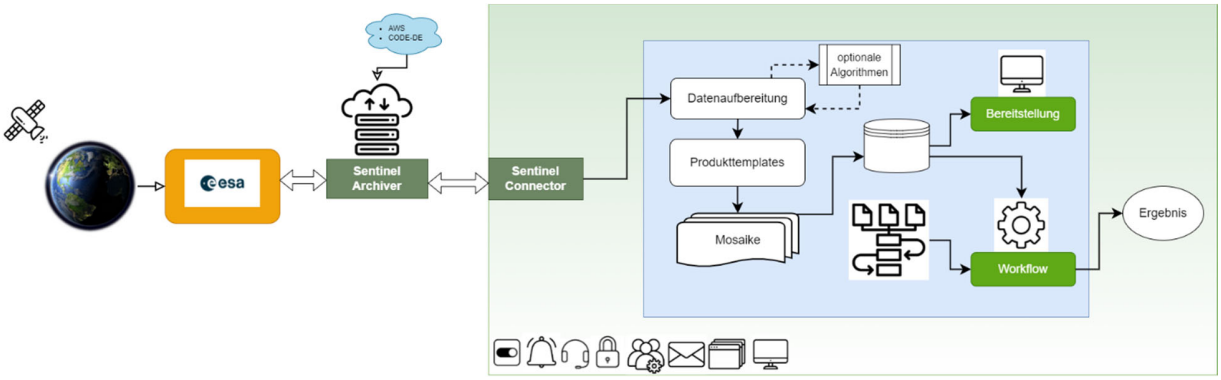


Abb. 2: Workflow zur Einbindung von Sentinel-2 Daten in das Digitale Lager

Das Modul *Sentinel Connector* als Teil der GDI-Komponente *novaFACTORY* registriert schließlich die bereitgestellten Sentinel-2 Daten in der GDI, koordiniert über sogenannte Produkttemplates voreingestellte Verarbeitungsschritte (z.B. die Generierung von wolkenfreien Mosaiken und Index Karten) und benachrichtigt die Nutzer bei Bereitstellung dieser Produkte.

Sind die vorverarbeiteten Sentinel-2 Daten schließlich im Digitalen Lager eingebunden, können sie den eigentlichen Nutzern über Download-Portale und Kartendienste (z.B. WMS) zur weiteren Verwendung dargeboten werden. Zudem stehen sie dann regelbasierten Veredelungsprozessen als Input zur Verfügung, so z.B. Change Detection (CD) Algorithmen (Abb. 3). Im Thüringer Landesamt für Bodenmanagement und Geoinformation (TLBG) wird die Aktualisierung des ATKIS Basis-DLM (Digitales Landschaftsmodell) nach ADV, 2019 bereits operationell mit Hilfe der *novaFACTORY* Datenhaltungs- und Workflow-Funktionalitäten durchgeführt. Der vom Anwender definierte Workflow benötigt sowohl Sentinel-2 als auch Digitale Orthophotos (DOPs) und ALKIS/ATKIS Daten die zeitgleich in *novaFACTORY* vorgehalten und automatisch regelbasiert bereitgestellt werden. Solch ein Workflow wird in der GDI skalierbar verarbeitet und das Ergebnis wieder im Digitalen Lager zur weiteren Verwendung verfügbar gemacht.

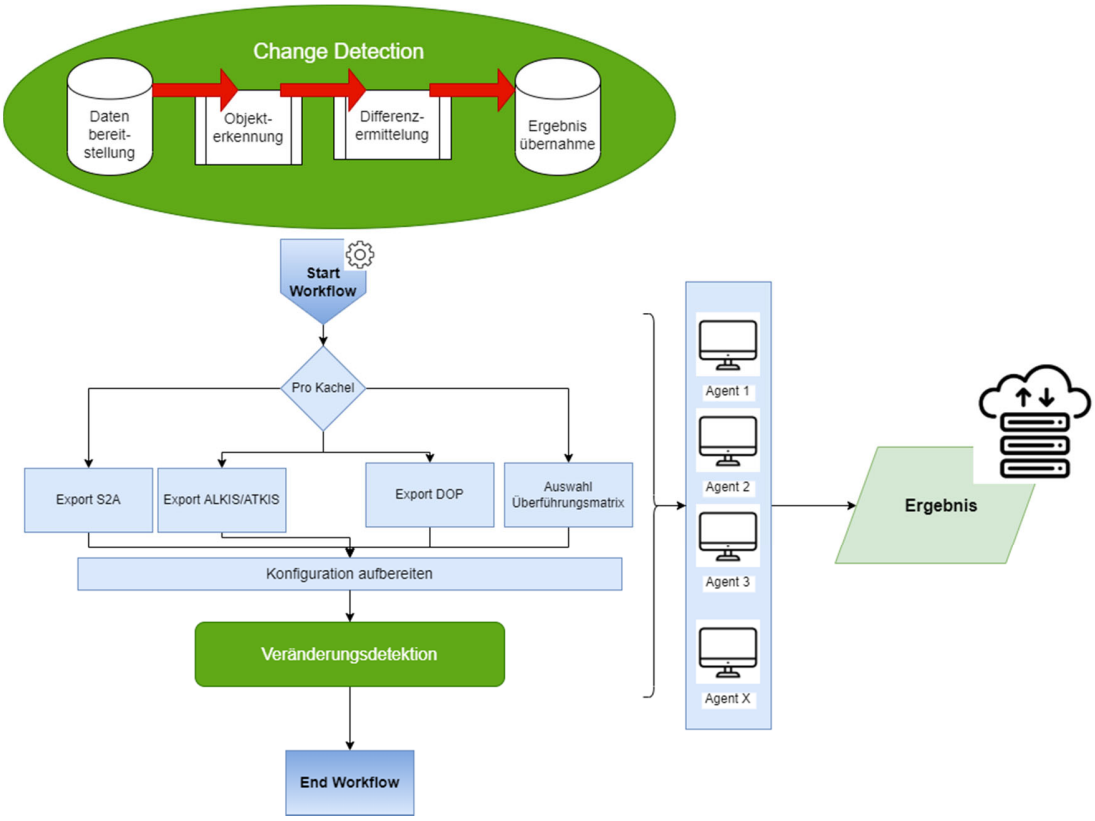


Abb. 3: Beispielhafter Change Detection Workflow

3 Fazit & Ausblick

Die Verarbeitung von Copernicus-Daten in der Cloud stellt für viele nutzerspezifische Fragestellungen immer noch eine Hürde in technischer Hinsicht dar. Beispielsweise benötigt die Erstellung von Waldschadenskarten oder die Change Detection mit Satelliten- und Luftbildern sowie vektorbasierten Karten häufig noch eine große Menge an lokal vorliegenden Daten und/oder basiert auf spezifischen Auswertungsmethoden. Daher ist die Einbindung von Satellitenprodukten (auch aus

unterschiedlichen Quellen) in eine lokale GDI bei operationellen Verarbeitungs- und Bereitstellungsprozessen für komplexe Anwendungen im Vorteil.
Der hier vorgestellte Ansatz zielt deshalb auf die Verknüpfung von cloudbasierten Copernicus Diensten mit einer modernen Lösungsarchitektur für die digitale Lagerhaltung geotopografischer Daten. Ein weiterer Vorteil der automatischen Einbindung der Satellitendaten in eine lokale GDI ist, dass den Nutzern bereits bekannte Anwendungsoberflächen, Weiterverarbeitungsschritte und (auch kommerzielle) Analysetools zur Verfügung stehen. Der Zeit- und damit Kostenaufwand reduziert sich so erheblich gegenüber einer Übertragung anderer Geodaten in die Cloud und der Neugenerierung und Bedienung von komplexen, datengetriebenen Workflows in Cloud Plattformen.
Darüber hinaus ist ein möglicher zusätzlicher Vorteil der Einbindung von Satellitendaten in eine GDI der einheitliche Zugriff auf Daten verschiedener Sensoren. Diese können durch leicht zu automatisierende Vorverarbeitungsschritte in der GDI in einheitlichen Geometrien, Zeitabschnitten sowie einer bestimmten Kanalauswahl bereitgestellt werden. Das erlaubt dann, diese Daten ohne detaillierte Kenntnis ihrer Formate zu höherwertigen Informationsprodukten zu verarbeiten. Diese Form der sogenannten Datenfusion ist ohne entsprechende Infrastruktur nur zeitaufwändig durch Erstellung von manuellen Vorverarbeitungsschritten möglich.

4 Literaturverzeichnis

AdV (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland), 2019: Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). Gesamtkonzept vom 01.06.2019.

Fassnacht, E.F., Poblete-Olivares, J., Rivero, L., Lopatin, L., Ceballos-Comisso, A. & Galleguillos, M., 2021: Using Sentinel-2 and canopy height models to derive a landscape-level biomass map covering multiple vegetation types. International Journal of Applied Earth Observation and Geoinformation, **94**, 102236, https://doi.org/10.1016/j.jag.2020.102236.

Jonas, F., Wiedemann, W., Fiedler, C., Jetter, F., Schmid, T., Holweg, D., Klöber, A., Richter, S., Christ, I., Lessing, R. & Walther, C., 2019: COP4EE - Entwicklung von Methoden und Diensten auf Basis von Satellitenbilddaten im Kontext von Copernicus zur Unterstützung des Einsatzes von erneuerbaren Energien. Abschlussbericht, DELPHI InformationsMusterManagement GmbH.

Klien, E. M., Krämer, M., Holweg, D. & Willkomm, P., 2019: Cloud4EO - Schnelle und effiziente Verarbeitung von Erdbeobachtungsdaten. Abschlussbericht, Fraunhofer Institut für Geographische Datenverarbeitung.

Malinowski, R., Lewiński, S., Rybicki, M., Gromny, E., Jenerowicz, M., Krupiński, M., Nowakowski, A., Wojtkowski, C., Krupiński, M., Krätzschmar, E., Schauer, P., 2020: Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. Remote Sensing, **12**(21), 3523, https://doi.org/10.3390/rs12213523.

M.O.S.S., 2019: Umsetzung eines Digitalen Lagers für Geodaten mit novaFACTORY. M.O.S.S. Computer Grafik Systeme GmbH, https://t1p.de/nFdigitaleslager20, letzter Zugriff am 11.01.2022.

Processing of Allied Air Forces World War II Strategic Air Reconnaissance Imagery

RUWEI NIE[1], MICHAEL CRAMER[1] & PATRICK LUBIG[2]

***Abstract**: Historical Aerial Images (HAIs) can provide long time series record of landscape change and landcover dynamic monitoring. With the development of computer vision, the challenges of interpretation and analysis of HAIs can be solved using methods like Structure from Motion (SfM). The key step of the application is to georeference the imagery to the reference coordinates system with a set of Ground Control Points (GCP). A coarse-to-fine approach was proposed to determine the position of GCPs and to generate georeferenced orthoimage and Digital Surface Model (DSM) in the study area. The georeferencing error was assessed by means of Root Mean Square Error (RMSE) computation. In total 4 steps of adding GCPs in weak areas with large RMSE values were implemented, to improve the georeferencing accuracy with a minimal number of GCPs expressed in terms of nGCPs/km^2. The refined accuracy for X, Y and Z axis were 1.63 m, 2.49 m and 7.59 m, respectively, with a ratio of 1.80 GCPs/km^2.*

1 Introduction

1.1 Motivation

In the years from 1939 to 1945, a transformational growth in aerial reconnaissance happened, especially in Britain and then in the United States. An amount of millions of historical aerial images was obtained by aerial reconnaissance in the time from 1938 until the end of the war and even beyond (MEIXNER & ECKSTEIN 2016). These Historical Aerial Images (HAIs) covered an aera of the majority of today's European countries and are now available in the allied archives such as the National Collection of Aerial Photography (NCAP) or the National Archives and Records Administration Washington (NARA). Recently, with the digitization of archival film-based historical aerial reconnaissance imagery, there is now free access through spatial data infrastructures and web services in many countries (GIORDANO et al. 2017). With the access of HAIs, a lot of research and imagery analysis were conducted by experts for the land survey and bomb detection. For example, the Luftbilddatenbank Dr. Carls GmbH did an extensive research and acquisition of stereoscopic interpretable series of aerial photographs from the time between 1938 and 1955 (MEIXNER & ECKSTEIN 2016). The Landesamt für Geoinformation und Landentwicklung in Baden-Württemberg has a collection of HAIs from the pre- and post-war period to produce true orthophotos and 3D models for landcover analysis and research. Like other historical records, the collection and interpretation of aerial reconnaissance intelligence are now a crucial contribution to ecosystem mapping and landscape features representation in small scale (TUOMINEN & PEKKARINEN 2005). Besides, they also provide long time series record of landscape

[1] Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, 70174 Stuttgart, E-Mail: michelle_11_15@outlook.com, michael.cramer@ifp.uni-stuttgart.de
[2] LBA Luftbildauswertung GmbH, Ludwigstraße 17B, 70176 Stuttgart, E-Mail: patrick.lubig@lba-luftbildauswertung.de

change and temporally continuous and spatially complete landcover dynamic monitoring (PAINE & KISER 2003).

However, along with the contributes and advantages of HAIs there are very specific characteristics of the imagery: analog photographs are acquired with different cameras in different acquisition conditions, the image quality is of huge variation and the images exhibit highly heterogeneous spatial resolutions (GIORDANO et al. 2018); the digital images are often scanned from old analog photographs, there remains only a coarse image localization. Metadata such as the physical size of the analog photos, the focal length of the camera and the camera calibration reports are not always available (MICHELETTI et al. 2015). Fortunately, with the rapid development of computer vision and the automatic feature-matching algorithms, Structure from Motion (SfM) resolves the problem of recovering the 3D structure from a series of overlapping aerial images. The approach uses conventional photogrammetric processing combined with computer vision SfM techniques to produce a robust solution that can self-calibrate and overcome the challenges of unknown uncalibrated cameras, inaccurate GPS systems and environmental conditions (EISENBEIß 2009). Hence, this method provides a possibility to the usage of the HAIs which are lack of ground control information and precise camera parameters to reconstruct the former geometry of landscape surfaces (NAGARAIAN & SCHENK 2016).

The paper is based on the research of a master thesis project in corporation of the Institut für Photogrammetrie (ifp) and LBA Luftbildauswertung GmbH. The practical application of the work in LBA requires to georeference the HAIs from World War II to present-day reference orthoimages or reference maps in order to search for the remains of arial bombing activities and to analyze the possibility of the unexploded ordnance on the land or underneath the earth surface. Hence, it is extremely important to locate the accurate position of the bombing area which is shown in the HAIs related to the present-day reference maps. As a result of this, the horizontal accuracy of the georeferencing process is of great interest for the practical work and is the most essential part for the applications. As known from the presented research, the georeferencing of HAIs is very difficult to implement and the accuracy can be very different because of the variance of the quality of the images and the condition of the study areas. One commonly applied method is to select Ground Control Points (GCPs) as ground reference manually by visual comparison between the scanned HAIs and the present-day orthoimages or Digital Surface Model (DSM) in GIS environment. Researches were conducted using this method: ONIGA et al. (2020) selected 150 GCPs in a 0.08 km^2 area and obtained a mean accuracy of 2.10 m with 1875 GCP/km^2; PERSIA et al. (2020) applied 75 GCPs in 67 HAIs covering 550 km^2 area surface and the mean accuracy was 2.66 m with 0.14 GCP/km^2. The goal of the master project was to obtain satisfying horizontal accuracy of georeferencing the HAIs at a lower cost of manual work as for a minimal number of GCPs, and to search for possible methods to improve the vertical accuracy at most. In the study area, the desired threshold for horizontal error was set to be below 5 m, while the expected vertical error was set to be below 10 m.

1.2 System Workflow

The system was designed subsequently according to the goals and requirements of the project. A coarse to fine approach was provided to generate fine absolute oriented orthoimage product. In the

coarse to fine approach, the first step was to compute a coarse absolute image orientation without the ground reference. The second step was to use coarse oriented orthoimage comparing with the present-day reference orthoimages and DSM to determine the GCPs, then with the ground reference the coarse absolute orientation was georeferenced to the reference coordinate system as fine absolute oriented orthoimage. The major challenge was the determination of ground references information, hence the georeferencing step was extremely essential to the accuracy of final absolute orientation. For the georeferencing process, a two-step method was proposed to achieve the desired accuracy with optimal GCPs. The Check Points (ChPs) provided a possibility to check and evaluate the georeferencing quality. At the end of the coarse to fine approach, the value of the ChPs coordinates error was assessed by the index Root Mean Square Error (RMSE). If the index RMSE was beyond the desired threshold of horizontal and vertical accuracy, the number of GCPs would then be increased to rerun the process until the desired accuracy is obtained.

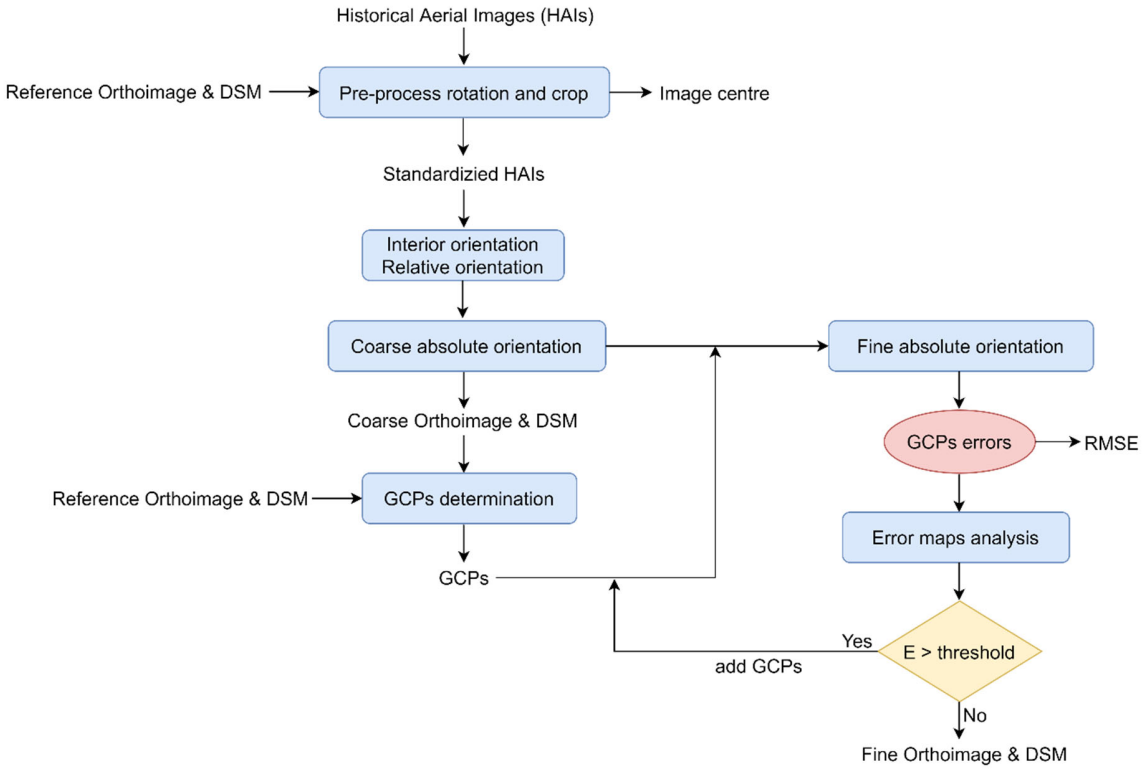


Fig. 1: System workflow

The designed workflow illustrated in Fig. 1 includes three main parts: the image pre-processing, the data processing, and the result analysis. In the image pre-processing part, the original HAIs were standardized in MATLAB, and the image centers were obtained. In the data processing part, the proposed coarse to fine approach was conducted with the GCPs determination in Agisoft Metashape. In the result analysis part, the GCPs coordinates error and the ChPs coordinates error were assessed to be RMSE and the error values were analyzed and visualized in error map productions in QGIS. The number of GCPs would be increased and added in the areas with high error values indicated by the error map to improve the georeferencing accuracy.

2 Data Sources and Study Area

2.1 Study Area

The study area as shown in Fig. 2 on the left side is in the capital city Erfurt in state Thüringen of Germany. In total 21 overlapping aerial images were processed covering a surface of approximately 20 km^2. The area was selected because the data sources were free available from Thüringen-Geoportal Open Geodata. What is more important is that the medieval town center of the city Erfurt with its many cultural monuments was not severely devastated by area bombings in the second World War, which made it easier to locate the ground reference for georeferencing. Reconstructing the aerial sightseeing from above with the HAIs from the end of the war and comparing the reconstructed scene with the reference orthoimages provided a method to discover the land cover change and city rebuilding after the second World War. Besides, the determination of the position of the area bombings is essential for the safety and stability of human activities and construction development. A second more rural study area with different landcover and identical features was applied to verify the general effectivity of the proposed approach but will not be discussed in detail (NIE 2021).

A sample of 21 black and white aerial images from two flight missions token above the city Erfurt were acquired as the source aerial images. As shown in Fig. 2 on the right side, the flight project 194539 contains 6 overlapping aerial images (green) which were captured on 23.03.1945; the flight project 194574 contains 15 overlapping aerial images (red) which were captured on 08.04.1945. Both flight projects were back and forth in the approximate direction of east-west. The 21 images were divided into three image groups, where there was not enough overlapping area in-between.

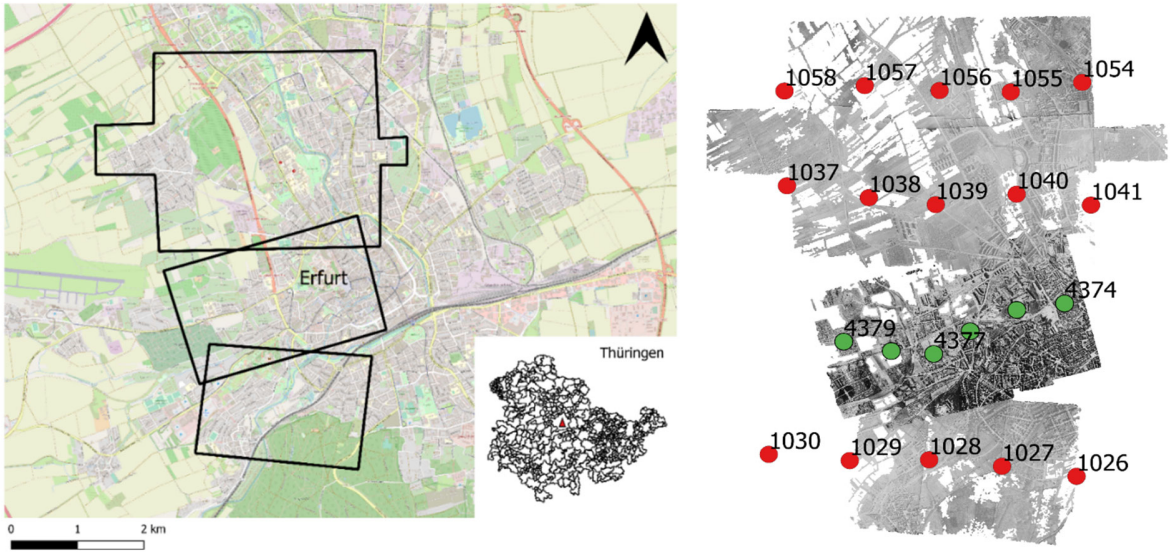


Fig. 2: Study area in Erfurt (left) and perspective center positions of two flight missions (right)

2.2 The Historical Aerial Images

Shown in Fig. 3 on the left side are two examples of the aerial images from flight missions 194539 and 194574. The zoom-in detailed identical features of the same scene from two different flight missions are shown on the right side. It is noteworthy that the two images are of different physical

size and image brightness. The image quality is not optimal to identify the features of streets and building structures. Additionally, the digital aerial images were scanned from analog aerial images and provided in GeoTIFF format. Hence, there was no calibration report for the camera information available, the manufacture information and interior orientation parameters were missing as well.

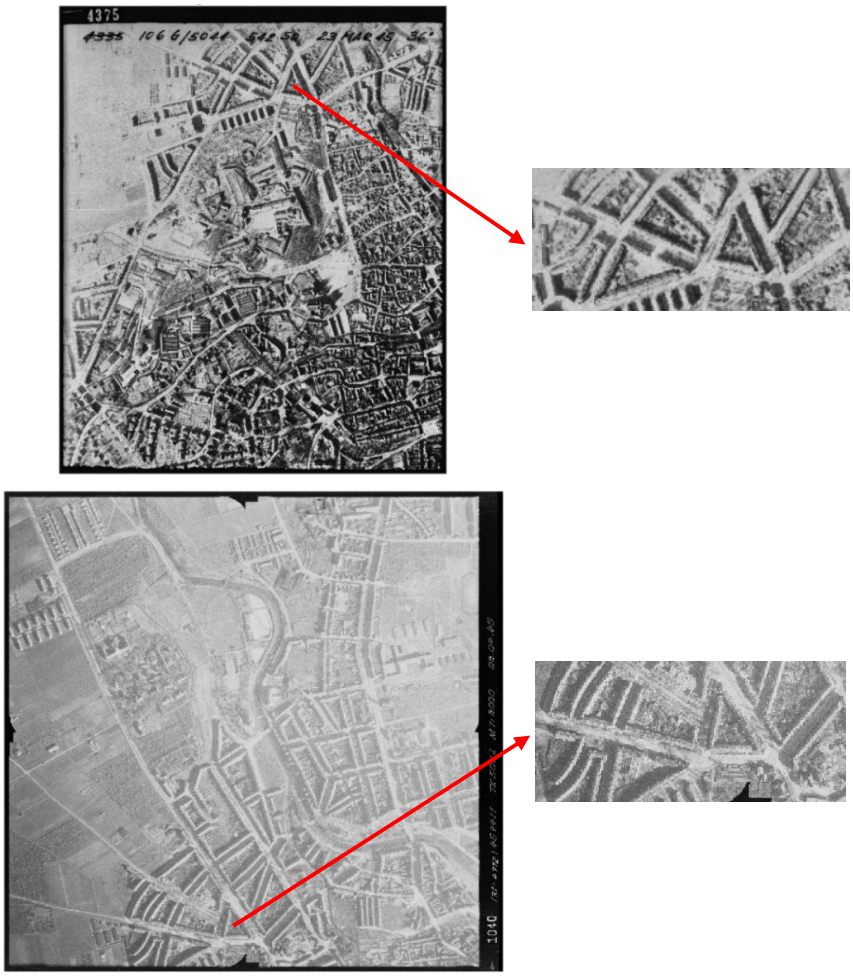


Fig. 3: Historical aerial images (left) with identical features (right) (top: 194539, bottom: 194574)

2.3 The Reference Orthoimage and DSM

The reference orthoimage and DSM shown in Fig. 4 were acquired from Thüringen-Geoportal Open Geodata in ETRS89 / UTM 32N coordinates system. The reference orthoimage colored in RGB were acquired on 23.03.2020 with a ground sampling distance of 0.2 m and standard deviation of 0.4 m. The State Office for Surveying and Geoinformation of Thüringen regularly used the airborne laser scanning to produce digital terrain and surface models, the DSM used as reference height model was obtained in January 2014 with a ground sampling distance of 1 m. The reference orthoimages and DSM rasters with the size of 1 km * 1 km were merged into one orthoimage and reference DSM covering the whole study area. The terrain in the north-east area is lower than in the south-west area, and the height difference is relatively large up to 170 m.

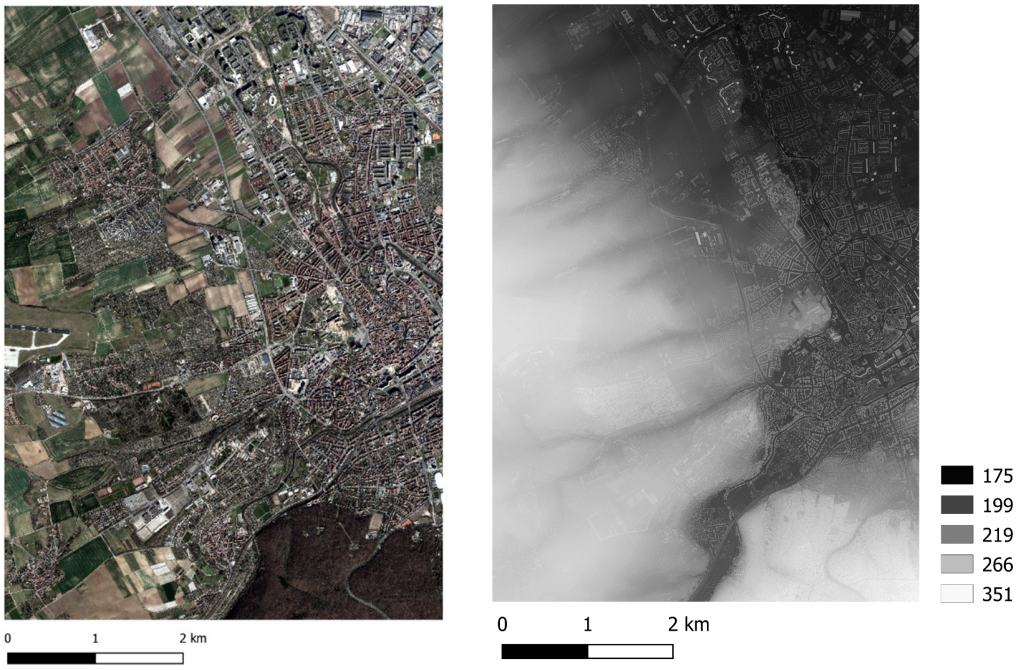


Fig. 4: Reference orthoimage (left) and DSM (right)

2.4 Ground Control Points

The GCPs are important to scale and orient the final orthoimage to the reference coordinates system and to assess the georeferencing accuracy of the ChPs. The horizontal and vertical positions of the selected control points were uncertain because they were identified manually using the present-day reference orthoimage and reference DSM. The land cover and the terrain surface might have changed significantly after a long time of manual constructions and human activities. To determine the uncertainty of the control points positions is extremely difficult. In order to minimize the uncertainty of the spatial positions of the control points, the selection of the positions where the GCPs and ChPs locate is essential. In this project, the control points were selected as the identities which were assumed to be stable and considered unchanged over time. The example features of GCPs and ChPs selection are shown in Fig. 5.

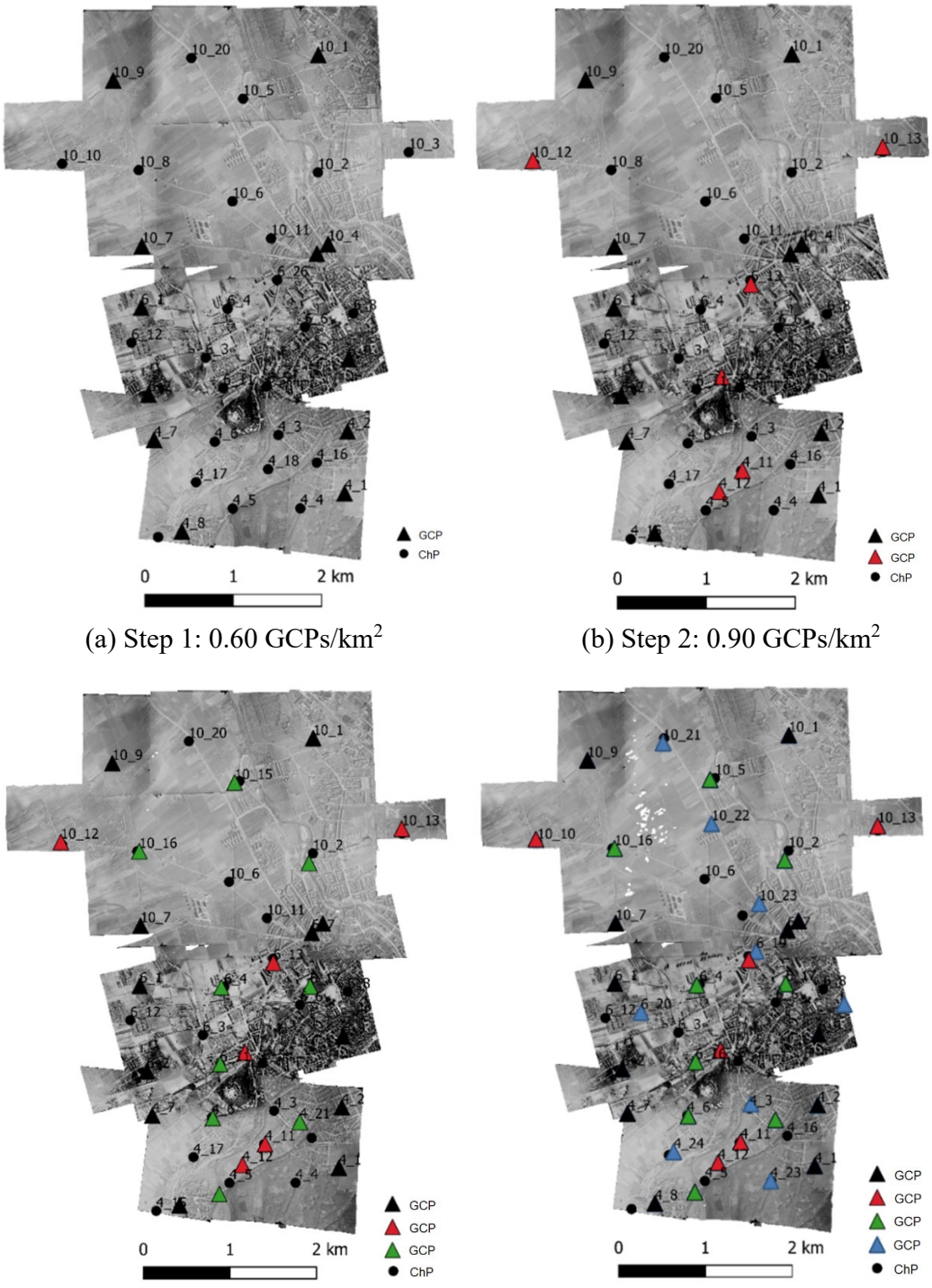
The study area is covering a large area of the city center and a small part of forests and farming fields. In the urban area, the GCPs and ChPs were mainly selected on the roof edges and some manmade features. For these identical features, the horizontal coordinates which were computed from reference orthoimage are more accurate than the vertical coordinates which were computed from reference DSM. A small change in the selected horizontal position on the roof edges might lead to a big difference in vertical height. Hence, they are usually used as 2D horizontal GCPs. The intersections of roads were used as reference points as well and are usually accurate in vertical and used as vertical GCPs. In the farming fields, the intersections of farming quarter were considered to be stable features. The stable features were selected as GCPs and ChPs manually by visual comparison of the coarse orthoimages of the three image groups and the reference orthoimage in QGIS. The X and Y coordinates were computed from the reference orthoimage

while the Z coordinates were extrapolated from the reference DSM. The reference coordinates system of the reference points was established in ETRS89 / UTM 32N.

Fig. 5: Reference points selection features (left: HAIs, right: reference orthoimage)

3 Implementation and Processing

In order to bring the scale and orientation of the coarse orthoimages to the reference coordinates system, the GCPs and ChPs were introduced to the georeferencing process. As shown in Fig. 6, in total 4 steps of adding GCPs to search for an optimal number of GCPs to improve the georeferencing results were implemented. The GCPs and ChPs were selected on the surfaces assumed to be stable over time and were distributed in the whole study area. In step 1, 12 GCPs in black triangles were selected in the corner of the image groups. More GCPs shown in red triangles, green triangles and blue triangles were added in each step, which were in total 18 GCPs in step 2, 27 GCPs in step 3 and 36 GCPs in step 4. The total surface of the study area is approximately 20 km^2. Hence, the index nGCP/km^2 in each step was 0.60 GCP/km^2, 0.90 GCP/km^2, 1.35 GCP/km^2 and 1.80 GCP/km^2. The number and position of the 24 ChPs remained unchanged, which provided check and evaluation to the georeferencing quality. The GCPs and ChPs in each step were added to the process in Agisoft Metashape and tagged to the HAIs manually. Besides, the georeferenced orthoimage and DSM of each step were generated as final georeferenced products.



(a) Step 1: 0.60 GCPs/km^2 (b) Step 2: 0.90 GCPs/km^2

(c) Step 3: 1.35 GCPs/km^2 (d) Step 4: 1.80 GCPs/km^2

Fig. 6: GCPs and ChPs in 4 steps

4 Results Analysis

4.1 RMSE of GCPs and ChPs

Once the aerial images were aligned with the selected GCPs and ChPs, the RMSE values referring to the corresponding GCPs and the ChPs in each step were calculated and exported, the labels of each selected GCPs and ChPS were listed along with the X, Y, Z coordinates, with X axis pointing towards east and Y axis pointing towards north. The error values of each GCPs and ChPs along the X, Y, Z axes were listed as X error, Y error and Z error, respectively (Tab. 1).

Tab. 1: RMSE (m) values referring to the corresponding GCPs and the ChPs in each step

Steps	Num.	X error	Y error	Z error	XY error	total error
Step 1	12 GCPs	2.12	2.29	0.52	3.12	3.16
	24 ChPs	4.23	5.45	42.27	6.90	42.83
Step 2	18 GCPs	2.36	2.25	0.38	3.26	3.28
	24 ChPs	2.74	4.18	28.04	4.99	28.48
Step 3	27 GCPs	1.86	1.98	0.83	2.72	2.84
	24 ChPs	2.83	3.57	12.53	4.56	13.33
Step 4	36 GCPs	1.99	1.86	0.67	2.73	2.81
	24 ChPs	2.05	3.48	18.65	4.04	19.08
Refine Step	36 GCPs	1.95	1.86	0.67	2.69	2.77
	24 ChPs	1.63	2.49	7.59	2.98	8.15

The RMSE values referring to all the 24 ChPs along X, Y, Z axes in each step are illustrated in Fig. 7. The horizontal axis consists of the 4 steps of adding GCPs for georeferencing and refine step with new ChPs selections. The left vertical axis in blue is the scale for X and Y errors, while the right vertical axis in red is the scale for Z errors. X error, Y error and Z error are represented in square, triangle and circle, the values of the errors along three directions in each step are marked clearly as well.

Pay attention to the RMSE values comparison of the ChPs, it is clear that from step 1 to step 2, the errors along the X and Y directions are decreasing slightly while the error along the Z direction is decreasing dramatically from 42.27 m to 28.04 m. The horizontal accuracy is satisfied in step 2 but the vertical error still overcomes the desired vertical accuracy. Hence more GCPs were added in step 3 to reduce the vertical error. From step 2 to step 3, the X error is slightly increasing while the Y error is slightly decreasing, which might be explained by the flight line direction. As mentioned, the flight direction was east-west direction which is along the X axis of easting. The Y axis is perpendicular to the X axis pointing northing. The Z error continues to reduce from 28.04 m to 12.53 m, which is not acceptable for vertical accuracy. In order to verify the prediction that the vertical error continues decreasing when more GCPs were added, step 4 was applied and proved that the Z error did not continue decreasing as predicted but increased from 12.53 m to 18.65 m. It means that there is no necessity to invest more manual work of selecting GCPs for

georeferencing for more steps. Hence the system design stopped in step 4 and the vertical accuracy in step 3 is considered the best georeferencing results we might obtain by adding GCPs.
In step 3, the X error and Y error are 2.83 m and 3.57 m, which are absolutely acceptable for horizontal accuracy considering the study area is 20 km^2 and only 27 GCPs were applied. The Z error with 12.53 m is, however, relatively large for vertical accuracy. The possible reasons might be the GCPs selection procedure. Since adding GCPs in step 4 did not improve the Z accuracy, other investments were applied to improve the accuracy in vertical direction in refine step (see section 4.3).

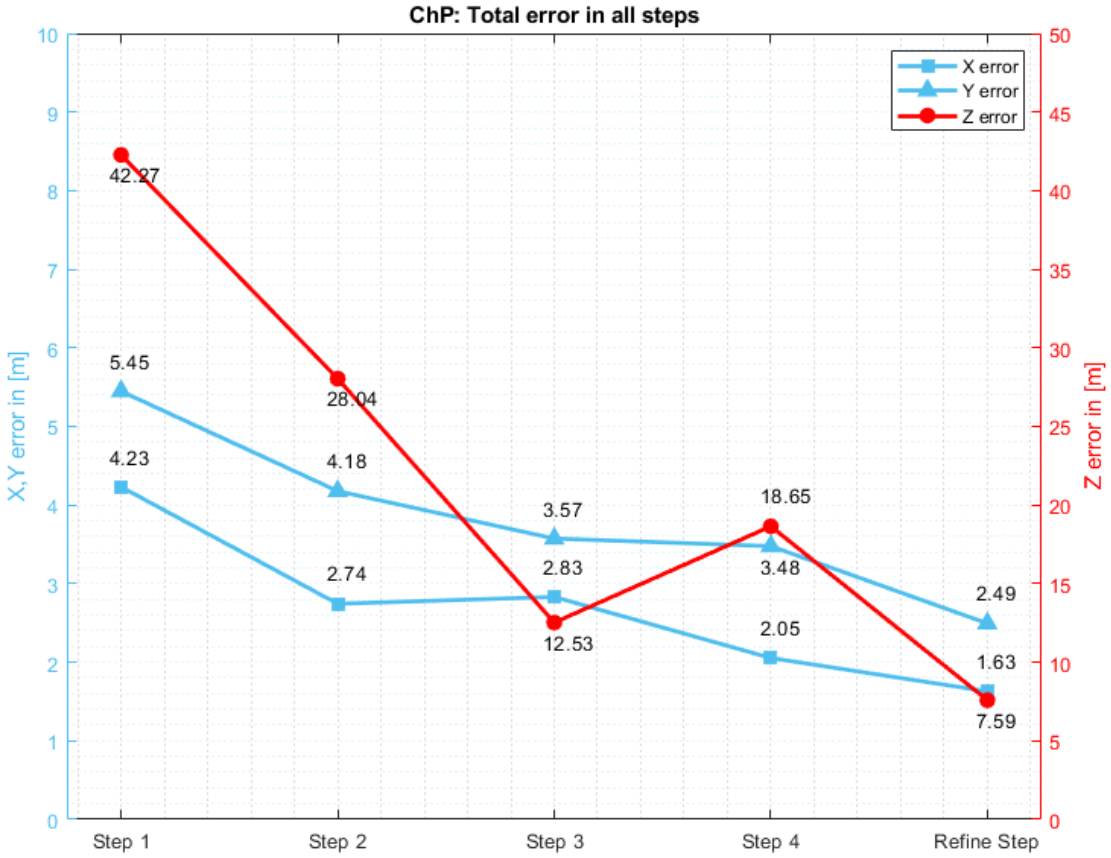


Fig. 7: RMSE values comparison of ChPs in all steps

4.2 Error Maps

To understand the spatial distributions of the RMSE values of the GCPs and the ChPs, the error maps of the GCP errors and ChP errors along X, Y, Z axes in step 3 are illustrated in Fig. 8. Error maps of the GCPs and ChPs along X axis, Y axis and Z axis are shown in red, green and blue with the triangles representing the GCPs errors and the circle points representing the ChPs errors. The color of each point represents the value of the errors. The value of the error changes from minimum to maximum as the color of the points changes from light to dark: the darker the point, the larger the error value. More GCPs were added in the areas with large absolute ChPs error values which were shown in the error maps in each step to improve the georeferencing accuracy.

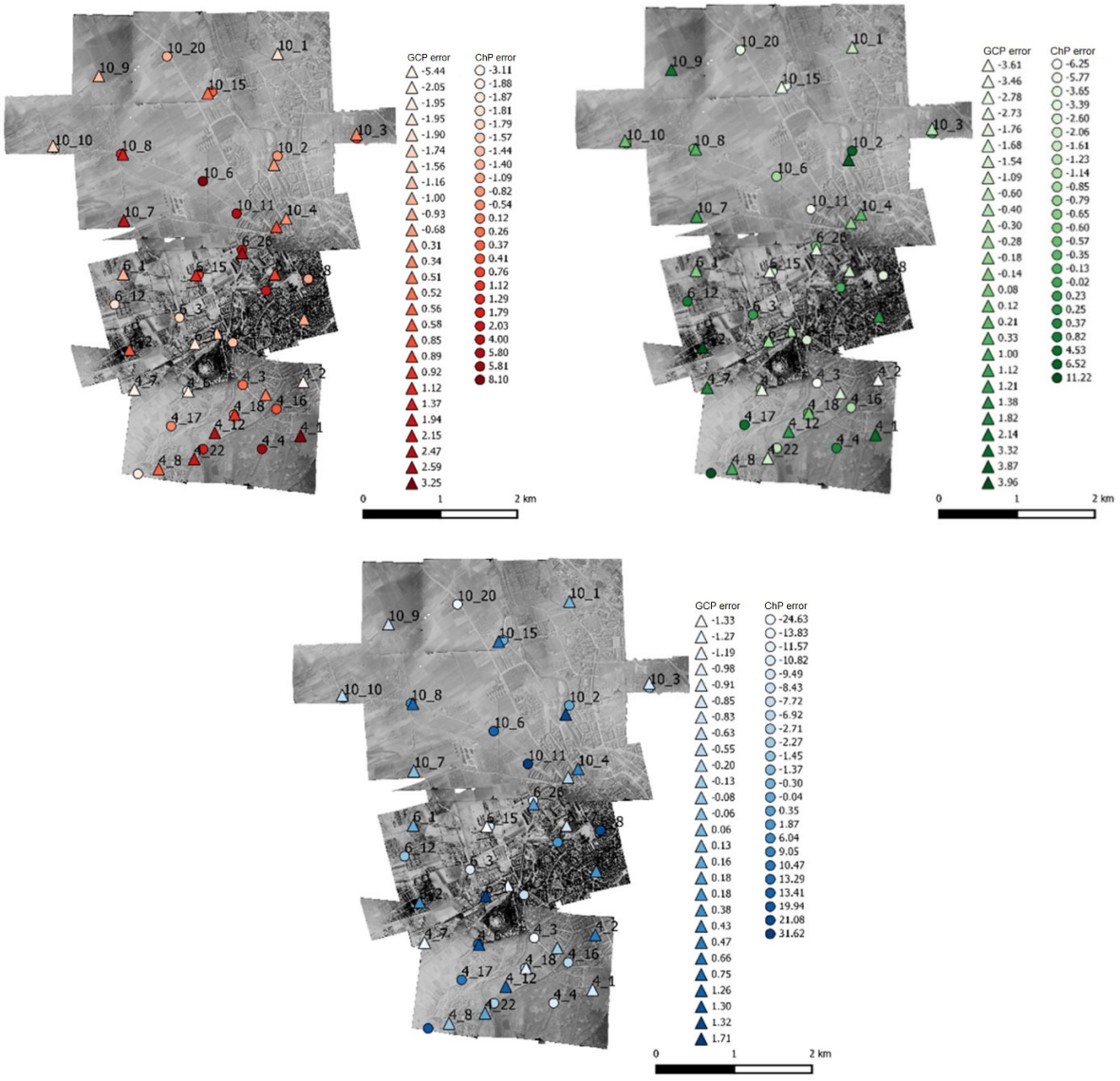


Fig. 8: Error maps along X (red), Y (green), Z (blue) axis in step 3

4.3 Refine Results

The study area in the city Erfurt was mainly covered by buildings and a small part of forests and farming fields. Hence, the horizontal error values were small and could be improved by adding GCPs easily because it was not difficult to find the identical features to determine the GCPs and the positions of the GCPs were relatively accurate. On the contrary, the vertical error values were the most uncertain element and could not be improved by adding GCPs in step 4. The possible error source might be that the GCPs and ChPs were selected mostly on the identical features such as the roof edges (Fig. 9). It might be the fact that the points were located on the road with height of 194 m besides the roof with height of 213 m or on the tree leaves above the roof, which could make a difference of 10 m to 15 m in vertical.

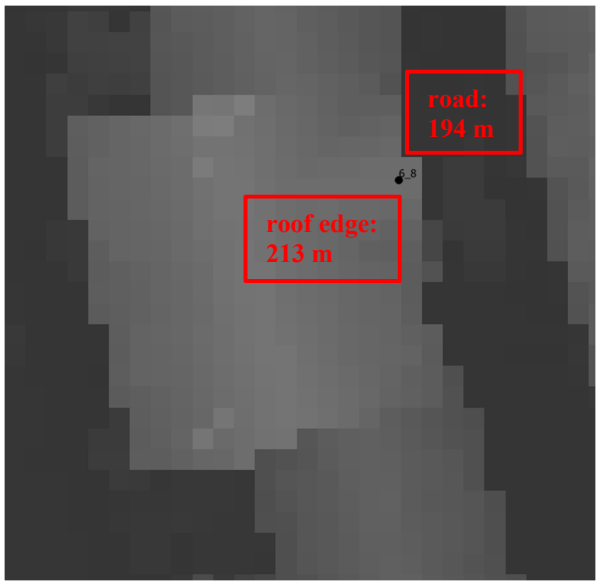


Fig. 9: Error source of ChPs selection

The vertical accuracy of the georeferencing process is of most uncertainty and adding GCPs in the areas with large errors could not improve the vertical accuracy in both study areas after 3 steps. Other methods were applied to determine the possible influence on vertical georeferencing elements, such as adding manual tie points to the matching images, separating the location GCPs and height GCPs. Unfortunately, all these methods did not work on improving the vertical accuracy. The method of the project was to increase the number of GCPs but did not change the position and quantity of the ChPs in the 4 steps. However, as explained before, the GCPs and ChPs determination procedure might have influenced the vertical accuracy and caused a difference of 10 m to 15 m in vertical direction. Hence, the ChPs with large RMSE values in Z axis were considered as outliers and were deleted in order to refine the vertical accuracy. After deleting the outliers, new ChPs were selected as shown in Fig. 10. The ChP 6_8 at the roof edge was with a high Z error and was deleted as an outlier and a new ChP 6_8 was selected in a new position.
Based on the error map along Z axis in step 4, in total 4 ChPs with large Z error values were newly selected. The RMSE values referring to the 36 GCPs and 24 ChPs with new selections are shown as refine step in Tab. 1 and from Fig. 7 we can see that the Z error reduced from 18.65 m in step 4 to 7.59 m in refine step. Additionally, it is noteworthy that the horizontal error values decreased as well with the new selections of ChPs. From the result that both the horizontal accuracy and vertical accuracy increased and the minimum and the maximum Z error values reduced dramatically, there is indeed a problem with GCPs and ChPs selection procedures, which caused the error values in Z axis. Hence, the advice for future works is to select the points correctly in the position where there is no great height difference between the selected point and the surroundings. Besides, the result indicates that the accuracy might be continuously improved when the ChPs with large error values are new selected.

Fig. 10: ChPs outlier (left) and new selection (right)

5 Conclusion and Future Work

The goal was to georeference the HAIs at a lower cost of manual work as for a minimal number of GCPs to obtain the desired horizontal accuracy, and to search for the possible methods to improve the vertical accuracy. The project has provided a workflow to process the images in a coarse to fine approach, in which the aerial images were processed without any camera calibration information as for coarse processed orthoimages, and then the coarse orthoimages were georeferenced with the selected GCPs and the ChPs to produce the error maps of X, Y, Z error values, according to which more GCPs were selected and added in the areas affected by large errors to achieve a higher accuracy for horizontal and vertical. Finally, when the accuracy could not be improved by adding more GCPs in the weak areas with large errors, the final fine georeferenced orthoimages were generated and mapped.

Two study areas with different landcover and identical features were tested with the workflow proposed in the project. In the study area in Erfurt, the best accuracy obtained by adding GCPs were 2.83 m for X axis, 3.57 m for Y axis, 12.53 m for Z axis, respectively, with a minimum quantity of only 27 GCPs in the area of 20 km^2. In the more rural study area, the best accuracy obtained by adding GCPs were 11.23 m for X axis, 6.47 m for Y axis, 12.80 m for Z axis, respectively. The horizontal error values were large compared with the horizontal error values of the study area in Erfurt but considered to be acceptable with a minimal number of only 12 GCPs in an area of 16.6 km^2.

In both study areas, the vertical accuracy stopped to improve after 3 steps of adding more GCPs in the weak areas. This fact proves that it is not the case that the more GCPs the better accuracy but there is an optimal number of GCPs with a lower cost of manual work within a certain area. The index nGCPs/km^2 which means the GCPs rate per unit area, is an effective index to improve the efficiency of georeferencing process: in the urban area with obvious identical features such as building roof and streets intersection, the ratio is 1.35 GCPs/km^2; in the area with vegetation landcover and farming field, the ratio is 0.72 GCPs/km^2. The index nGCPs/km^2 is now utilized as

a reference for practical application of the work in LBA, to reduce the manual investigation to select GCPs in different landcover and surface conditions and to georeference the HAIs with a desired horizontal accuracy better than 5 m.
Additionally, the vertical accuracy in both study areas was approximately 12 m in step 3. The selection procedure of GCPs and ChPs caused a difference of 10 m to 15 m in vertical direction, which could not be improved by adding GCPs to georeferencing process. After new selection of ChPs in refine step, the vertical accuracy was improved to be 7.59 m. The research of refining the vertical accuracy indicated how important the determination procedure of GCPs and ChPs was. Therefore, a method of how and where to locate the position of GCPs for the practical application was suggested: to select identical features which are not direct near trees or other vegetations and are not at the exact corner of the roof edges near streets or roads.
Many practical applications require accurate horizontal georeferencing of old images and present-day maps referring to the present regions. From the final result analysis, it is apparent that the selection of the GCPs and ChPs is the fundamental procedure in the georeferencing process of HAIs. The final refined accuracy of the proposed methodology satisfied the expectation: the RMSE values for X, Y and Z axis were 1.63 m, 2.49 m and 7.59 m, respectively, with a ratio of 1.80 GCPs/km^2. The workflow provides a method to georeference the HAIs efficiently with the index nGCPs/km^2 and automatically with the computer vision SfM techniques.
However, it is worth mentioning that the method of adding GCPs in weak areas with large errors was not investigated in this work. The GCPs were added in or within the areas where the RMSE values were large but without a certain rule and the distribution of the GCPs was not regulated as well. Theoretically, to define the optimal number of GCPs in a certain size of study area statistically is of significant research potential. In addition, to determine how the spatial distribution of the GCPs as well as the ChPs affects the final georeferencing accuracy is also an interesting topic to follow on. Another topic which might deserve further research lies in different study areas with different conditions, such as areas with great land cover change, which could make it difficult to find GCPs, or areas with large terrain change, which makes it even more challenging to determine error values in the vertical direction.

6 References

EISENBEIß, H., 2009: UAV photogrammetry. Doctoral dissertation, ETH Zurich, https://doi.org/10.3929/ethz-a-005939264.

GIORDANO, S., LE BRIS, A. & MALLET, C., 2017: Fully automatic analysis of archival aerial images current status and challenges. IEEE Joint Urban Remote Sensing Event (JURSE), 1-4, https://doi.org/10.1109/JURSE.2017.7924620.

GIORDANO, S., LE BRIS, A. & MALLET, C., 2018: Toward automatic georeferencing of archival aerial photogrammetric surveys. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **4**, 105-112. https://hal.archives-ouvertes.fr/hal-02386709.

MICHELETTI, N., LANE, S. N. & CHANDLER, J. H., 2015: Application of archival aerial photogrammetry to quantify climate forcing of alpine landscapes. The Photogrammetric Record, **30**(150), 143-165. https://doi.org/10.1111/phor.12099.

MEIXNER, P. & ECKSTEIN, M., 2016: MULTI-TEMPORAL ANALYSIS OF WWII RECONNAISSANCE PHOTOS. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, **41**. https://doi.org/10.5194/isprs-archives-XLI-B8-973-2016.

NAGARAJAN, S. & SCHENK, T., 2016: Feature-based registration of historical aerial images by Area Minimization. ISPRS Journal of Photogrammetry and Remote Sensing, **116**, 15-23. https://doi.org/10.1016/j.isprsjprs.2016.02.012.

NIE, R., 2021: Processing of Allied Air Forces World War II Strategic Air Reconnaissance Imagery. Master dissertation, Universität Stuttgart, Institut für Photogrammetrie, https://www.ifp.uni-stuttgart.de/lehre/masterarbeiten/625-Nie.

ONIGA, V. E., BREABAN, A. I., PFEIFER, N. & CHIRILA, C., 2020: Determining the suitable number of ground control points for UAS images georeferencing by varying number and spatial distribution. Remote Sensing, **12**(5), 876. https://doi.org/10.3390/rs12050876.

PAINE, D. P. & KISER, J. D., 2012: Aerial photography and image interpretation. John Wiley & Sons.

PERSIA, M., BARCA, E., GRECO, R., MARZULLI, M. I. & TARTARINO, P., 2020: Archival aerial images georeferencing: a geostatistically-based approach for improving orthophoto accuracy with minimal number of ground control points. Remote Sensing, **12**(14), 2232. https://doi.org/10.3390/rs12142232.

TUOMINEN, S. & PEKKARINEN, A., 2005: Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote sensing of Environment, **94**(2), 256-268. https://doi.org/10.1016/j.rse.2004.10.001.

Implementation of a low-cost Ambulance Management System

FRANZ OKYERE[1], THEA MINNICH[2], MORITZ SPROLL[2], EDMUND MENSAH[3], LILIAN AMARTEY[1], CARLOS OTOO-KWOFIE[3] & ANSGAR BRUNN[2]

Abstract: Ambulance services are a critical component of health service delivery. The importance of a functioning and efficient management of ambulances to achieve the most gain of highly expensive resources cannot be overemphasised. Resources needed in health provision are expensive and commercially available management systems have high costs in installation and maintenance. In most deprived regions, the available ambulances to convey patients may not be smartly deployed. This is due to the lack of an effective fleet management system. Furthermore, poor road conditions only add insult to injury. Mobile health equipment takes advantage of the road network to make resource allocation and management optimal. We seek to develop a system that will reduce the cost of implementation of an ambulance management system. In this article, we present the implementation of a low-cost effective ambulance management system (EAMS) for Ashaiman, a large town in the Greater Accra Region of South Ghana and is the capital of the Ashaiman Municipal District. The solution is based on the principles of a geographic information system (GIS) and gives the operator clear acting paths. As part of planning, constant consultation with stakeholders was carried out making sure the basic user needs assessments were established. OpenStreetMap (OSM), a crowdsourced volunteered geographic information (VGI), data served as the source of spatial data for the system; thus the spatial database was linked to an open-source GIS desktop application. The OSM is used as well as cost-free Google satellite and Google hybrid as background map. Bespoke functionalities created within the custom application realise necessary processes; the functions were integrated into the software via custom python plugins. Moreover, communication of the requests for service from users to the ambulances and vice versa is done by email and messenger and Internet services. The operating system used was also free software with zero cost. We made use of software that can be installed at will on any computer void of the compatibility issues common with different computer platforms. The system has been installed in a clinic in Ashaiman (Ghana). The system has been tested by live simulation.

1 Introduction

In Africa, a higher proportion of individuals generally lack access to health facilities and emergency treatment. After an emergency, mortality is linked to income and a country's or region's overall degree of development (EGGER et al. 2018: 490). Meanwhile, Ghana has one of the worst emergency supplies in the world. In the entire country. There are only 56 ambulances in use, some

[1] KAAF University College, Kasoa, Ghana,
E-Mail: Franz.Okyere@kaafuni.edu.gh, LilianAmartey31@gmail.com
[2] University for Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany,
E-Mail: Thea.Minnich@student.fhws.de, Moritz.Sproll@googlemail.com, Ansgar.Brunn@fhws.de
[3] Cape Coast Technical University, Cape Coast, Ghana,
E-Mail: Carlos Otoo-Kwofie: noraspinto@yahoo.co.uk, Edmund Mensah: eddiebeezy5@gmail.com

of which are even not used at all due to technical difficulties. The few ambulances concentrate on the regional capitals (HUPP et al. 2019: 133). As a result, circumstances for rescuing individuals in medical situations are inadequate in Ashaiman, a city in Ghana's Greater Accra area (Accra metropolitan). These circumstances impact many, who live within the Ashaiman township.
Effective resource allocation and use will require a minimum of resources. But rather we find there are frequently insufficient cars available to evacuate sick people in critical conditions (HUPP et al. 2019: 132). It is vital to address these difficulties and improve the situation, especially of disadvantaged groups such as people living in informal settlements. The goal of this study is to see how well QGIS (previously Quantum GIS), an open-source geographic information system (GIS), can manage and help coordinate a fleet of ambulances. Additionally, we need to ensure efficiency in communication between the ambulance drivers and patients alike. With a stable server system, we should be able to track the ambulance or receive the position of an ambulance from external software. The health facilities should be located on a map and supplied with pertinent information in the application. This should ensure administrators or operators at the headquarters can choose the appropriate clinic in the case of an emergency. Furthermore, the system should be able to take over the ambulance's routing and send the driver information with the destination's coordinates or a predetermined route. To guarantee that the crises are documented, it is also critical that reports on the different activities be generated in the system.
This work aims to answer the following questions with the help of various plugins, settings and options in QGIS:

- To what extent can QGIS be used as a fleet management system?
- What adjustments need to be made?
- How can the appropriate geospatial data be integrated?
- How far can QGIS software be minimised to provide the user with a light version?
- How can communication with the ambulance take place?
- Which plugins are required for implementation?
- How can these components of the system be brought together in QGIS?

2 Study Region

The study region is the municipality of Ashaiman (Ghana) and neighbouring settlements, s. Fig. 1. Until 2008, Ashaiman was still part of the city of Tema and currently has an area of approx. 45 km². The area is characterized by a very high level of migration and unemployment. The population tripled between 1984 and 2000 (ARDAYFIO-SCHANDORF et al. 2012: 25). According to the statistical office in Ghana, Ashaiman had a population of about 235,000 in 2019. A high number of unreported cases can be assumed due to the high level of migration and the slums in the area. If you convert this statistical number of inhabitants to the area of Ashaiman, you get a population density of approx. 5200 people per km². The general location of Ashaiman in the Greater Accra Region, Ghana is 5°42′N 0°02′W and it is bounded by minimum longitude and latitude of 0° 3' 35.33" W and 5° 40' 41.16211" N respectively; and by maximum longitude and latitude of 0° 0' 46.80176" W, 5° 43' 23.100" N respectively.

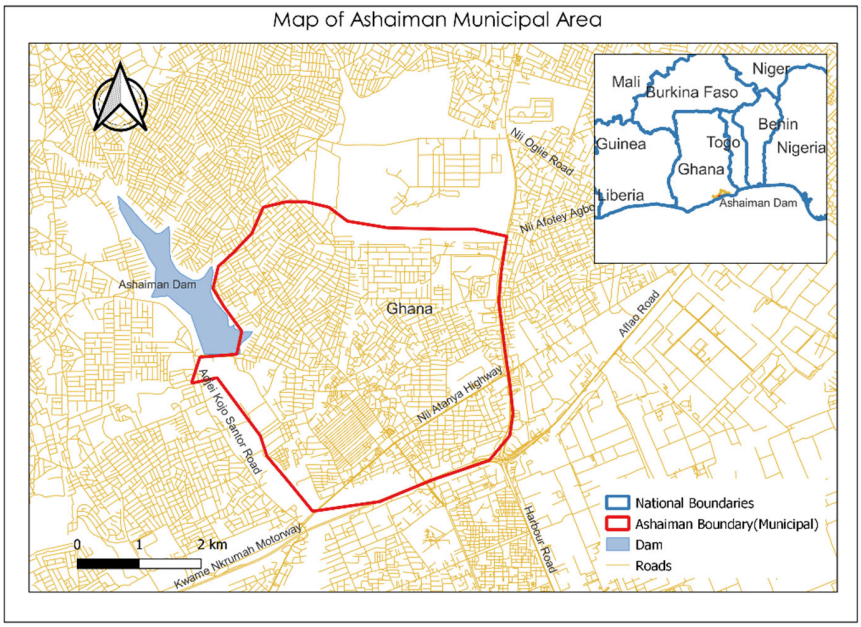


Fig. 1: Target Area Map

3 Problem Statement

The road infrastructure and condition of the roads in Ashaiman are inadequate. The few roads that exist are mostly not passable. A survey in 2012, in which 172 residents in Tulaku Township in Ashaiman were questioned, showed that more than half of the respondents said that a fire truck could not reach their home due to the poor road conditions (SARPONG 2013: 67). These survey results can be easily transferred to a scenario for medical transport since fire engines and ambulances usually have a similar requirement profile.

Strong environmental impacts, such as rain, have an impact on the roadways. During the rainy season, even paved roads in the region are vulnerable to flooding. The problem is exacerbated by the absence of drainage channels and the fouling of existing drains. The majority of Ashaiman's streets are not paved, making them even more vulnerable to the aforementioned influences: The streets are softened, and water collection holes are constructed. The fact that automobiles travel on these roads exacerbates the problem by allowing potholes to appear. Many of these routes are impassable during the rainy season as a result of these circumstances. During the project's duration, it's critical to determine which routes are accessible to an ambulance and which are not. To improve the route finding process, the latter might be eliminated from the routing. Traffic is also a major problem in Ashaiman. Concerning the size of the city, there are few street signs and speed bumps to regulate speeds and traffic (GUMAH 2015:12). As a result, accident rates in these locations are often high, necessitating greater medical care via ambulances. Traffic congestion is common, especially during the so-called "rush hour," due to insufficient infrastructure. Because of the absence of traffic organization and regulation, as well as the numerous traffic jams, motorized vehicles in metropolitan areas may only go at a speed of five to ten kilometres per hour on average.

As a result, getting a vehicle to a patient and a healthcare institution in the Ashaiman region takes a lengthy time. In an emergency, the ambulance management system should aid gain valuable minutes, which may frequently save a patient's life.

4 Related Works

Planning and administration are critical in hospitals or trauma centres for timely audit and treatment of trauma patients to prevent mortality caused by accidents and unforeseen events (AMINZADEH et al. 2017). As a result of this, we need to make use of effective systems that ensure timeliness in the transport of patients who may not be detained.

A sophisticated system developed by SANKAR et al. (2016)used a server, mobile applications, and GPS to track all ambulances, detect accident locations, send the nearest ambulance to the accident site, and ultimately, monitor the ambulance driver's performance. Google Maps layers were utilized to determine the best route.

RAABER et al. (2016)also employed prehospital geographical information systems (GIS) to handle the problem of overcrowding and handover. The use of GIS also influenced nurses' perceptions of GIS data as a tool for maximizing resource utilization and improving the quality of all patients' reception, whether they were critically or non-critically ill. Using traffic speed data, CHO & YOON (2015)proposed a GIS-based method to calculate the k-minute journey time contour to reflect the response coverage in Seoul, South Korea, considering time-of-day and day-of-week effects on travel time.

ESRI are producers of suites of commercial software for diverse problem domains. Esri® ArcLogisticsTM software provides a comprehensive solution for complicated routing and scheduling problems, allowing you to cut costs, increase productivity, and improve customer experience (ESRI 2010). We can also look at a prototype they developed in 2021. The functional requirements for the AJAX and Adobe Flex-based prototype fleet management application encompass- geographic coverage area, vehicle locations from an XML file, display of the vehicles location and related functions. The prototype makes use of the ArcWeb Services API family, which includes SOAP, REST, OpenLS, and Java ME, all of which are hosted by ESRI (GUP 2021).

It is worth noting that a decade or more ago, PASHA (2006) already developed an Ambulance Management System prototype using GIS/GPS/GSM using ArcGIS9.1.

Hexagon Geospatial is also a commercial software company that provides extra features to their GeoMedia tiers' fundamental capabilities. Municipalities, airports, seaports, departments of transportation (DOTs), train firms, waterway agencies, and pipeline operations can use GeoMedia Transportation Manager to efficiently examine their transportation infrastructure. This includes the ability to track network data as well as execute specialized analytics (HEXAGON 2021).

KARKAR (2019) proposed a smart ambulance system based on GIS that can designate a route as an emergency route. It allows paramedics to transport patients to the hospital more quickly. The system, made up of three parts: the server program, the user emergency end-user application, and the paramedic end-user application.

As can be seen from the few works mentioned, the commercial approaches take centre stage when it comes to these systems. Attempts have, however, been made to use open source software to

create EAMS. For example, Traccar is an open-source GPS tracking system that works with a variety of GPS tracker devices. The software was first released in 2009 by Russian developer Anton Tananaev and has been actively developed by developers all around the world since then. It is composed of the Jetty Java HTTP server, the MySQL database, and the Netty network pipeline architecture. Traccar server is only needed to control and record GPS tracker position (SENIMAN et al. 2020).
An open-source development environment, MapServer, was used to construct the mapping component of the fleet management system (MEDAGAMA et al. 2008). In this system, a webserver and mail server connected to GPS utilised an Oracle database. Oracle Database (commonly referred to as Oracle DBMS or simply as Oracle) is a multi-model database management system produced and marketed by Oracle Corporation (ORACLE 2017). So here we have a system that is not entirely free because of the use of a commercial database, Oracle.

5 Methods

5.1 Fleet Management

Managing ambulances requires knowledge of fleet management. This is vital for the optimal deployment of the vehicles. Fleet management, in general, includes the targeted disposition, management and control of a vehicle fleet, taking into account internal and external influencing factors (JUNG & VAN LAAK 2001). In the particular case of our customised system, we took into consideration the road network and to some extent the constraints that exist and can determine the availability of ambulances within the fleet of vehicles.
Fleet management systems should have certain requirements and capabilities and be tailored to specific application needs. The most important component for such a system is the tracking and tracing of the ambulance. Tracking is understood to mean the current position at the time of observation. In the case of tracing, on the other hand, position data from the past are included. The snapshots of the tracking are quasi strung together to make the change in position visible depending on the time (JUNG & VAN LAAK 2001: 34). The system developed in the EAMS, made it possible to track the vehicles permanently in real-time so that they could be dispatched as needed in an emergency. At the control centre, the computers or as is in the ideal situation, the server hosts the GIS application instance, database server; all of this connect to the internet for communication.

5.2 Geographic Information Systems

A Geographic Information System (GIS) is a computer-based system that provides the following four sets of capabilities to handle georeferenced data: data capture and preparation; data management, including storage and maintenance; data manipulation and analysis; data presentation.
We think of geographic phenomena as a collection of geographic objects when it isn't present everywhere in the research region but 'sparsely' populates it. Typically, such things are easily identified and named, and their location in space is determined by a combination of one or more parameters (HUISMAN & BY 2009). In this study, an ambulance management system is to be implemented in a GIS. This special case can be classified as a geographic information system in

transport and logistics (GIS-T). QGIS has been used at the core of the system to create our own GIS-T to manage the fleet of ambulances. The following plugins(functions) were developed - *Ambulance Location*, *Report Incident, GeoCoding*, *Location Help, Sending Mail*, *Information health facilities and Capture Coordinates.*
Qt designer was used for the user interface whiles python was used to write code that gave the controls functionality.

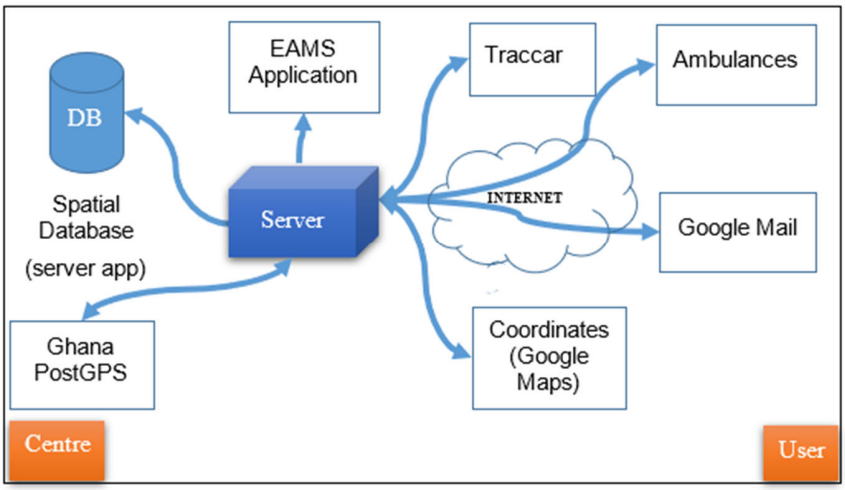


Fig. 2 System and Data Architecture

5.3 Components of an EAMS management system

The EAMS make use of several available components and new programmed Python code. All components of the system are shown in Fig. 2 and explained in the following sections.

5.3.1 Background Data layers

For the background data layer, the OpenStreetMap, Tileserver; Geofabrik geodata sources were used. Free geodata is extracted, filtered, and processed by the Geofabrik. Shapefiles, map pictures, tiles, and entire web map applications are among the services Geofabrik provide.
A tile server is a service that generates rendered images (tiles) from a database. Maps that are created with the help of such tiles are very efficient in terms of zoom and shift functions. The individual tiles on a map usually have an area of 256×256 pixels and are processed in PNG format. The tiles are cached on the server and client sides. Tile servers are superior to other technologies such as maps that have been loaded with the help of a WMS (Web Map Service) (GEOFABRIK GMBH 2020).
The standard OpenStreetMap (OSM) map is the first tile server to be integrated into QGIS. OSM dubbed the "Wikipedia of maps," is a free alternative to commercial map products that are backed by a big community of people who are interested in reproducing and comparing aerial images. In urban areas, the quality and timeliness of OSM data are even better compared to most commercial products. In contrast, the quality of the information in rural regions cannot always be judged to be good due to the small number of project supporters there (KEßLER 2015). In terms of coverage,

OSM data on roads are available for most parts of our study region but may not be guaranteed by parts characterised with informal settlements.

5.4 Open Source GIS Software Customisation - QGIS

It is possible to develop a complete application in a Python environment. Python in QGIS can be used to execute commands from the console, execute code automatically when QGIS starts, write custom expressions and actions, create new processing algorithms, create plugins and create standalone applications in QGIS (GANDHI 2020). The plugins were created to address the needs of the system by using the python programming language in QGIS.

5.5 Data Exchanges and Messaging Services

One of the most important factors in the project is the transmission of the ambulance coordinates to the control centre. It is essential that there are few failures as possible during the transmission and that an almost smooth process can be guaranteed. We need to ensure that the transmission works almost without errors. The three communication channels used by the control centre to interact with the ambulance, WhatsApp, Telegram, and Traccar, are integrated into the system via different interfaces. One of the application's requirements is that as many features as feasible be stored in QGIS. A link between the ambulance vehicles and QGIS was essential. Control panels were constructed with the Plugin Builder to integrate WhatsApp, Telegram, and Traccar with QGIS, allowing the platforms to be launched in external windows.

In addition to the live location from Traccar, there is the option of automatically importing the coordinates from WhatsApp into QGIS in case of a failure. It is important to update the current live position of the Google Maps Live location sent via WhatsApp and to centre the map as far as possible on the currently required ambulance.

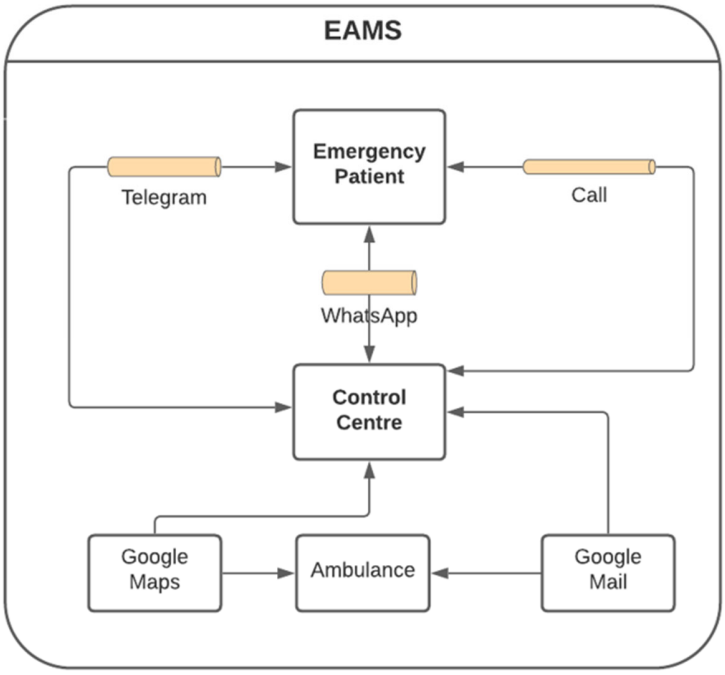


Fig. 3: EAMS Communication Network (Human)

Traccar is a GPS tracking system that is open-source. The software distinguishes itself by its flexibility, adaptability, and 'opensourceness'. Traccar enables you to track devices in real-time and view their coordinates at any time. Traccar has been set up on a DigitalOcean virtual private server (VPS) and can be accessed via a uniform resource locator (URL). A VPS is a virtual machine sold as a service by an internet hosting provider. This makes it possible for us to access an instance of Traccar server from anywhere in the world via an Internet protocol (IP) address.
Another task that must be completed as part of the project is to electronically convey information about an occurrence to the ambulance drivers. The Sending Mail plug-in was created as part of the project for this purpose.

5.6 Geocoding

Since the caller who wants to request help from an ambulance often does not know the coordinates of his accident/place of operation, it is essential to integrate a function for geocoding into the QGIS system. With geocoding, a text-based description of a place, e.g. an address or the name of a location, is entered and geographic coordinates, i.e. longitude and latitude, are returned to identify the location on the earth's surface.
The system uses the QGIS extension GeoCoding for geocoding, which makes it possible to use data from Nominatim, the geocoder from OpenStreetMap, or the Google Geocoder, within QGIS (NOMINATIM 2021). If the plug-in finds more than one place that fits the address or description, this extension shows several options in which the selection must now be made manually by the user.
In some countries, such as Ghana in this example, it is not uncommon that the street names are not known. For this case, the GhanaPostGPS system, which is common there, was additionally integrated as geocoding via API. In this system, there are unique codes for addresses in the form of squares with a side length of five metres.

5.7 Spatial Database and Routing

A database consists of a collection of data that are logically related to one another and organized by a database management system (DBMS). The data is saved according to predefined rules and is easily accessible. A programming language can be used to access the database and manage the data. By far the most important international database interface is SQL (Structured Query Language) (SCHICKER 2017).
Routing is a critical component of the system. The ambulance should be given the choice of routing to the individual in need and then to the nearest healthcare institution based on location data. There are numerous methods for routing and determining the shortest pathways between places. Edsger W. Dijkstra initially described and published the Dijkstra algorithm in 1959 (DIJKSTRA 1959: 269ff). In the initialization step, this method is given a coherent weighted graph (a network) with nodes and edges, as well as a beginning node from which the observation begins.
The A * algorithm is considered to be a generalization and extension of the Dijkstra algorithm. It is also known for searching for the shortest path, but in contrast to Dijkstra's algorithm, not only the path lengths are part of the calculations. With a database set up with which data about the hospitals can be displayed. For routing, the roads vector layer of Ashaiman was used together with

its routing system based on a database, leveraging on the QGIS shortest path (point to point) function.
A separate routing function has been integrated into the system. Road data from the project area was downloaded from the OpenStreetMap database and used for routing. The routing function uses the A* algorithm to find the shortest route from one point to another.
The system differs from other routing options in that the availability (free or occupied) of the ambulances is also included. To see which ambulance is currently occupied and which is not, the individual health facilities have been assigned geo-fences in Traccar. If an ambulance now receives a new order, its status changes to occupied. If the ambulance now enters the area of a health facility and delivers the patient there, the status is changed back to free and the dispatcher who operates the EAMS sees that the ambulance can accept a new order.

5.8 User Interface

The user interface of plugins in QGIS is created with the help of the Qt Designer. To achieve a reduction in the surface area, it is necessary to use the surface adaptation function in QGIS. All options on the graphical user interface that do not make sense for the application are hidden. Thus, the user is provided with a light version of QGIS tailored to the needs of an ambulance management system.

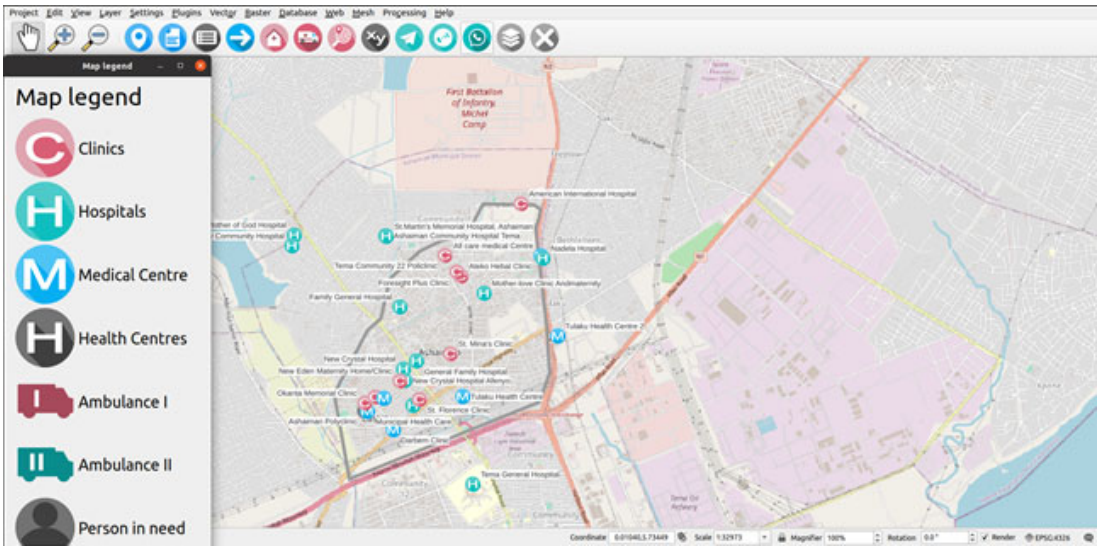


Fig. 4: QGIS Customised User Interface

6 Results and Tests

To test the developed application, the scenario of an emergency and an incoming emergency call was simulated. Concerning the functions of the system, the EAMS was developed with a local doctor in Ashaiman. He tested the developed functions and suggested improvements. As someone who has never worked with GIS before, the intuitiveness of the system and the practical application of the functions could be tested alongside the functions themselves.

On the integration of Traccar's functions into QGIS, no ambulances were available for testing purposes so far. However, their function could be simulated with faked GPS locations via a smartphone. This also made it possible to test Traccar's functions and their integration into QGIS. With regards to installation and subsequent use of plugins on another system, ZIP files can be set up using the *Install from ZIP* function in the QGIS plugin menu. This makes it possible to install and use the plugins on other computers. If a plugin uses add-ons such as browser history, it is also necessary to install them on the respective computer before using an extension.
Furthermore, the tests with the application show that QGIS is very well suited for the implementation of the project in Ashaiman. This geographic information system has the property of being highly configurable and adapting to any scenario with the help of plugins or other developments.
In the course of this work, a method was developed to automatically update the locations of the ambulances in the QGIS system in real-time, as is usual with a commercial fleet management system. The automatic receiving of the coordinates is done through a Python interface that was programmed. This reads the real-time coordinates of the ambulances from Traccar and passes them to QGIS. QGIS can then display this data directly. This interface automatically queries the location every 2 seconds.

7 Conclusions & Outlook

This study sort to find out the possibilities within the implementation of GIS for a fleet management system. The existing and available system for starters is almost free, low-cost and based on open-source software save the necessary but marginal cost of static IP address subscription to be used for the Traccar system. Nonetheless, there is the need to do adjustments and setting modifications and alterations to the open-source cost software to convert it to the semblance of a fleet management system or something closer. There is a need to consider certain factors concerning data integration – i.e. how do we bring data and integrate or make it a part of the system? We found out that the possibilities exist in making connections between the different components of the system; making use of the internet. So we have been able to achieve somewhat data integration and this was possible through the use of the interoperability between Traccar software, data server hosting PostgreSQL and the use of GPS data from mobile devices. To ensure that the application is useful and accessible to the user in the most portable way is to warrant that they could run the application without limitations. That is to say, the operating system and machine specifications do not hinder the installation process, for example. An ambulance system to work requires interconnected parts. We need to ensure these communication processes are functional and real-time, and we were able to achieve this by making use of the web applications that allowed effective and reliable communication. QGIS plugins are possible to create from scratch or customize. With the use of a programming language like Python, this was possible to do within the application and without much cognitive overhead – the development environment was amenable to allow for adaptation to meet user needs.
Furthermore, further tests with simulated incidents are being planned with ambulances on site, also for the training of new staff in the Control Centre and ambulances. It is also interesting for the use

of Traccar that, in addition to the live location of the ambulances, we also enable the geofences to display in QGIS which ambulance is occupied and, when entering the geofences of the assigned health facility, to display it is free again. More tests and analyses are possible and will reveal the effectiveness and accuracy issues with the current system.

8 Bibliography

AMINZADEH, S., RABIEE, M., REZAEI, Y., FARNAGHI, M. & BAHMAN-ABADI, M., 2017: The role of GIS in Information Systems, Management and Decision Making in Hospitals and Ambulance Centers. International Journal of Mind, Body and Culture, **4**, 18-23, https://doi.org/10.22122/ijbmc.v0i0.89.

ARDAYFIO-SCHANDORF, E., YANKSON, P.W. & BERTRAND, M., 2012: The Mobile City of Accra. African Books Collective, https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers13-04/010056996.pdf

CHO, J. & YOON, Y., 2015: GIS-Based Analysis on Vulnerability of Ambulance Response Coverage to Traffic Condition: A Case Study of Seoul. IEEE Conference on Intelligent Transportation Systems, 1402-1407, https://doi.org/10.1109/ITSC.2015.230.

DIJKSTRA, E.W., 1959: A note on two problems in connexion with graphs. Numerische Mathematik, **1**(1), 269-271, https://doi.org/10.1007/BF01386390.

EGGER, M., LOW, N., ZÜRCHER, K. & RAZUM, O., 2021: Globale Gesundheit. Public Health Kompakt, 4. Auflage, De Gruyter, 507-530, https://doi.org/10.1515/9783110673708-010.

ESRI, 2010: ArcLogisticsTM : The Routing and Scheduling Solution for Fleet Management. www.esri.com.

GANDHI, U., 2020: Customizing QGIS with Python (Full Course Material). https://courses.spatialthoughts.com/pyqgis-in-a-day.html. https://courses.spatialthoughts.com/pyqgis-in-a-day.html

GEOFABRIK GMBH, 2020: Tiles. https://www.geofabrik.de/de/maps/tiles.html.

GUMAH, G., 2015: Spatio-temporal analysis of road traffic accident on the Accra-Tema motorway: causes and risk factors (MPhil Thesis). 10227422, 194, http://ugspace.ug.edu.gh.

GUP, A., 2021: Easily integrate Web-based GIS using the ArcWeb Services JavaScript API. www.esri.com.

HEXAGON, 2021: GeoMedia: Dynamic GIS Mapping Software for a Changing World. https://www.hexagongeospatial.com/products/power-portfolio/geomedia/geomedia-dynamic-gis-mapping-software-changing-world.

HUISMAN, O. & DE BY, R. A., 2009: Principles of Geographic Information Systems. ITC Educational Textbook Series, **1**(17).

HUPP, M., SCHMIDT, M., OSCHMANN, M., SCHOLER, M., AIT, L., STÄBLEIN, S., GÖB, M., BECKMANN, M., KATZENBERGER, M., HILPERT, A., WEIS, R., GOEB, M., WOLF, N., HAJI, N., HOSSEIN, M. & BRUNN, A.,2019: Ein Konzept für einen effektiven Krankenwagendienst zur Reduktion der Müttersterblichkeit, in der Schwangerschaft und nach der Geburt, sowie der Sterblichkeit generell in Ashaiman (Tema, Ghana). Projektbericht, Hochschule für angewandte Wissenschaften Würzburg/Schweinfurt.

JUNG, J., & VAN LAAK, B. L., 2001: Flottenmanagementsysteme – Grundlegende Technologien,

Funktionen und Marktüberblick. Arbeitsbericht des Instituts für Wirtschaftsinformatik, Universität Koblenz. **28**.

KARKAR, A.G., 2019: Smart ambulance system for highlighting emergency-routes. 3rd World Conference on Smart Trends in Systems, Security and Sustainability, WorldS4 2019, 255-259, https://doi.org/10.1109/WorldS4.2019.8903948.

MEDAGAMA, M., GAMAGE, D., WIJESINGHE, L. & LEELARATNA, N., 2008: GIS/GPS/GPRS and Web-based Framework for Vehicle Fleet Tracking. Journal of the Institution of Engineers, Sri Lanka, **41**(5), https://ent.uom.lk/dialog/documents/Geoinfomatics_final_compact.pdf.

NOMINATIM, 2021: Open-source geocoding with OpenStreetMap data. https://nominatim.org/.

ORACLE, 2017: Multimodel Database with Oracle Database 12c Release 2. Oracle White Paper, https://download.oracle.com/otndocs/products/spatial/pdf/12c/Multimodel_Database_with_Oracle_Database_12c_Release_2.pdf

PASHA, I., 2006: Ambulance management system using GIS. Master thesis, Linkoping University Sweden, https://www.diva-portal.org/smash/get/diva2:22123/FULLTEXT01.pdf.

RAABER, N., DUVALD, I., RIDDERVOLD, I., CHRISTENSEN, E.F., & KIRKEGAARD, H., 2016: Geographic information system data from ambulances applied in the emergency department: Effects on patient reception. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, **24**(1), 1-9, https://doi.org/10.1186/S13049-016-0232-5/FIGURES/2.

SANKAR, S.H., JAYADEV, K., SURAJ, B. & APARNA, P.,2016: A comprehensive solution to road traffic accident detection and ambulance management. IEEE International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEES), 43-47, https://10.1109/ICAEES.2016.7888006.

SARPONG, A.O., 2013: Fire Risk Vulnerability in Informal Settlements. The Case of Ashaiman. 174, http://ugspace.ug.edu.gh.

SCHICKER, E., 2017: Datenbanken und SQL. Springer Fachmedien Wiesbaden.

SENIMAN, ZENDRATO, N., ARISANDI, D. & LUBIS, F., 2020: Traffic Light Controlling for Emergency Vehicle Line Based on GPS Tracking and Position Using GPRS Network. Journal of Physics: Conference Series, 1566(1), https://doi.org/10.1088/1742-6596/1566/1/012014.

Monitoring von blütenreichen Flächen mittels Fernerkundung

SOPHIE PROKOPH[1], JASMIN CHEEMA[2], ANITA KIRMER[3], ANGELA LAUSCH[4] & LUTZ BANNEHR[1]

Zusammenfassung: Seit einigen Jahren werden in der Landwirtschaft gezielt blütenreiche Flächen angelegt, um die Nektar- und Pollenquellen für Bestäuber und andere Nützlinge zur Verfügung zu stellen, Trittsteine in der Landwirtschaft zu schaffen und damit dem Rückgang der Artenvielfalt entgegenzuwirken. Für den Erfolg dieser Maßnahmen ist die Anzahl der etablierten Pflanzenarten und der Blühaspekt über die Vegetationsperiode entscheidend. Dieser Artikel gibt einen Einblick, wie Fernerkundung dabei helfen kann, die Erfolgs- und Effizienzkontrolle, welche bisher mit aufwändigen Vor-Ort-Erfassungen geprüft wird, einfacher zu gestalten. Untersucht wurde einerseits, ob und wie sich blütenreiche Flächen anhand von hochauflösenden RGB-Luftbildern klassifizieren und von anderen Landbedeckungsarten unterscheiden lassen. Die Fernerkundungsdaten wurden mit dem Gyrocopter der Hochschule Anhalt von Versuchsanlagen (Feldraine, mehrjährige Blühstreifen, Buntbrachen) auf und um den Campus Strenzfeld der Hochschule Anhalt (bei Bernburg, Sachsen-Anhalt) erhoben. Das Vorgehen bei der Identifikation von Blühflächen bestand im Wesentlichen aus drei Schritten: Segmentierung, GLCM-Berechnung und Klassifikation. Die Ergebnisse sind als sehr positiv zu bewerten (Overall Accuracy > 0.98), da bei der Klassifikation nur wenige Segmente zwischen den Klassen „Blühflächen“ und „nicht-Blühflächen“ verwechselt wurden. Des Weiteren wurde anhand hochaufgelöster RGB-Daten versucht, den Artenreichtum, die Blütenfarbe und -menge sowie den Einfluss von Gräserreichtum auf die Ergebnisse über eine RGB-Farbraumanalyse abzuleiten. Ebenso sollte geprüft werden, welche Datengrundlage und räumliche Bodenpixelauflösung für die Fragestellungen geeignet sind. Die Ergebnisse zeigen, dass eine qualitative Einschätzung des Blütenreichtums möglich ist; für eine quantitative Einschätzung spielt allerdings die Bodenpixelauflösung eine entscheidende Rolle. Der Artenreichtum blühender Pflanzen korreliert mit der Anzahl der vorkommenden Blühfarben und kann somit grob abgeschätzt werden. Ein hoher Gräserreichtum wirkt sich nicht negativ auf die prozentual ermittelte Blütenmenge aus. Bei der Einschätzung der vorwiegenden Blühfarbe besteht noch Optimierungsbedarf.

1 Einleitung und Stand der Forschung

Seit 1990 wurde in Naturschutzgebieten ein besorgniserregender Biomasserückgang der Fluginsekten von etwa 75 % festgestellt (HALLMANN et al. 2017; SORG et al. 2013; RADA et al. 2018). Als bedeutende Einflussgröße wird die Landwirtschaft gesehen, da eine verstärkte Flächeninanspruchnahme, Monokulturen, der Einsatz von Schädlingsbekämpfungsmitteln, sowie verändertes

[1] Hochschule Anhalt, FB3 AFG, Bauhausstraße 8, D-06846 Dessau, E-Mail: [Sophie.Prokoph, Lutz.Bannehr]@hs-anhalt.de
[2] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden, E-Mail: jasmin.cheema@tu-dresden.de
[3] Hochschule Anhalt, FB1 LOEL, Strenzfelder Allee 28, D-06406 Bernburg, E-Mail: anita.kirmer@hs-anhalt.de
[4] Helmholtz-Zentrum für Umweltforschung GmbH – UFZ, Department Landschaftsökologie, Permoserstraße 15, D-04318 Leipzig, E-Mail: angela.lausch@ufz.de

Flächenmanagement, etc. als Faktoren zusammenwirken (NABU - NATURSCHUTZBUND DEUTSCHLAND E.V. 2022).

Heutzutage wird das Landschaftsbild in Deutschland von großen Nutzflächen geprägt, in denen es kaum noch integrierte oder vernetzte Biotope für Pflanzen und Tiere gibt. Diesem Trend versucht man, z. B. durch Agrarumwelt- und Klimamaßnahmen (AUKM) entgegenzuwirken. In Sachsen-Anhalt wird u. a. die Anlage von mehrjährigen Wildpflanzen-Blühstreifen oder -flächen und Buntbrachen über AUKM subventioniert (FENCHEL et al. 2015). Darüber hinaus können artenarme, grasdominierte Randstrukturen zwischen Wegen und Ackerflächen durch die Ansaat von artenreichen Wildpflanzenmischungen in kräuter- und blütenreiche Feldraine umgewandelt werden, die in ausgeräumten Landschaften eine wichtige Funktion als Verbindungskorridore und Trittsteine übernehmen (KIRMER et al. 2018, 2019).

Der Erfolg der Ansaaten ist abhängig von der verwendeten Saatgutmischung sowie vom Management und kann beispielsweise über die Anzahl und Deckung der angesäten und spontan etablierten Pflanzenarten (Zielarten, Gräser, Ruderalarten, invasive Neophyten) bzw. über die vorkommenden Insektenarten kontrolliert werden (KIRMER et al. 2018; SCHMIDT et al. 2020, 2021).

Mit verschiedenen Fernerkundungstechnologien können Zustände, Zustandsänderungen und Wechselwirkungen von Pflanzeneigenschaften (plant traits) großflächig, kontinuierlich und zunehmend harmonisiert erfasst werden (LAUSCH et al. 2016; PAUSE et al. 2020).

Dass sich mit Fernerkundung Vegetationsveränderungen, insbesondere auch bei Blühpflanzen aufzeigen lassen, zeigen folgende Beispiele:

- MÜLLEROVÁ et al. (2013) hat nachgewiesen, dass sich der invasive Neophyt Riesen-Bärenklau (auch als Herkulesstaude bekannt) durch seine riesigen Blütenstände in Satellitenaufnahmen detektieren lässt.
- SMIGAJ und GAULTON (2021) haben mit einer Random Forrest Klassifikation anhand von hochaufgelösten RGB- und multispektralen Daten den Blütenreichtum in Hecken untersucht.
- LANDMANN et al. (2015) hat in afrikanischen Savannen mithilfe von hyperspektraler Satellitenfernerkundung Blühpflanzen kartiert.
- FEILHAUER et al. (2016) konnte unter Einsatz von hyperspektralen airborne Fernerkundungs-Technologien drei verschiedene Bestäubungsarten klassifizieren.

In diesem Artikel wird einerseits eine Möglichkeit aufgezeigt, Blühstreifen/ -flächen anhand von hochaufgelösten Luftbilddaten zu klassifizieren, andererseits wird ein Ansatz vorgestellt, um die Blütenfarbe und -menge abzuleiten. Die Untersuchung erfolgte mittels KI-Klassifikationsalgorithmen, statistischen Verfahren sowie einer RGB-Farbraumanalyse.

Die verwendeten fernerkundlichen Methoden sollen dazu beitragen, die Erfolgs- und Effizienzkontrolle von Ansaaten einfacher zu gestalten. Folgende Fragestellungen wurden dabei bearbeitet und sind Gegenstand des vorliegenden Papers:

- Wie lassen sich blütenreiche Flächen anhand von multisensoralen Luftbildern in einem Gebiet klassifizieren und von anderen Landbedeckungsarten unterscheiden?
- Lassen sich Artenreichtum, Blütenfarbe und -menge anhand hochaufgelöster RGB-Daten ableiten? Hat das Vorhandensein von Gräsern auf diese Erfassung eine Auswirkung?
- Welche Datengrundlage ist für die Fragestellungen notwendig/ausreichend?

2 Untersuchungsgebiet und Datenerfassung

Das Untersuchungsgebiet Strenzfeld liegt in der Nähe von Bernburg in Sachsen-Anhalt, Deutschland (Abb. 1). Hier befinden sich auf bzw. um den Campus Strenzfeld der Hochschule Anhalt verschiedene Versuchsanlagen für Feldraine, mehrjährige Blühstreifen und Buntbrachen (Abb. 2).

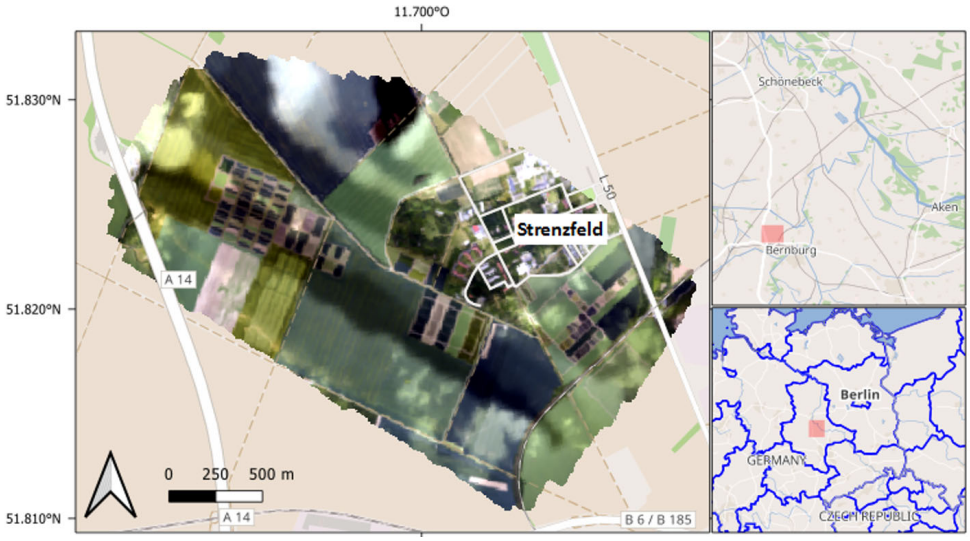


Abb. 1: Lage des Untersuchungsgebietes als RGB-Orthomosaik.

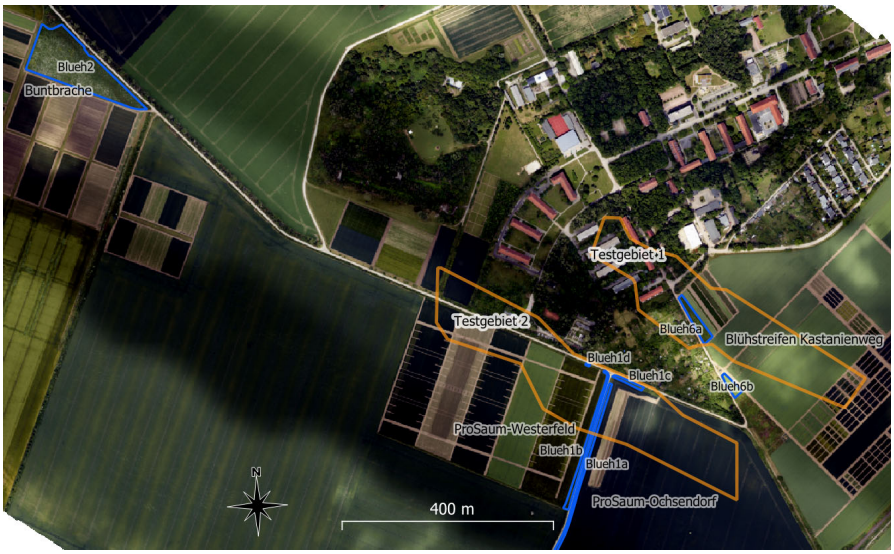


Abb. 2: Lage der ausgewählten Versuchsanlagen für Blühstreifen und -flächen (blau) und Testgebiete (orange). Bei dem Feldrain Blueh1a (ProSaum-Ochsendorf) handelt es sich um einen 2010 angelegten Blockversuch aus sechs Varianten und fünf Wiederholungen, in dem vier Varianten mit einer artenreichen Ansaatmischung aus 49 gebietseigenen Wildpflanzen angesät wurden (Kirmer et al. 2018). Hier wurden die ersten 12 Parzellen betrachtet (im Norden). Blueh1b ist ein 2011 angesäter artenreicher Feldrain. Beide Feldraine beinhalten unterschiedliche Mahdtermine (Mai, Juni, September). Blueh1c und Blueh1d sind kleine Flächen mit Ansaaten. Bei Blueh2 handelt es sich um eine Buntbrache. Blueh6a und Blueh6b sind mehrjährige Blühstreifen neben dem Kastanienweg. Testgebiet 1 beinhaltet Blueh6a. Testgebiet 2 beinhaltet Blueh1c, Blueh1d und Teile von Blueh1a sowie von Blueh1b.

Das Untersuchungsgebiet wurde 2021 zu verschiedenen Zeitpunkten (Mai, Juni, August, September) mit dem Gyrocopter der Hochschule Anhalt beflogen. Der Vorteil dieser Trägerplattform gegenüber Satelliten besteht darin, dass eine wesentlich höhere räumliche Auflösung der Bilddaten erreicht werden kann (sensorabhängig, < 10 cm). Diese ist zwar in der Regel nicht so hoch ist wie bei UAVs (Unmanned Aerial Vehicle), aber im Vergleich dazu ist eine höhere Gebietsabdeckung pro Überflug möglich sowie eine größere Vielfalt an parallel installierter Fernerkundungssensorik möglich.
Der Gyrocopter ist mit einer komplexen Sensorik, wie einem Hyperspektralscanner (HySpex VNIR 1600), einer hochauflösenden 100 Mpix RGB-Kamera (Phase One iXM-100) sowie einer Thermalkamera (FLIR A655sc) ausgestattet (BANNEHR et al. 2021). Die Befliegungen wurden in Flughöhen von 300 m bzw. 500 m durchgeführt, um einzelne Blütenstände, die nur wenige Zentimeter groß sind, erfassen zu können. Mittels eines GPS (Global Positioning System) und einer IMU (Inertial Measurement Unit) können die resultierenden Bilddaten lagegenau verortet werden. Nach einer GPS-Korrektur mit dem Inertial Explorer erfolgte die weitere Datenprozessierung (im Wesentlichen geometrische Entzerrung) mit der Software Agisoft Metashape für die Thermal- und RGB-Daten (getrennt) und mit Parge für die Hyperspektraldaten.
Für diese Untersuchungen wurden vorerst nur die Daten vom 13. Juni 2021 betrachtet. Während der Befliegung war wechselhaftes Wetter, was sich besonders im RGB-Bild durch verschieden stark belichtete Bereiche widerspiegelt (Abb. 1) und eine Herausforderung bei der Datenauswertung darstellt. Die resultierenden Luftbilder haben die folgende räumliche Auflösung: 4 cm (RGB-Orthomosaik), 20 cm (bildbasiertes digitales Oberflächenmodell, bDOM), ca. 15 cm (Hyperspektraldaten) sowie ca. 65 cm (Thermaldaten). Für die Analysen in dieser Studie wurden zunächst die RGB-Daten herangezogen. Die Hyperspektral- und Thermaldaten sollen aber zukünftig noch bei der Beantwortung weiterer Fragestellungen eingesetzt werden. In Tabelle 1 sind einige weitere Metainformationen zu den verwendeten Luftbildern zusammengefasst. Bei der Genauigkeitsangabe der räumlichen Auflösung ist zu beachten, dass der angegebene Wert ohne Berücksichtigung einer Forward Motion Compensation-Korrektur berechnet wurde und somit die von Metashape berechnete, aber nicht die tatsächliche räumliche Auflösung darstellt.

Tab. 1: Räumliche und radiometrische Auflösung sowie spektraler Bereich für die jeweiligen Bilddaten von der Befliegung am 13.06.2021.

	Räumliche Auflösung		**Radiometrische Auflösung**	**Spektraler Bereich**
	Bei 300 m Flughöhe	**Bei 500 m Flughöhe**		
RGB-Orthomosaik	≈ 4 cm	≈ 5,4 cm, 11 cm	16 bit	rot, grün, blau
bDOM	≈ 8,6 cm	20 cm	-	-

Zudem erfolgten am 15. und 16. Juni 2021 vor Ort Erhebungen von Referenzdaten. Unter anderem wurden Blütenreichtum (BRT), Gräserreichtum (GRT) und Artenreichtum (ART) mittels einer Skala von 0 (kein Vorkommen) bis 5 (sehr häufiges Vorkommen) abgeschätzt sowie die dominanten Blütenfarben und Pflanzenarten, Wuchshöhe und der Mahdstatus dokumentiert.

3 Methodik/Datenanalyse

3.1 Klassifikation blütenreicher Flächen

Bei dem Vorkommen mehrerer Pflanzenarten innerhalbeiner blütenreichen Fläche handelt es sich nicht um eine Objektklasse, die auf einer einzelnen spektralen Signatur mit Variationen beruht, sondern um eine Reihe von Klassen, deren Reflexionsverhalten sich massiv voneinander unterscheiden können und die räumlich zusammen vorkommen. Pixelbasierte Methoden, die die benachbarten Pixel bei der Klassenzuweisung nicht beachten, geraten deshalb an ihre Grenzen. Demnach liegt es nahe, den räumlichen Kontext der umliegenden Pixel und somit deren spektrale Information in die Klassifikation mit einzubeziehen. Hier wurden zwei Ansätze ausgewählt, um die relative Lage der Pixel zu berücksichtigen. Der erste besteht darin, eine überwachte objektbasierte Klassifikation basierend auf einem GEOBIO Ansatz durchzuführen (BLASCHKE 2010). Das bedeutet, dass die Aufnahme im Vorfeld in Gruppen von Pixeln aufgeteilt wird, die bis zu einem gewissen Grad homogen sind und von denen ausgegangen wird, dass sie zu einer Klasse gehören (Segmentierung).

Anschließend werden Statistiken für diese Segmente kalkuliert, die dann die Basis für die darauffolgende Klassifikation bilden. In diesem Fall werden Maximal-, Minimal- und Mittelwert sowie die Standardabweichung und der Median berechnet. Da es nur darum geht, die blütenreichen Flächen zu erfassen, handelt es sich um eine One Class Classification (OCC). Das bedeutet, dass es nur zwei Klassen gibt: „Blühfläche“ und „nicht-Blühfläche“. Für die Blühflächen wurden Referenzsegmente manuell ausgewählt, während für die Repräsentation der nicht-Blühflächen zufällig 20 Prozent der restlichen Segmente zum Training verwendet wurden. Die verbleibenden Segmente wurden dann nach dem Training klassifiziert. Als Klassifikationsmethode wurde Support Vector Machines (SVM) gewählt, welche auch bei einer geringen Menge an Referenzdaten anwendbar ist. Außerdem ist sie als nicht-parametrischer Algorithmus nicht auf Klassen mit einer Normalverteilung angewiesen. Als Datengrundlage dienten das RGB-Orthomosaik und das darauf basierende DOM, insgesamt somit vier Eingangskanäle. Letzteres wurde einbezogen, da rein visuell eine mögliche Schwierigkeit in der Unterscheidung von Bäumen und Blühflächen vermutet wurde.

Zum Testen dieses Ansatzes wurden zwei Testgebiete ausgewählt (Abb. 2). Das erste enthält eine artenreiche Blühfläche und das zweite enthält mehrere Flächen, in denen fast keine bzw. wenige Blüten vorhanden sind. Die räumliche Verkleinerung ist auch durch die ressourcenintensiven Aufgaben (Segmentierung/Statistikberechnung) begründet.

Für die Segmentierung wurden die drei Bänder aus dem RGB-Datensatz genommen. Neben diesen und dem DOM wurden noch Texturparameter extrahiert, die ebenfalls in die Klassifikation eingeflossen sind. Für das Generieren von Texturbändern wurde die Grey Level Co-Occurrence (GLCM) Matrix verwendet. Mit dieser Methode können verschiedene Größen, welche die Textur auf unterschiedliche Weisen charakterisieren, ausgegeben werden. Da die möglichen Ergebnisbänder teilweise korreliert sind, wurden lediglich folgende fünf Texturparameter verwendet: Kontrast, Entropie, Mittelwert, Varianz und Korrelation. Diese wurden auf jedes Eingangsband angewendet und ebenso die Statistiken dafür berechnet. Somit ergab sich eine Anzahl von 120 Werten pro Segment (4(Kanäle)*5(Statistik)+4(Kanäle)*5(Statistik)*5(Textur)).

Das Vorgehen besteht demnach im Wesentlichen aus drei Schritten: Segmentierung, GLCM-Berechnung und Klassifikation.

Abb. 3: Gegenüberstellung eines sehr blütenarmen (Blueh1a_P3, oben), wenig blütenreichen (Blueh1a_P6, Mitte) und sehr blütenreichen (Blueh6a, unten) Gebietes. Zu sehen ist ein RGB-Ausschnitt mit 11 cm (links) und 5,4 cm (Mitte) Auflösung im Vergleich zu einem vor Ort aufgenommenen Foto (rechts).

3.2 RGB-Farbraumanalyse

Es liegt die Vermutung nahe, dass sich Blüten aufgrund der bunten -nicht grünen- Blütenfarben in den sowohl radiometrisch als auch räumlich hochaufgelösten RGB-Luftbildern (als nicht grüne Pixel) erkennen und durch eine RGB-Farbraumanalyse bzw. Darstellung der pixelweisen Farbhäufigkeiten einschätzen lassen. Die Variationen des Blütenreichtums verschiedener Untersuchungsgebiete sind mit bloßem Auge in den Daten sichtbar (Abb. 3). Aus den Blühflächen wurden ein oder mehrere gleichgroße Ausschnitte für die weitere Untersuchung ausgewählt. Es stellte sich heraus, dass die Region Blueh6b stark überbelichtet ist, so dass diese Flächen von der weiteren Untersuchung ausgeschlossen wurden. Somit verblieben 27 Flächen für die weitere Analyse.

Mit möglichst wenig Aufwand (Rechenzeit, Datenvorbereitung, …) wurden zwei Möglichkeiten untersucht, über einen Algorithmus sowohl Blütenmenge als auch die überwiegende Blütenfarbe zu ermitteln.

Ein erster Ansatz beruht auf der Idee, sich den RGB-Farbraum als dreidimensionales Koordinatensystem vorzustellen, bei dem die drei Achsen durch die drei Kanäle rot (R), grün (G) und blau (B) gebildet werden. Diese Vorstellung gleicht einem Farbwürfel, bei dem der Nullpunkt dieses Systems schwarz ist, die Ecke mit den Maximalwerten weiß repräsentiert und die übrigen Eckpunkte die Farben rot, grün und blau (entlang der Koordinatenachsen) oder gelb, cyan und magenta abbilden (Abb. 4).

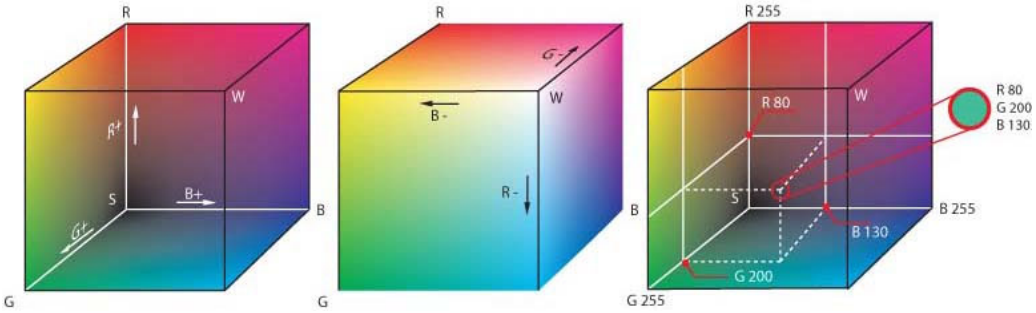


Abb. 4: Darstellung des RGB-Farbraumes als dreidimensionales Koordinatensystem [Bildquelle: https://upload.wikimedia.org/wikipedia/commons/0/03/RGB_farbwuerfel.jpg]

Es bestand die Annahme, dass blütenreiche Gebiete auffällig viele Punkte im entsprechenden Farbbereich haben. Für diese Untersuchung wurden für jeden der 27 Ausschnitte die Grauwerte pro Kanal jeweils in fünf Klassen (lineare Einteilung) zusammengefasst. Anschließend wurde für jede Klasseneinteilung die Anzahl der hierin vorkommenden Pixel ermittelt. Die Ergebnisse bestehen aus drei Diagrammen pro Ausschnitt, die für die verschiedenen Blickrichtungen die Verteilung und Häufigkeit der Pixelfarben je nach Farbkanal aufzeigen, um zu prüfen wie sich die Farbräume blütenreicher und blütenarmer Gebiete unterscheiden (Abb. 7 bis 9, S. 229f). Dazu wurde die freie Software R verwendet.

Da die Farbräume allerdings nicht der oben aufgeführten linearen Einteilung entsprechen, wurden die Farbräume der in den Pixeln erkennbaren Blütenfarben (rot, weiß, gelb, helllila, hellgelb) aus dem RGB-Orthomosaik abgegriffen (Abb. 5). Vereinzelte blaue Blüten waren in den Pixeln nicht erkennbar.

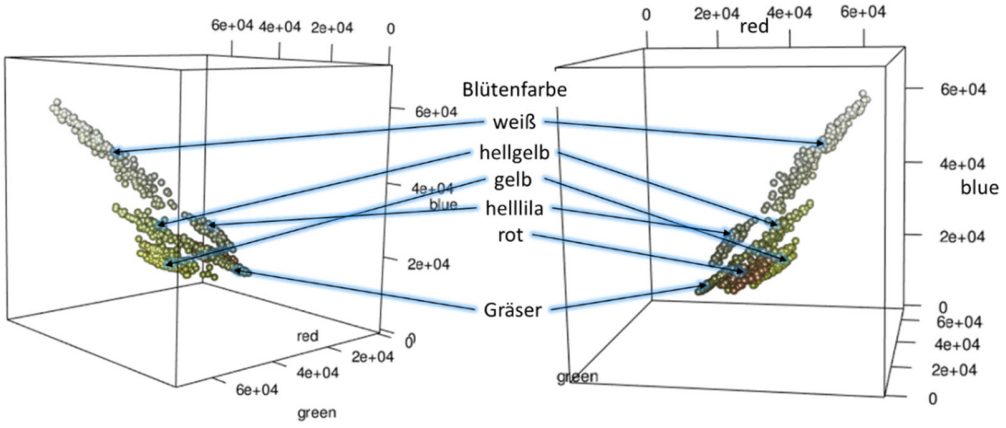


Abb. 5: R-Plot mit Punkthaufen der vorkommenden Blütenfarben und Gräser im 3D-Farbraum. Die Punkthaufen der einzelnen Blütenfarben haben eine ellipsoidische Form.

In einem zweiten Ansatz wurde für die einzelnen 27 Untersuchungsflächen geprüft, welche Pixel innerhalb eines Farbraums (Mittelwert ± einfache Standardabweichung) liegen. Ermittelt wurden sowohl die Anzahl für alle Farbräume als auch die prozentuale Blütenabdeckung (Verhältnis aus Anzahl der nicht grünen Farbpixel und Gesamtpixelzahl der jeweiligen Untersuchungsfläche). Dies wurde ebenso für den Farbraum der Gräser durchgeführt, um später zu prüfen, ob sich der Gräserreichtum auf die Erfassung von Blütenfarbe und -menge auswirkt.
Die Blütenmenge entspricht der prozentualen Blütendeckung und wurde mit den vor Ort erfassten Werten des Blütenreichtums verglichen. Die vorherrschende Blütenfarbe wurde als richtig gewertet, wenn die meisten Pixel diesem Farbraum zugeordnet werden konnten. Bei zwei vorherrschenden Blütenfarben wurden die beiden Farbräume mit den meisten Pixeln zur Einschätzung herangezogen.

4 Ergebnisse und Diskussion

4.1 Klassifikation blütenreicher Flächen

Die Resultate für die Segmentierung und die Klassifikation sind in Abb. 6 dargestellt. Die Segmentierung hat sehr gute Ergebnisse gebracht, die zur weiteren Bearbeitung genutzt wurden. Defizite gibt es bei Blueh1a in Bezug auf die Erfassung der kleinen Trennungsstücke zwischen den Parzellen. Bei Blueh6a ist nur ein kleiner Bereich am nordöstlichen Rand problematisch, in dem ein Teil der Blühfläche dem Segment des Nachbarschlages zugeteilt wurde. Zur überwachten Klassifikation wurden drei Trainingssegmente aus Blueh6a gewählt, mit dem Ziel die leicht unterschiedlichen Muster aufzugreifen. Eine gleiche Herangehensweise wurde bei den fünf gewählten Segmenten in Testgebiet 2 gewählt.
Beide Klassifikationsergebnisse zeigen eine hohe Overall Accuracy (Testgebiet 1: 0.997, Testgebiet 2: 0.984), unterscheiden sich jedoch in der Quote der richtig erkannten Blühflächen-Segmente, der Producers'Accuracy, (Testgebiet 1: 0.917, Testgebiet 2: 0.571). Das bedeutet, die hohen Zahlen für die Overall Accuracy resultieren, insbesondere für Testgebiet 2, von den richtig erkannten nicht-Blühflächen-Segmenten (Testgebiet 1: 1.0, Testgebiet 2: 0.996). Diese sind in sehr

viel höherer Zahl vorhanden und beeinflussen deshalb die Overall Accuracy positiv. Das deutet möglicherweise auf Overfitting des angewendeten Modells, eine zu starke Anpassung an die Trainingsdaten, hin. Davon unabhängig legt die Tatsache, dass nur ein Segment in Testgebiet 1 falsch klassifiziert wurde, den Verdacht nahe, dass ein hoher Blütenreichtum das Alleinstellungsmerkmal für Blühflächen ist. Eine weitere Schlussfolgerung ist, dass die Klassifizierung von nicht blütenreichen Blühflächen anfälliger für Verwechselungen ist. Ursachen für solche Uneindeutigkeiten können Grasflächen oder Teile von Feldern sein, die ähnliche spektrale Eigenschaften aufweisen aber nicht dieselben Funktionen erfüllen können.

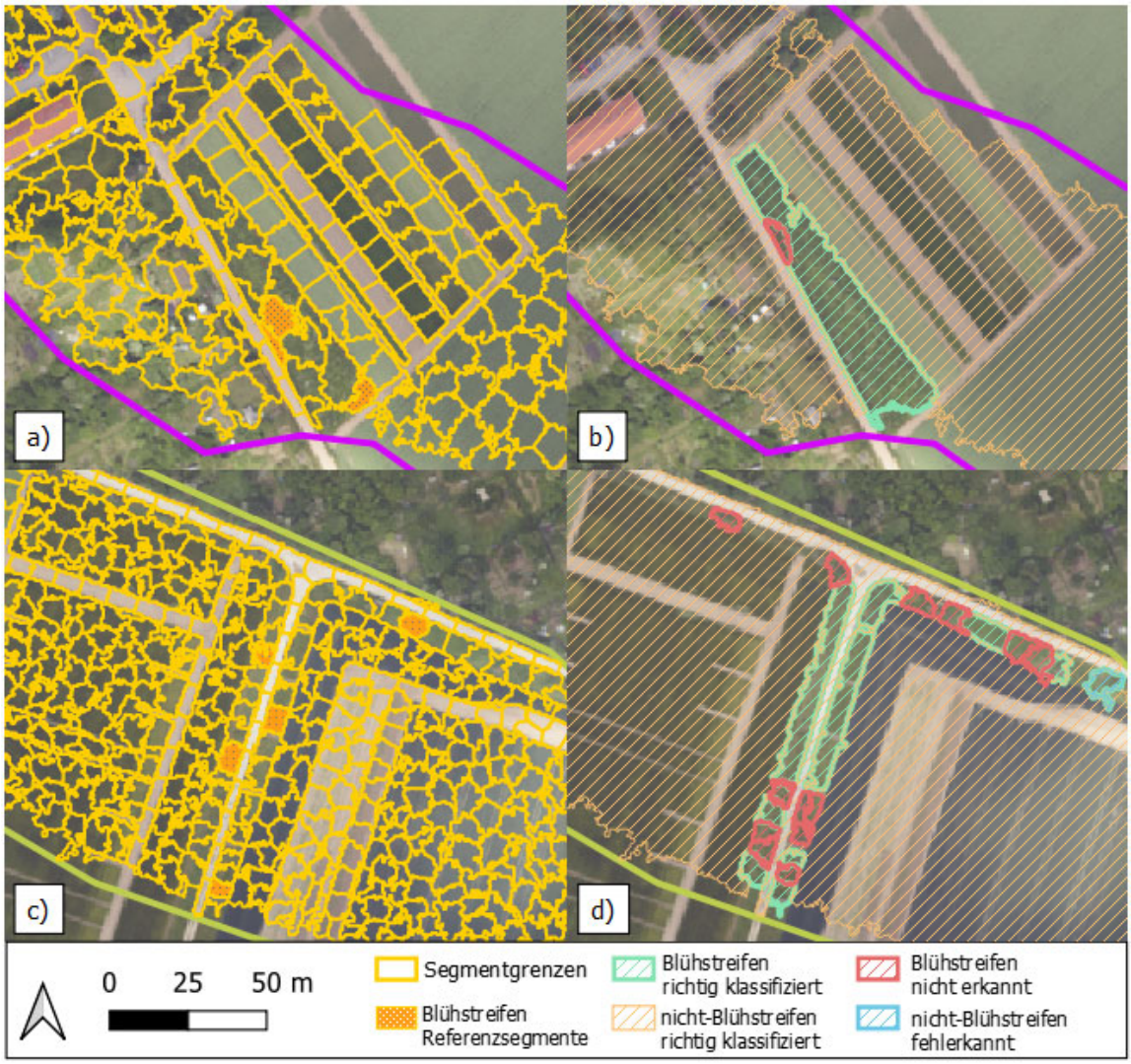


Abb. 6: Ergebnisse der Segmentierung und Klassifikation. Segmentierungsergebnis und die daraufhin ausgewählten Trainingsgebiete für die Blühflächen der Testgebiete a) 1 und c) 2 mit den dazugehörigen Klassifikationsergebnissen b) und d). In b) und d) sind die Teile der Blühflächen, die richtig klassifiziert wurden, grün markiert, während rot für die nicht erkannten Segmente steht. Hellorange sind die richtig erkannten nicht-Blühflächen. Segmente, die fälschlicherweise als nicht-Blühflächen klassifiziert wurden, sind blau dargestellt.

4.2 RGB-Farbraumanalyse

Mittels einer RGB-Farbraumanalyse wurden der Blütenreichtum, der Einfluss des Gräserreichtums auf den berechneten Blütenreichtum, die Erkennung der vorherrschenden Blühfarbe sowie der Artenreichtum untersucht.

4.2.1 Blütenreichtum

Um den Blütenreichtum abzuschätzen, wurde zunächst die Verteilung der RGB-Farbwerte (11 cm Auflösung) in einem dreidimensionalen Farbraum betrachtet. Pro Ausschnitt wurden drei Diagramme generiert, die die Häufigkeit und Verteilung der Pixelfarben darstellen. Dabei sind die Grauwerte der Kanäle jeweils in fünf Klassen zusammengefasst (1 entspricht Grauwerten von 0 bis 13106, 2 entspricht Grauwerten von 13107 bis 26213, usw.). In den Abb. 7 bis 9 sind die Diagramme für Ausschnitte mit verschieden hohem Blütenreichtum und verschiedenen Blühfarben dargestellt, „band1“ entspricht dabei dem roten Kanal, „band2“ dem grünen und „band3“ dem blauen Kanal. Bei dem blütenarmen Gebiet Blueh1a_P3 liegt die Mehrzahl der Pixel bei {3,3,2} (Abb. 8), dies lässt sich jedoch nicht für alle blütenarmen Gebiete bestätigen.
Bei Blühfarben im weiß-hellgelben Bereich gilt: Je blütenreicher, desto mehr sind Pixel mit hohen Datenwerten für die drei Bänder vertreten. Bei den helllila Blüten lässt sich dieser Trend allerdings nicht feststellen, jedoch sind hier auch deutlich mehr Pixel in anderen Farbklassen als bei blütenarmen Gebieten vorhanden. Allgemeingültige Aussagen zum Blütenreichtum und zu den vorherrschenden Blühfarben lassen sich anhand der Diagramme nur insofern treffen, als dass bei höherem Blütenreichtum eine stärkere Streuung der Werte entlang der Hauptdiagonale zu erkennen ist.

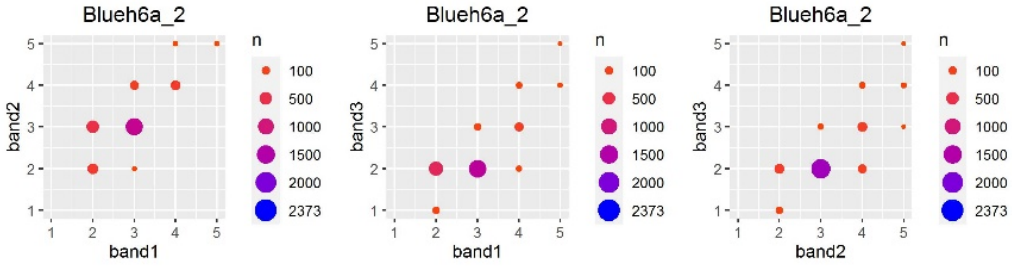


Abb. 7: Beispiel für eine RGB-Farbraumanalyse eines sehr blütenreichen Gebietes (Blueh6a_2, BRT=5, GRT=1, ART=4, vorwiegende Blühfarbe=hellgelb). Band1, band2 und band3 steht für die Farbkanäle rot, grün und blau, n gibt die Anzahl der Pixel dieses Ausschnittes an.

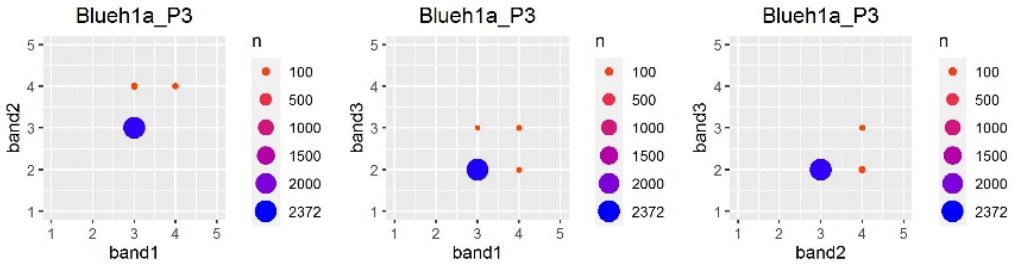


Abb. 8: Beispiel für eine RGB-Farbraumanalyse eines sehr blütenarmen Gebietes (Blueh1a_P3, BRT=1, GRT=3, ART=2, vorwiegende Blühfarbe=hellgelb). Band1, band2 und band3 steht für die Farbkanäle rot, grün und blau, n gibt die Anzahl der Pixel dieses Ausschnittes an.

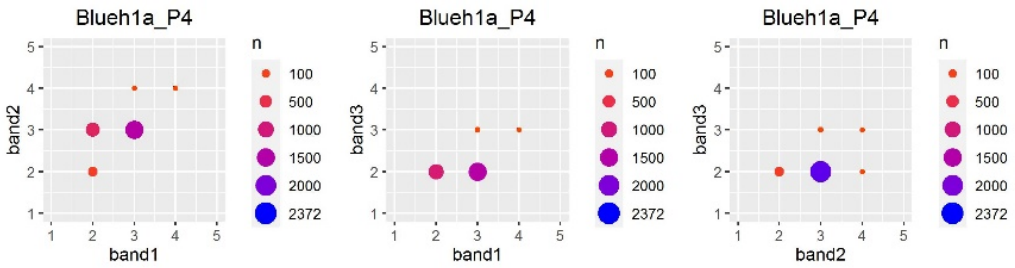


Abb. 9: Beispiel für eine RGB-Farbraumanalyse eines blütenreichen Gebietes (Blueh1a_P4, BRT=3, GRT=3, ART=2, vorwiegende Blühfarbe=helllila). Band1, band2 und band3 steht für die Farbkanäle rot, grün und blau, n gibt die Anzahl der Pixel dieses Ausschnittes an.

Entsprechend des zweiten Ansatzes wurden sowohl die Anzahl der Pixel für die vorherrschenden Farbräume (rot, weiß, gelb, helllila, hellgelb) als auch die prozentuale Blütendeckung (Verhältnis aus Anzahl der Blühfarbenpixel und Gesamtpixelzahl der jeweiligen Untersuchungsfläche) ermittelt.
Die berechnete prozentuale Blütenabdeckung und der vor Ort geschätzte Blütenreichtum wurden gegenübergestellt. Ein Vergleich dieser Gegenüberstellung für verschiedene räumliche Auflösungen soll bei der Planung und Optimierung zukünftiger Befliegungen helfen. Eine niedrigere räumliche Auflösung wäre leichter zu realisieren (kürzere Befliegungsdauer, da höhere Gebietsabdeckung pro Überflug) und mit weniger Kosten verbunden.
Unabhängig von der Bodenpixelauflösung (Groud Sampling Distance, GSD) ist der gleiche Trend zu erkennen (Vergleiche beide Diagramme in Abb. 10).

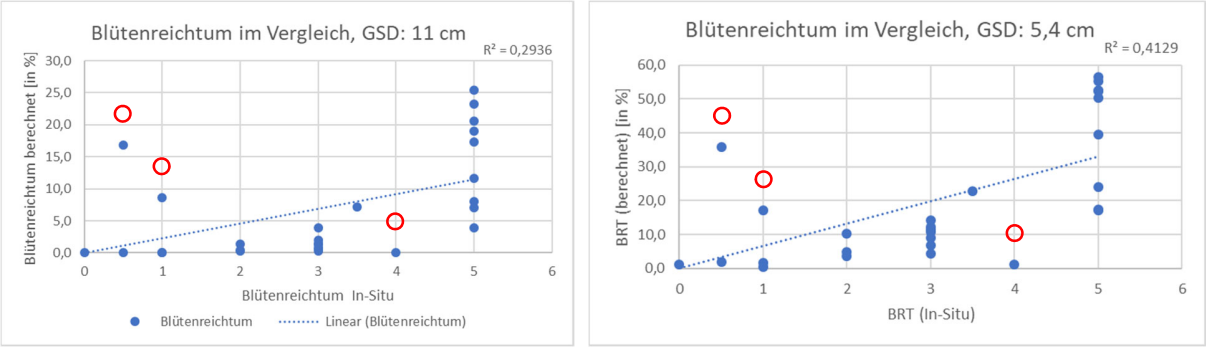


Abb. 10: Prozentual berechneter Blütenreichtum im Vergleich zu dem vor Ort geschätzten Blütenreichtum bei einer GSD von 11 cm (links) und 5,4 cm (rechts). Zwischen dem berechneten und dem In-Situ Blütenreichtum wurde ein linearer Zusammenhang angenommen. Die Regressionskoeffizienten R^2=0.29 (links) und R^2=0.41 (rechts) werden stark durch drei Ausreißer (rote Markierungen) und die hohen Schwankungen bei den als sehr blütenreich (5) eingestuften Gebieten beeinflusst.

Bei genauerer Betrachtung fällt auf, dass sich der für die einzelnen Gebiete berechnete Blütenreichtum je nach Bodenpixelauflösung stark voneinander unterscheidet. So sind die Werte bei einer höheren GSD (hier 5,4 cm) etwa doppelt so hoch wie bei einer geringeren Auflösung von 11 cm. Eine mögliche Ursache dafür wird in der Bildung von Mischpixeln vermutet. Das wiederum deutet darauf hin, dass die Bodenpixelauflösung der Luftaufnahme bei der Erfassung

der kleinen Blüten einen deutlich größeren Einfluss hat, um quantitative Aussagen bezüglich der Blütenmenge treffen zu können. Eine genauere Untersuchung der prozentualen Blütenmenge anhand von Referenzdaten, bei der die prozentuale Blütenabdeckung erfasst wurde, soll zukünftig durchgeführt werden.
Außerdem fällt bei einer Gegenüberstellung der vor Ort geschätzten Blütenmenge und der prozentual ermittelten Blütenmenge aller Ausschnitte auf, dass es besonders in den als sehr blütenreich (5) eingestuften Gebieten hohe Variationen der berechneten Blütenmenge gibt. Die Ursache hierfür wird darin gesehen, dass bei einer sehr hohen Blütenmenge das subjektive Empfinden die Schätzwerte beeinflusst und nicht weiter unterschieden werden kann zwischen *sehr vielen* und *noch viel mehr* Blüten. Ein Lösungsansatz zum Vermeiden dieses Problems ist, den berechneten prozentualen Blütenreichtum in Klassen zusammenzufassen (Abb. 11). Daher wurde die berechnete prozentuale Blütendeckung ebenfalls in sechs Klassen eingeteilt (0 keine Blüten, 5 sehr hoher Blütenreichtum). Die Stufen orientieren sich dabei an der für Vegetationsaufnahmen gebräuchlichen LONDO-Skala (TRAXLER 1997). Abhängig von der räumlichen Auflösung musste die Skala allerdings leicht angepasst werden.

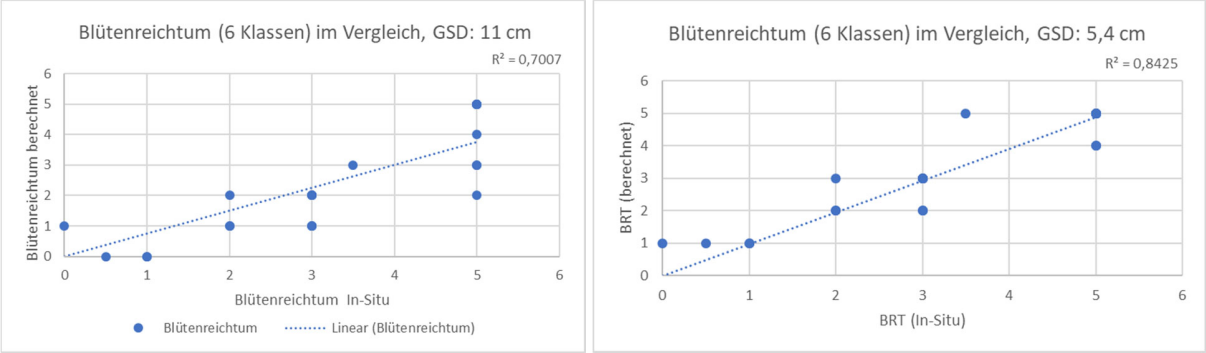


Abb. 11: Gegenüberstellung des in Klassen zusammengefassten berechneten Blütenreichtums im Vergleich zu dem vor Ort geschätzten Blütenreichtum je nach GSD ohne die drei stärksten Ausreißer: Blueh1a_P3, Blueh1a_P8, Blueh1b_P2. Links für 11 cm Auflösung (Klasseneinteilung: 0: 0, 1: Werte > 0 und <= 1 %, 2: Werte > 1 % und <= 5 %, 3: Werte > 5 % und <= 10 %, 4: Werte > 10 % und <= 15 %, 5: Werte > 15 %); Rechts 5,4 cm Auflösung (Klasseneinteilung: 0: 0, 1: Werte > 0 und <= 3 %, 2: Werte > 3 % und <= 8 %, 3: Werte > 8 % und <= 15 %, 4: Werte > 15 % und <= 20 %, 5: Werte > 20 %).

Für die beiden untersuchten Auflösungen sehen die Ergebnisse ähnlich aus, wobei das Ergebnis bei einer GSD von 5,4 cm mit einem Regressionskoeffizienten von 84 % im Vergleich zu dem Luftbild mit 11 cm GSD (mit R^2=70 %) noch besser ist. Bei den Diagrammen in Abb. 11 wurden die drei stärksten Ausreißer (Blueh1a_P3, Blueh1a_P8 und Blueh1b_P2) nicht mit berücksichtigt. Die Ursache für diese Abweichungen bleibt ungeklärt.

4.2.2 Einfluss des Gräserreichtums auf den Blütenreichtum

Der Verdacht, dass sich ein hoher Gräserreichtum aufgrund der helllila bis gelblichen Farbe der Gräser positiv auf den berechneten Blütenreichtum auswirkt (durch fehlerhafte Zuordnung der Pixel), konnte widerlegt werden. Zwar liegen die Farbräume der Klassen nah beieinander, aber bei einem hohen Gräserreichtum steigt weder die Zahl der hellgelben oder helllila Pixel noch der pro-

zentuale berechnete Blütenreichtum (Abb. 12). Allerdings haben die untersuchten Flächen mit hohem Blütenreichtum auch nur einen geringen Gräserreichtum und die Flächen mit hohem Gräserreichtum nur einen geringen Blütenreichtum.

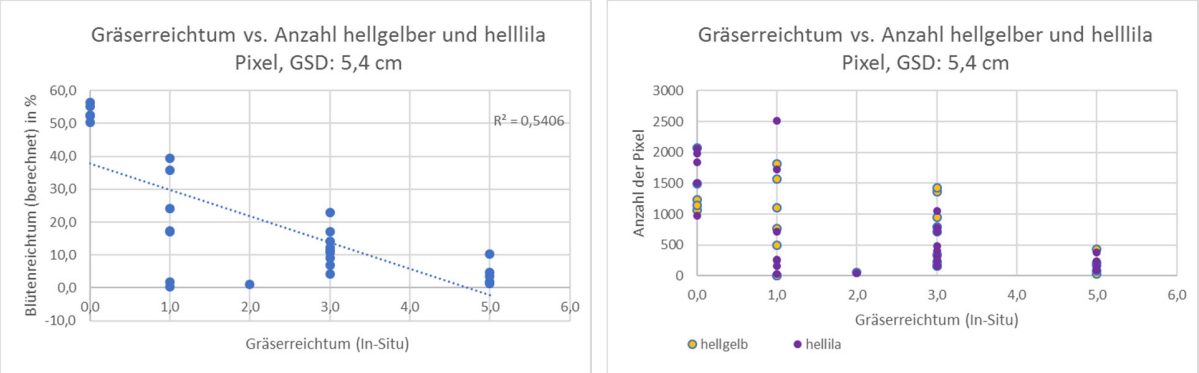


Abb. 12: Bei einem hohen Gräserreichtum steigt weder der prozentuale berechnete Blütenreichtum (links) noch die Zahl der hellgelben oder helllila Pixel (rechts). Dies geht mit der Beobachtung vor Ort einher, dass bei hohem Gräserreichtum weniger Blüten vorhanden waren.

4.2.3 Erkennung der vorherrschenden Blütenfarbe

Als vorherrschende Blütenfarben traten weiß, gelb, helllila und hellgelb sowie Kombinationen davon auf. Die über den vorgestellten Ansatz ermittelte vorherrschende Blütenfarbe wurde als richtig gewertet, wenn die meisten Pixel diesem Farbraum zugeordnet werden konnten. Bei zwei vorherrschenden Blütenfarben wurden die beiden Farbräume mit den meisten Pixeln zur Einschätzung herangezogen (daher ist hier auch die Gesamtzahl höher als die Zahl der Untersuchungsflächen). Die Ergebnisse ließen sich mit Hilfe einer Konfusionsmatrix auswerten.
Für beide untersuchten Bodenpixelauflösungen geben die Konfusionsmatritzen die gleiche Overall Accuracy von 57 % an. Auch die Zahl der richtig erkannten vorwiegenden Blühfarben (20) und die der falsch erkannten (15) ist bei beiden Bodenpixelauflösungen gleich hoch.
Bei genauerer Betrachtung kann festgestellt werden, dass bei einer GSD von 11 cm die nicht blühenden Gebiete besser erkannt wurden. Allerdings kam es zu häufigen Verwechselungen bei den vorwiegend hellgelb blühenden Flächen. Bei einer GSD von 5,4 cm haben hingegen viele Gebiete die meisten helllila-farbenen Pixel, obwohl die vorwiegende Blühfarbe eine andere ist. Bei den beiden verschieden hohen Bodenpixelauflösungen wurden vorwiegend gelbblühende Flächen nicht als solche erkannt. Dies deutet darauf hin, dass die Erkennung der vorwiegenden Blühfarbe(n) über diese Methode nur bedingt praktikabel ist.
Mögliche Ursachen bilden zum einen die Ähnlichkeit der Farbräume (z. B. weiß und helllila, gelb und hellgelb) und zum anderen, dass es bei der Bestimmung der vorwiegenden Blühfarbe mitunter nur wenige Pixel Unterschied gab, was zu einer fehlerhaften Zuordnung führte.
Zudem hat die Genauigkeitseinschätzung der vorwiegenden Blühfarbe über die genannte Methode mit einer Konfusionsmatrix den Nachteil, dass bei einer geringen Anzahl der Untersuchungsgebiete Fehlklassifizierungen die Gesamtgenauigkeit stark beeinträchtigen.
Eine Erhöhung des Stichprobenumfangs durch Einbeziehung weiterer Befliegungsdaten zu anderen Zeitpunkten ist geplant. Ebenso steht die Prüfung der gewählten Methode zur Diskussion.

Als Alternative könnte beispielsweise die Zuordnung der Pixel in die entsprechende Farbklasse über einen anderen Klassifikationsalgorithmus, der die ellipsoidische Form der einzelnen Farbräume berücksichtigt, untersucht werden.

4.2.4 Erkennung der vorherrschenden Blütenfarbe

Der vor Ort erfasste Artenreichtum wurde in Bezug zu der Anzahl der anhand der RGB-Luftbilder ermittelten vorkommenden Blühfarben gesetzt. Hierbei muss erwähnt werden, dass bei der Auflösung von 11 cm die vorkommenden Blütenfarben gezählt wurden, sobald mindestens drei Pixel darin erkannt wurden. Bei der Auflösung von 5,4 cm GSD wurde die Farbe erst gezählt, wenn mehr als zehn Pixel der jeweiligen Farbklasse erkannt wurden.
Bei dem Vergleich zwischen Artenreichtum und des anhand der RGB-Luftbilder ermittelten Farbreichtums ist insbesondere bei einer GSD von 11 cm ein klarer proportionaler Zusammenhang erkennbar (Abb. 13). Bei dem Punkt (3,5|0) handelt es sich um einen der oben bereits benannten Ausreißer Blueh1a_P8. Der Artenreichtum kann grob über die Anzahl der vorkommenden Blühfarben abgeschätzt werden, eine genaue Zahl der vorkommenden Arten kann über eine Farbraumanalyse jedoch nicht ermittelt werden, da verschiedene Arten die gleiche Blütenfarbe haben. Außerdem wurden in dieser Untersuchung nur die blühenden Arten betrachtet.

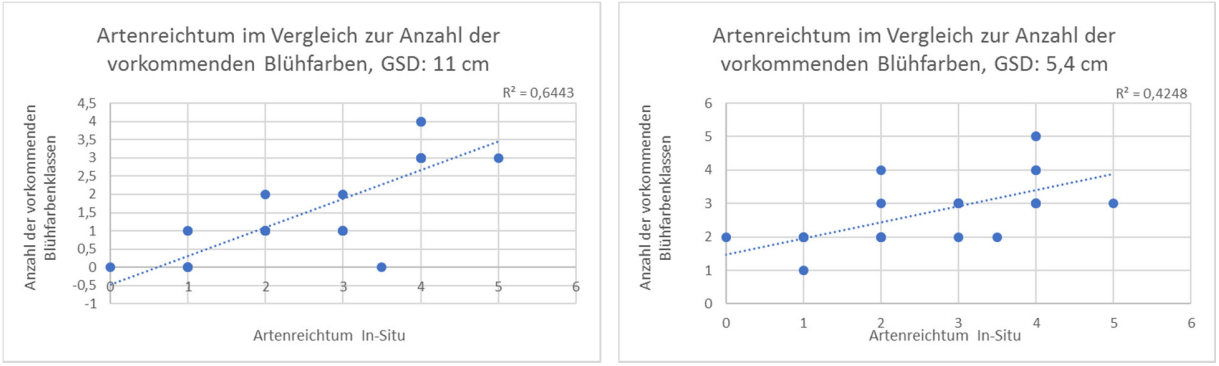


Abb. 13: Vergleich zwischen dem vor Ort erfassten Artenreichtum und der Anzahl der ermittelten vorkommenden Blütenfarben mit einer GSD von 11 cm (links) und einer GSD von 5,4 cm (rechts).

5 Fazit und Ausblick

Zusammenfassend lässt sich feststellen, dass im Rahmen dieser Erarbeitung die vorausgegangenen Fragestellungen beantwortet werden konnten und Potential für weiterführende Untersuchungen aufweisen.
Es konnte ein Algorithmus entwickelt werden, mit dem die Identifikation von Blühflächen anhand weniger bekannter Gebiete möglich ist.
Ebenso wurde gezeigt, dass eine qualitative Einschätzung des Blütenreichtums möglich ist, wobei insbesondere nach der Eliminierung von Ausreißern und des in Klassen zusammengefassten berechneten prozentualen Blütenreichtums gute Ergebnisse vorliegen (Overall Accuracy > 70 %).
Eine quantitative Abschätzung der Blütenmenge ist stark abhängig von der Bodenpixelauflösung, da bei einer höheren räumlichen Auflösung deutlich mehr farbige Blütenpixel erkannt werden, was

bei der Berechnung zu einer höheren prozentualen Abdeckung des Blütenreichtums führt. Der Artenreichtum blühender Pflanzen korreliert mit der Anzahl der vorkommenden Blühfarben und kann darüber grob abgeschätzt werden. Ein hoher Gräserreichtum wirkt sich nicht negativ auf die prozentual ermittelte Blütenmenge aus. Bei der Einschätzung der vorwiegenden Blühfarbe besteht noch Optimierungsbedarf. So kann die Farbeinschätzung z. B. mit Hilfe eines Klassifikationsverfahrens durchgeführt werden.
Zur besseren Ergebnisvalidierung sollten mehr Datensätze z. B. von weiteren Zeitpunkten und Gebieten einbezogen werden. Diese Gebiete sollten mehrere ähnliche Blühflächen aufweisen, um die Übertragbarkeit besser abschätzen zu können, und damit die Trainings- und Testsegmente nicht vom gleichen Blühstreifen stammen (weniger Korrelation). Die zu verschiedenen Zeitpunkten und Regionen erfassten Fernerkundungsdaten unterliegen allerdings verschiedenen Beleuchtungsbedingungen, die das Ergebnis beeinflussen. Daher sollte bei einem Vergleich zukünftigen Befliegungsdaten eine Atmosphärenkorrektur vorgenommen werden oder die Verwendung von Ratios (Bildung von Kanalverhältnissen, um z. B. verschiedene Beleuchtungsbedingungen zu kompensieren) bevorzugt werden.
Die erforderliche räumliche Bodenpixelauflösung richtet sich nach dem jeweiligen Anwendungsfall: Oft ähneln sich die Ergebnisse und weisen den gleichen Trend auf, haben allerdings verschiedene Genauigkeiten. In weiteren Testreihen sollte die prozentuale Blütendeckung mit Vor-Ort-Daten verglichen werden, um so auch spezifischere Aussagen bezüglich der quantitativen Einschätzung der Blütenmenge treffen zu können.
Bezogen auf die Klassifikation von blütenreichen Flächen sollte mehr Forschung betrieben werden. Mit einer Unterteilung von Blühstreifen in Subklassen, deren Charakteristiken weniger weitläufig sind, könnte man Klassifikatoren besser trainieren und das Potential von Fehlklassifizierungen verringern. Um Overfitting zu reduzieren, könnte ein anderer Klassifikator herangezogen werden, der mit dem Input von einer positiven und einer negativen Klasse, die aus mehreren sehr unterschiedlichen Subklassen besteht, besser umgehen kann. Auch sollten die Hyperspektraldaten näher untersucht werden: welche Bänder sind inwiefern ausschlaggebend und auf welche Bänder könnte verzichtet werden, ohne Genauigkeit einzubüßen.
Es ist vorstellbar, dass der entwickelte Klassifikationsalgorithmus auf Satellitendaten übertragen werden könnte (ggf. mit Anpassungen), um die globale Verteilung der blütenreichen Flächen zu erkennen. Desweiteren sollen zusätzliche Fernerkundungsdaten, die zu anderen Vegetationszeitpunkten erhoben wurden, auf zeitliche Veränderungen des Blühverhaltens untersucht werden.

6 Literaturverzeichnis

BANNEHR, L., GABLER, M. & PROKOPH, S., 2021: Forschungsgyrocopter der Hochschule Anhalt - Sensorik. Sensorik. Institut für Geoinformation und Vermessung. https://gyrocopter.afg.hs-anhalt.de/index.php/sensorik, letzter Zugriff 13.01.2022.

BLASCHKE, T., 2010: Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, **65**(1), 2-16, https://doi.org/10.1016/j.isprsjprs.2009.06.004.

FEILHAUER, H., DOKTOR, D., SCHMIDTLEIN, S. & SKIDMORE, A. K., 2016: Mapping pollination types with remote sensing. J Veg Sci, **27**(5), 999-1011. https://doi.org/10.1111/jvs.12421.

FENCHEL, J., BUSSE, A., REICHARDT, I., ANKLAM, R., SCHRÖDTER, M., TISCHEW, S., MANN, S. & KIRMER, A., 2015: Hinweise zur erfolgreichen Anlage und Pflege mehrjähriger Blühstreifen und Blühflächen mit gebietseigenen Wildarten (mit Hinweisen zu einjährigen Blühstreifen und Blühflächen sowie Schonstreifen). MLU Sachsen-Anhalt (Hrsg.), Magdeburg.

HALLMANN, C. A., SORG, M., JONGEJANS, E., SIEPEL, H., HOFLAND, N. & SCHWAN, H., 2017: More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS ONE, **12**(10), e0185809. https://doi.org/10.1371/journal.pone.0185809.

KIRMER, A., RYDGREN, K. & TISCHEW, S., 2018: Smart management is key for successful diversification of field margins in highly productive farmland. Agriculture, Ecosystems & Environment, **251**, 88-98. https://doi.org/10.1016/j.agee.2017.09.028.

KIRMER, A., JESCHKE, D. KIEHL, K. & TISCHEW, S., 2019: Praxisleitfaden zur Etablierung und Aufwertung von Säumen und Feldrainen. 2. Auflage 2/2019. Bernburg: Hochschule Anhalt Hochschulbibliothek. https://www.offenlandinfo.de/themen/saeume-feldraine-und-bluehstreifen/, letzter Zugriff 13.01.2022.

LANDMANN, T., PIIROINEN, R., MAKORI, D. M., ABDEL-RAHMAN, E. M., MAKAU, S., PELLIKKA, P. & RAINA, S. K., 2015: Application of hyperspectral remote sensing for flower mapping in African savannas. Remote Sensing of Environment, **166**, 50-60. https://doi.org/10.1016/j.rse.2015.06.006.

LAUSCH, A., BANNEHR, L., BECKMANN, M., BOEHM, C., FEILHAUER, H., HACKER, J. M., HEURICH, M., JUNG, A., KLENKE, R., NEUMANN, C., PAUSE, M., ROCCHINI, D., SCHAEPMAN, M.E., SCHMIDTLEIN, S., SCHULZ, K., SELSAM, P., SETTELE, J., SKIDMORE, A.K. & CORD, A.F., 2016: Linking Earth Observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. Ecological Indicators, **70**, 317-339. https://doi.org/10.1016/j.ecolind.2016.06.022.

MÜLLEROVÁ, J., PERGL, J. & PYŠEK, P., 2013: Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species Heracleum mantegazzianum (giant hogweed). International Journal of Applied Earth Observation and Geoinformation, **25**, 55-65. https://doi.org/10.1016/j.jag.2013.03.004.

NABU – NATURSCHUTZBUND DEUTSCHLAND E.V., 2022: Intensive Landwirtschaft ist Hauptgrund für Insektensterben - NABU. https://www.nabu.de/natur-und-landschaft/landnutzung/landwirtschaft/artenvielfalt/insektensterben/index.html, letzter Zugriff 13.01.2022.

PAUSE, M., HAHN, A. & LAUSCH, A., 2020: Verbesserung der Informationslage zur Flächennutzung mittels aktueller und zukünftiger Fernerkundungsdaten. 12. Dresdner Flächennutzungssymposium, https://doi.org/10.26084/12dfns-p024.

RADA, S., SCHWEIGER, O., HARPKE, A., KÜHN, E., KURAS, T., SETTELE, J. & MUSCHE, M., 2019: Protected areas do not mitigate biodiversity declines: A case study on butterflies. Divers Distrib, **25**(2), S. 217-224. https://doi.org/10.1111/ddi.12854.

SCHMIDT, A., KIRMER, A., HELLWIG, N., KIEHL, K. & TISCHEW, S., 2021: Evaluating CAP wildflower strips: High-quality seed mixtures significantly improve plant diversity and related pollen and nectar resources. Journal of Applied Ecology, 1-12, https://doi.org/10.1111/1365-2664.14102.

SCHMIDT, A., KIRMER, A., KIEHL, K. & TISCHEW, S., 2020: Seed mixture strongly affects species-richness and quality of perennial flower strips on fertile soil. Basic and Applied Ecology, **42**, 62-72, https://doi.org/10.1016/j.baae.2019.11.005.

SMIGAJ, M. & GAULTON, R., 2021: Capturing hedgerow structure and flowering abundance with UAV remote sensing. Remote Sens Ecol Conserv, **7**(3), 521-533, https://doi.org/10.1002/rse2.208.

SORG, M., SCHWAN, H., STENMANS, W. & MÜLLER, A., 2013: Ermittlung der Biomassen flugaktiver Insekten im Naturschutzgebiet Orbroicher Bruch mit Malaise Fallen in den Jahren 1989 und 2013. Mitteilungen aus dem Entomologischen Verein Krefeld, **1**, 1-5. https://www.boerenlandvogels.nl/content/ermittlung-der-biomassen-flugaktiver-insekten-im-naturschutzgebiet-orbroicher-bruch-mit-mala, letzter Zugriff 05.11.2021.

TRAXLER, A., 1997: Handbuch des vegetationsökologischen Monitorings. Teil A Methoden, Methoden, Praxis, angewandte Projekte. Wien: Umweltbundesamt (Monographien / Umweltbundesamt, 89A).

Digital 3D City Models Towards Urban Data Platform using OGC 3D GeoVolumes API

THUNYATHEP SANTHANAVANICH[1], PATRICK WUESRTLE[1], RUSHIKESH PADSALA[1] & VOLKER COORS[1]

Abstract: Nowadays, the digital 3D city models are the basis of the urban data platform. It plays an essential role in various fields of industry and research, from urban resource planning, environmental simulation, disaster management, and many more. Especially, there is more and more use of 3D city models in 3D visualization applications. However, one common issue is the accessibility difficulty according to different data formats generated from different data providers. Accordingly, there is a need for an interoperable 3D geospatial data delivery method to serve data in a standardized way. In this research, we introduce the use of the 3D GeoVolumes API, the open specification from Open Geospatial Consortium (OGC) to maximize interoperability, replicability, reusability, and accessibility of 3D geospatial data. The use cases had been implemented in the OGC Container and Tiles Pilot (2020) and OGC Interoperable Simulation and Gaming Sprint (2020 - 2021) using the 3D GeoVolumes API to deliver various 3D geospatial data formats. As a proof of concept, the 3D data from GeoVolumes API are visualized and showcased in game engines and web application clients.

1 Introduction

A 3D city model is a digital representation of a built environment and forms a fundamental building block for digital twins of cities. In a recent study, digital twins are seen as the ultimate tool for urban planners and city governments to design and build their infrastructure in a cost-effective way as well as share their vision with the general public (ABI RESEARCH 2021). The digital twins of cities is an accurate scale digital representation of city objects enabled to be paired with dynamic predictive models using simulations, data analytics, and machine learning to help decision-making by understanding cities' behavior in different "what if" scenarios. In the backend of the cities' digital twin is the urban data platform where the 3D city model, along with other urban datasets, are stored. Applications using 3D city models have already shown huge potential with respect to the built environment and motivated many public and private organizations to make use of it. For example, it has been used in several domains such as noise mapping (CZERWINSKI et al. 2007), the Helsinki 3D+ Digital Twins applications (AIRAKSINEN 2019), neighborhood planning (AGUGIERO et al. 2020), integrating dynamic sensor data with 3D city models (SANTHANAVANICH & COORS 2021), public participation method (WUERSTLE et al. 2021), urban infrastructure simulations (PADSALA et al. 2021), and many others. As a result, the demand for using 3D city models has increased drastically. It has become vital to create, analyze, store, visualize and share the 3D data for better planning and decision making. With the gap between desktop and web geo-visualization slowly minimizing, 3D visualization of geodata on the web increasingly has gained importance.

[1] Stuttgart University of Applied Sciences (HFT Stuttgart), Center for Geodesy and Geoinformatics, Schellingstraße 24, D-70174 Stuttgart, Germany,
E-Mail: [Thunyathep.Santhanavanich, Patrick.Wuerstle, Rushikesh.Padsala]@hft-stuttgart.de

One particular challenge in accessing, distributing, and visualizing 3D geodata on the web is in the data distribution method. As the need for web 3D geospatial visualization has risen, several solutions have been developed by different data vendors based on technical and internet bandwidth situations resulting in different data distributed methods. As a result, it has increasingly become challenging to access and integrate a variety of 3D content from different providers.
Thus, in this paper, we present the use of OGC 3D GeoVolumes, which seeks to address the above challenge by providing a corresponding Application Programming Interface (API) to integrate various data distribution methods into a single, open standard solution. The 3D GeoVolumes API defines a Web API enabling servers and clients to access 3D models and promotes interoperability providing the capabilities for 3D clients to access multiple servers and 3D formats such as 3D Tiles and CityGML. As a result, end-users and applications are able to search and access 3D contents using a common API, irrespective of its underlying distribution mechanism. The recent introduction of plugins to well know game engines like Unity and Unreal Engine now allow importing massive 3D city models into game engines to develop graphically rich and immersive next-generation interactive digital twin applications. Accessing, importing, and integrating 3D city models in such game engines or web applications with different underlying data distribution formats without data conversion represents a significant step forward in geodata interoperability.

2 Background

2.1 History of the 3D GeoVolumes API

The 3D GeoVolumes API concept had been first named and developed in the OGC 3D Container and Tiles pilot (MILLER et al. 2020a) as a candidate 3D delivery specification to address this issue by allowing software clients to access 3D geospatial contents from a single *HTTP* endpoint. Later, the 3D GeoVolumes had been successfully used as a data delivery specification in the OGC Interoperable Simulation and Gaming Sprint Year 1 (DALY & SERICH 2021) and Year 2 (DALY & PHILLIPS 2021). As a result, it shows the interoperability with multiple 3D data formats, integrations, and visualization.

2.2 3D GeoVolumes API Concept

The 3D GeoVolumes API is developed with the aim to conform with the OGC API-Common foundation resources, including landing page, conformance declaration, and collections. Being part of the OGC API family of standards, the 3D GeoVolumes API makes it easy for anyone to share, consume, filter the 3D geospatial resources through the web using the defined resource-centric APIs (HEAZEL 2020). Instead of accessing 3D data from different vendors, users can use 3D GeoVolumes API to manage the data heterogeneity and access data from a single source. The overview of the 3D GeoVolumes API resource path is shown in Tab. 1. In the GeoVolumes API, the collection is a group of spatial aggregation contents in the same regional level, e.g., city, country, continent, etc. Each collection is able to contain single to multiple children 3D containers. Each 3D container contains 3D resources in single to multiple 3D formats (see Chapter 3.2).

Tab. 1: The path and result from the 3D GeoVolumes API resources.

Resource	Path	Result
Landing page	/	Landing page in JSON or HTML.
Conformance declaration	/conformance	Conformance in JSON.
API definition	/api	API definition in JSON
Collection	/collections	All 3D collections in JSON/HTML.
Collection (with bbox)	/collections?bbox=[bbox]	Filtered 3D collections in JSON/HTML matching the bounding box [bbox].
3D container	/collections/[container_id]	3D Collections of [container_id] in JSON.
3D container (with bbox)	/collections/[container_id]?bbox=[bbox]	Filtered 3D collections of [container_id] in JSON matching the bounding box [bbox].
3D resources	/collections/[container_id]/[resource_format]	3D resources of [container_id] in the requested formats.

3 Use Cases

As a proof of concept, the 3D GeoVolumes API server had been implemented to serve the 3D city models through the web during our participation as part of the OGC 3D Container and Tiles pilot (MILLER et al. 2020a) and the OGC Interoperable Simulation and Gaming Sprint Year 1 (DALY & SERICH 2021) and Year 2 (DALY & PHILLIPS 2021). The overall structure of the development workflow for the 3D data management (Fig. 1) includes three main sections. Firstly, the 3D data sources in several formats were collected, integrated, and processed in the 3D data conversion pipeline (see Chapter 3.1). Secondly, these converted 3D data were served on the 3D GeoVolumes API server (see Chapter 3.2). Then, these data were evaluated on the web-, mobile-, and game engine applications (see Chapter 3.3). Moreover, we have integrated sensor data into the 3D building models to show the interoperability of the 3D GeoVolumes API with other OGC standards. The sensor data were integrated using the OGC SensorThings API standard based on the CityThings concept (SANTHANAVANICH & COORS 2021).

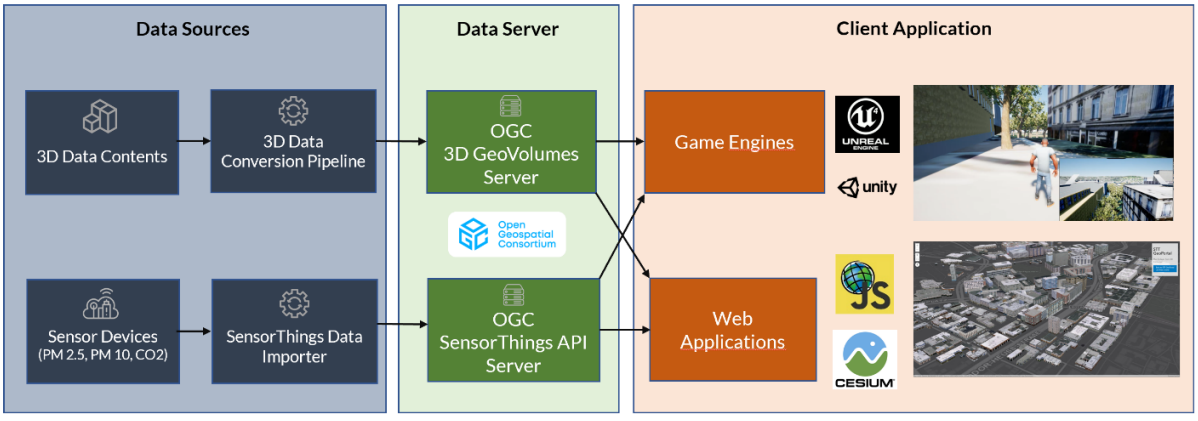


Fig. 1: Overall development workflow of the 3D data management.

3.1 3D Data Formats and Data Processing

Several 3D geospatial formats (Tab. 2) had been explored and processed to support 1) 3D visualizing in the web clients and game engines and 2) the tiling scheme to improve the rendering

and streaming performance. The data conversion processes were done in the Feature Manipulation Engine (FME) software.

Tab. 2: The 3D data formats used to serve in the 3D GeoVolumes API.

3D Data Formats	Web Client and Game Engines Support	Tiling Support	Data Processing
OGC CityGML	No	No	Convert to glTF, I3S, and 3D Tiles format.
OGC CityJSON	No	No	Convert to glTF, I3S, and 3D Tiles format.
OGC Common Database (CDB)	No	No	Convert to glTF, I3S, and 3D Tiles format.
Esri 3D Shapefiles	No	No	Convert to glTF, I3S, and 3D Tiles format.
Esri Multipatch	No	No	Convert to glTF, I3S, and 3D Tiles format.
Wavefront Object file (Obj)	No	No	Convert to glTF, I3S, and 3D Tiles format.
Filmbox (FBX)	Yes	No	Convert to I3S and 3D Tiles format.
Graphics Language Transmission Format (glTF)	Yes	No	Convert to I3S and 3D Tiles format.
Esri Indexed 3D Scene Layer (I3S)	Yes	Yes	No need to convert.
Cesium 3D Tiles	Yes	Yes	No need to convert.

3.2 3D GeoVolumes API Development

The data server had been implemented using Node.js and Express.js web framework and shared publicly at our university's repository (SANTHANAVANICH 2020). This implementation was done following the development and design guide of the 3D GeoVolumes API architecture (MILLER et al. 2020b). The 3D GeoVolumes API was deployed and publicly accessible[2]. The converted 3D data in the area of Stuttgart city, Germany, and New York City, USA, had been served on this 3D GeoVolumes API service (Fig. 2).

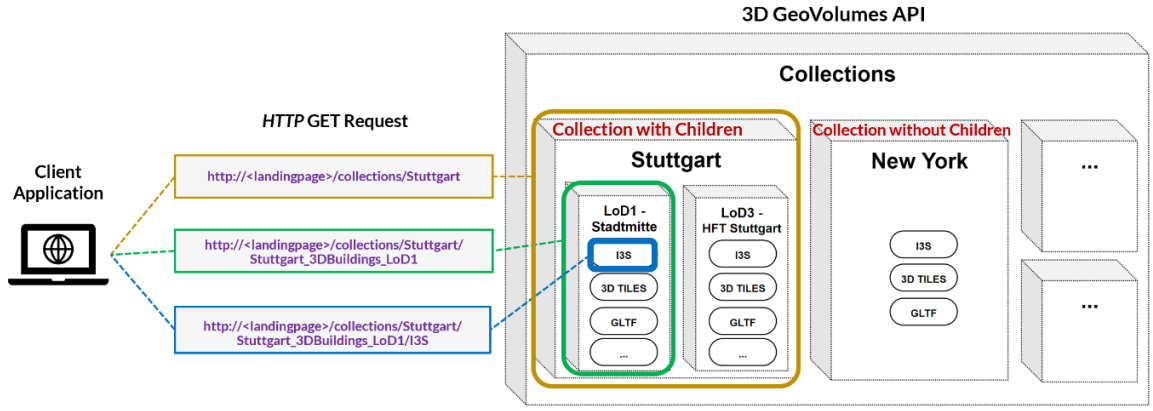


Fig. 2: The structure of developed 3D GeoVolumes API and example resource path URL to the 3D collections and 3D contents.

[2] Available online at https://steinbeis-3dps.eu/3dgeovolumes

3.3 Client Applications

The game engine and web client applications had been developed to evaluate different 3D data formats served through the 3D GeoVolumes API server. The CesiumJS and ArcGIS for JS web libraries were used as candidate web clients, while the Unreal Engine 4 and Unity 3D were used as our candidate game engine application. The convert 3D data were made ready in ESRI I3S, Cesium 3D Tiles, glTF, and FBX format and streamed from the 3D GeoVolumes server. As a result, Tab. 3 shows the evaluation result with the 3D visualization support.

Tab. 3: The 3D visualization support of web and game applications by 3D data formats streamed from the 3D GeoVolumes API.

Data format	Web Clients		Game Engines	
	CesiumJS	ArcGIS for JS	Unreal Engine 4	Unity3D
Esri I3S	Support with our custom plugin.	Support natively	Support with ArcGIS Maps SDK for game engines.	Support with ArcGIS Maps SDK for game engines.
Cesium 3D Tiles	Support natively	Not support	Support with Cesium plugin	Support with Unity3DTiles plugin
glTF	Support natively	Support natively	Support natively	Support natively
FBX	Not support	Not support	Support natively	Support natively

4 Conclusion

This paper explored the use of 3D GeoVolumes API in different client applications with varying degrees. It has shown an excellent solution as a standardized specification for delivering 3D resources to web and game engine applications. This gives the possibility of central data management for multiple web applications. Therefore, the 3D resources only have to be updated and managed at a single location. Additionally, it allows data owners to fully manage data accessibility. Additionally, it delivers 3D resources over the internet, which users can stream and visualize 3D resources in real-time without prior download. Even though the data streaming performance strongly depends on the internet connection, but it would have a significant impact on the application projects that require substantial 3D resources. The 3D GeoVolumes API has also shown interoperability with other OGC sets of standards such as OGC SensorThings API, OGC API Features, etc. In conclusion, the 3D GeoVolumes API can be used as an Urban Data Platform where 3D data sources with their metadata can be stored and provided in a standardized way. In future work, the 3D GeoVolumes API will be further explored and developed in the 3D data management project in cooperation with the Runder Tisch GIS, Technical University of Munich, the Stuttgart Technology University of Applied Sciences, and the State Office for Geoinformation and Land Development at Baden-Württemberg and Bayern. This project aims to use the 3D GeoVolumes API to serve the 3D data of different cities for the various use cases of applications.

5 Bibliography

ABI RESEARCH, 2021: Digital Twins the 'Ultimate Tool' for City Infrastructure Planning. https://gcn.com/articles/2021/08/12/digital-twin-city-planning.aspx, last accessed on 29.11.2021.

AIRAKSINEN, E., 2019: The Kalasatama Digital Twins Project. https://www.hel.fi/static/liitteet-2019/Kaupunginkanslia/Helsinki3D_Kalasatama_Digital_Twins.pdf, last accessed on 29.11.2021.

AGUGIARO, G., GONZALEZ, F. & CAVALLO, R., 2020: The City of Tomorrow from... the Data of Today. ISPRS International Journal of Geo-Information, **9**(9), 554, https://doi.org/10.3390/ijgi9090554.

CZERWINSKI, A., SANDMANN, S., STOCKER-MEIER, E. & PLUMER, L., 2007: Sustainable SDI for EU Noise Mapping in NRW - Best Practice For Inspire. International Journal of Spatial Data Infrastructures Research, **2**, 90-111.

DALY, L. & PHILLIPS, R., 2021: Interoperable Simulation and Gaming Sprint Year 2 Engineering Report. OGC Public Engineering Report. http://docs.ogc.org/per/20-058.html, last accessed on 26.11.2021.

DALY, L. & SERICH, S., 2021: Interoperable Simulation and Gaming Sprint Engineering Report. OGC Public Engineering Report. http://docs.ogc.org/per/20-057.html, last accessed on 26.11.2021.

HEAZEL, C., 2020: OGC API - Common - Part 2: Geospatial Data. OGC Implementation Standard. https://ogcapi.ogc.org/common/, last accessed on 26.11.2021.

MILLER, T., TRENUM, G. & SIMONIS, I., 2020a: 3D Data Container and Tiles API Pilot Summary Engineering Report. OGC Public Engineering Report. https://docs.ogc.org/per/20-031.html, last accessed on 26.11.2021.

MILLER, T., TRENUM, G. & LIEBERMAN, J., 2020b: 3D Data Container Engineering Report. OGC Public Engineering Report. https://docs.ogc.org/per/20-029.html, last accessed on 20.11.2021.

PADSALA, R., GEBETSROITHER-GERINGER, E., PETERS-ANDERS, J. & COORS, V., 2021: Inception of Harmonising Data Silos and Urban Simulation Tools Using 3D City Models for Sustainable Management of the Urban Food Water and Energy Resources. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., **VIII**-4/W1-2021, 81-88. https://doi.org/10.5194/isprs-annals-VIII-4-W1-2021-81-2021.

SANTHANAVANICH, T. & COORS, V., 2021: CityThings: An Integration of the Dynamic Sensor Data to the 3D City Model. Environment and Planning B: Urban Analytics and City Science, **48**(3), 417-432, https://doi.org/10.1177/2399808320983000.

SANTHANAVANICH, T., 2021: OGC 3D GeoVolumes API. Zenodo. https://doi.org/10.5281/zenodo.5731579.

WUERSTLE, P., SANTHANAVICH, T., PADSALA, R. & COORS, V., 2021: Development of a Digital 3D Participation Platform – Case Study of Weilimdorf (Stuttgart, Germany). ISPRS - Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XLVI**-4/W1-2021, 123-129. https://doi.org/10.5194/isprs-archives-XLVI-4-W1-2021-123-2021.

Genauigkeitspotential direkt georeferenzierter UAS-Messungen am Beispiel eines Bahnhaltepunktes

ROBIN GILBERT SCHULZE[1], DAVID MADER[1] & ANETTE ELTNER[1]

Zusammenfassung: In diesem Beitrag wird das Genauigkeitspotential direkt georeferenzierter UAS-Messungen mit einer DJI Phantom 4 RTK analysiert. Es wird der Einfluss unterschiedlicher Befliegungsmuster, Passpunktverteilungen und Kamerakalibrierungen untersucht. Die erreichte Genauigkeit wird anhand von RMS-Abweichungen an den Kontrollpunkten und durch einen Vergleich der bildbasierten Punktwolken mit einer TLS-Punktwolke bewertet. Die Ergebnisse zeigen, dass durch geeignete Aufnahmekonfigurationen Lage- und Höhengenauigkeiten von bis zu 1 cm bei einer Flughöhe von 50 – 70 m über Grund möglich sind. Durch die Kombination von Nadir- und Schrägaufnahmen oder durch das Hinzufügen von zusätzlichen Parametern (Fourier-Reihen) in den Simultankalibrierungsprozess der Kamera besteht die Möglichkeit, gänzlich auf Passpunkte zu verzichten.

1 Einleitung

Unbemannte Luftfahrzeuge (Unoccupied Aerial Vehicle; UAV) bzw. Luftfahrtsysteme (Unoccupied Aircraft System; UAS) haben sich in den letzten Jahren aufgrund der flexiblen Einsatzmöglichkeiten zu einem etablierten Aufnahmesystem für vermessungstechnische Anwendungen entwickelt (LINDSTAEDT & KERSTEN 2018). Viele 3D-Modelle, basierend auf UAS-Bilddatensätzen, sind jedoch mit systematischen Blockverformungen behaftet. Diese Fehler betreffen dabei vor allem die Höhenkomponente des Modells und sind auch als „Dome“ oder „Bowl“ Effekt (Kuppel- oder Schalenfehler) bekannt (SANZ-ABLANEDO et al. 2021). Die Korrektur oder Kompensation der Blockverformungen erfolgt in der Regel über eine indirekte Georeferenzierung mit einer hohen Anzahl an Passpunkten (PP). Die Markierung und Einmessung der PP bedeutet jedoch einen hohen Feldarbeitsaufwand und kann außerdem in schwer oder nicht zugänglichen Geländen nicht umgesetzt werden. Um diesen zeitlichen sowie arbeitsintensiven Mehraufwand zu minimieren und mögliche Blockverformungen zu vermeiden, kann auf RTK (Real-Time Kinematic) oder PPK- (Post-Processed Kinematic) GNSS-fähige UASs zurückgegriffen werden, die die Bestimmung der Bildaufnahmepositionen mit einer hohen Genauigkeit erlauben. Dies ermöglicht die direkte Georeferenzierung der Bildverbände und es kann die notwendige Passpunktanzahl erheblich reduziert werden (z.B. GERKE & PRZYBILLA 2016; TADDIA et al. 2019; PRZYBILLA & BÄUMKER 2020). PRZYBILLA & BÄUMKER (2020) konnten zeigen, dass bei der Verwendung von nur vier PP eine Höhengenauigkeit von 1.2 cm möglich ist und wenn kein PP verwendet wird, eine Höhengenauigkeit von 15 cm erreicht wird. Die UAS-Aufnahmen wurden in Flughöhen von 50 bis 60 m gemacht. TADDIA et al. (2019) nutzten ebenfalls direkte Georeferenzierung und nur einen PP und

[1] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden,
E-Mail: [robin_gilbert.schulze, david.mader, anette.eltner]@tu-dresden.de

erreichten eine Höhengenauigkeit von 5 cm bei einer Flughöhe von 80 m. Das Einbinden von bestimmten Flugmustern, wie z.B. Querflüge, unterschiedliche Flughöhen und Schrägaufnahmen, ermöglicht eine bessere Kamerakalibierung (PRZYBILLA & BÄUMKER 2020). Den Vorteil der Einbindung von Schrägaufnahmen für die Bestimmung der Parameter der inneren Orientierung wurde durch SANZ-ABLANEDO et al. (2021) bestätigt. Neben der Beeinflussung der äußeren Orientierung mittels des Flugmusters muss die innere Orientierung der Kamera hinsichtlich der Auswirkung auf das zu berechnende 3D-Modell beachtet werden. UASs, insbesondere im low-cost Bereich, besitzen Kameras, die starke Verzeichnungen aufweisen können. Diese können nicht ausreichend mit dem Brown-Modell (BROWN 1971) beschrieben werden und erfordern während der Simultankalibrierung (HASTED et al. 2021) die Schätzung zusätzlicher Parameter (Fourier-Reihen; TANG et al. 2012.

In dieser Arbeit wird das Genauigkeitspotential direkt georeferenzierter UAS-Messungen analysiert. Es wird untersucht, welchen Einfluss verschiedene Aufnahmekonfigurationen, Anzahl an PP und Kamerakalibieransätze auf die Genauigkeit der Ergebnisse haben. In Kapitel 2 werden Testgebiet, Fluggerät, Befliegungsmuster, Datenverarbeitung und die für die Validierung notwendigen Referenzdaten vorgestellt. Kapitel 3 zeigt die Auswirkung der verschiedenen Aufnahme- und Passpunktkonfigurationen sowie Kameramodelle und den Vergleich mit den Referenzdaten. Kapitel 3 beschreibt das Fazit und gibt einen kurzen Ausblick.

Abb. 1: Oben: Übersicht der Passpunktverteilung (gelb), Höhenprofil (rot) und TLS-Positionen (grün). Unten v.l.n.r.: Passpunktmarkierungen (schwarz-weiß Targets, aufgesprühte Farbmarkierungen, markante Punkte in der Örtlichkeit), DJI Phantom 4 RTK.

2 Methodik

2.1 Untersuchungsgebiet

Die UAS-basierten Bilddatensätze wurden über einem Haltepunkt der Deutschen Bahn aufgenommen (Abb. 1). Das Testgebiet erstreckt sich über eine Fläche von 200 × 50 m^2. Der Bahnsteig liegt auf einem ca. 5 m erhöhten Bahndamm.

2.2 Flugplanung und -durchführung

Die Aufnahme der Bilddatensätze erfolgte mit der DJI Phantom 4 RTK. Dieses UAS-Modell ist mit einem integrierten Zweifrequenzempfänger (RTK-GNSS-Modul) ausgestattet. Durch das Empfangen von Korrekturdaten über eine Echtzeitverbindung zum SAPOS-Dienst (SAPOS 2021), ist eine zentimetergenaue Positionierung realisierbar. Zusätzlich speichert die Phantom 4 RTK die Satellitenbeobachtungs- und Ephemeridendaten während des Fluges in einem speziellen Dateiformat (RINEX.obs) ab. Mit geeigneter Software ist somit auch eine PPK-Auswertung möglich. Die an einem Gimbal fixierte Kamera verfügt über einen 1“ CMOS-Sensor mit 20 Megapixel und einer Brennweite von 8,8 mm. Die erreichbare Bodenauflösung (GSD) beträgt ca. 1,3 cm bei einer Flughöhe von 50 m über Grund.
Die Flugplanung erfolgte mit der DJI GS Pro App. Es wurden vier Teilflüge durchgeführt, die sich aus verschiedenen Flugmustern zusammensetzten:

- Konventionelle Befliegung
- Kreuzbefliegungen
- Befliegungen in unterschiedlichen Höhen (50 – 70 m)
- Nadir- und Schrägaufnahmen (30° Offset zu Nadir)
- Längs- und Querüberlappung mind. 80/60

Anschließend wurden aus den Teilflügen fünf verschiedene Aufnahmekonfigurationen bzw. Testszenarien (TS) festgelegt (Tab. 1 und Abb. 2). Die Konfiguration TS1 *Regular 50* beinhaltet Bilder, die mit dem konventionellen Flugmuster (Nadiraufnahmen) von einer Flughöhe von 50 m aufgenommen wurden. TS2 *Kreuz 50* beinhaltet eine Kreuzbefliegung (Nadiraufnahmen), die ebenfalls auf einer Höhe von 50 m durchgeführt wurde. TS3 *Regular 50/60* entspricht einem konventionellen Befliegungsmuster auf zwei Flughöhen (50 und 60 m). TS4 *Regular 50/ Kreuz Oblique 70* umfasst eine konventionelle Befliegung auf 50 m Flughöhe und eine Kreuzbefliegung mit Nadiraufnahmen, die von 70 m Flughöhe aufgezeichnet wurden. Die letzte Aufnahmekonfiguration TS5 *Multi-Oriented* berücksichtigt verschiedene Flughöhen, Schrägaufnahmen sowie Kreuzbefliegung und besteht nur aus Bildern, die während des vierten Teilfluges aufgenommen wurden.

Tab. 1: Zusammenfassung der zusammengestellten Testszenarien (TS). R = Regular/ Standardbefliegung, R* = Flugstreifen, R** = 2 Flugstreifen in entgegengesetzter Flugrichtung, K = Kreuzbefliegung, N = Nadiraufnahmen, O = Obliqueaufnahmen, GSD = Ground Sampling Distance.

Testszenarien	Aufnahmekonfiguration	~ GSD	Flugzeit	Fotos
1 – *Regular 50*	R/ N/ 50	1,23 cm	6 min	94
2 – *Kreuz 50*	R/ N/ 50 K/ N/ 50	1,23 cm	9 min	122
3 – *Regular 50/60*	R/ N/ 50 R*/ N/ 60	1,28 cm	8 min	108
4 – *Regular 50/ Oblique 70*	R/ N/ 50 K/ O/ 70	1,59 cm	11 min	170
5 – *Multi-Oriented*	R*/ N/ 50 R**/ O/ 70 K/ O/ 70	1,86 cm	8 min	120

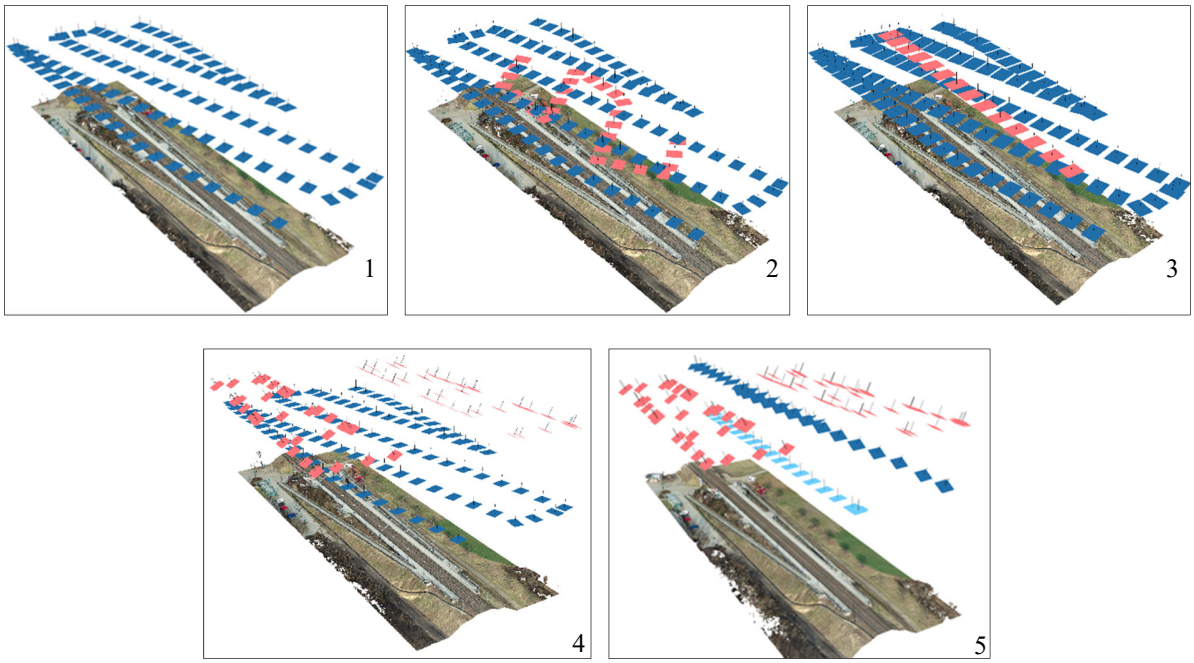


Abb. 2: Kombinationen der vier Teilflüge zu fünf Aufnahmekonfigurationen: 1) TS1 *Regular 50*, 2) TS2 *Kreuz 50*, 3) TS3 *Regular 50/60*, 4) TS4 *Regular 50/ Kreuz Oblique 70*, 5) TS5 *Multi-Oriented*

2.3 Direkte Georeferenzierung

Da während der Flüge keine Echtzeitverbindung zum SAPOS-Dienst möglich war, wurden die UAS-Positionen in einer PPK-Auswertung mit der Software RTKLIB (TAKASU 2021) bestimmt. Die für den Beobachtungszeitraum notwendigen Satellitenkorrekturdaten wurden über den SAPOS-GPPS-Dienst bezogen (SAPOS 2021). Bei 94 % der Bilder konnte das Problem der Mehrdeutigkeiten gelöst und eine zentimetergenaue Positionierung erreicht werden. Die Standardabweichungen betragen 1 cm bei den ersten drei Teilflügen und 2 cm beim vierten Teilflug. Da sich die PPK-bestimmten Positionen zunächst auf das Antennenphasenzentrum beziehen, musste noch der Versatz (leverarm) zwischen diesem und dem Kameraprojektionszentrum berücksichtigt werden (TADDIA et al. 2019).

2.4 Verarbeitung der Bilddaten

Die photogrammetrische Verarbeitung der Bilddatensätze und Erstellung der 3D-Modelle erfolgte in Agisoft Metashape (Version 1.7.4) mit den Standardeinstellungen. Die Bildverbände aller Aufnahmekonfigurationen wurden einerseits ausschließlich auf den mittels PPK-GNSS berechneten Bildpositionen gelagert. Andererseits wurden unterschiedliche Passpunktkonfigurationen in der Bündelblockausgleichung (BBA) berücksichtigt:

- ein PP in der Blockmitte
- vier PP an den Blockecken
- fünf PP in der Blockmitte und an den -ecken
- elf gleichmäßig verteilte PP

Neben den verschiedenen Aufnahme- und Passpunktkonfigurationen wurden zwei verschiedene Modelle der inneren Kamerageometrie für alle Flugmuster- und Passpunktkombinationen berücksichtigt; einerseits das Standard Brown-Modell (BROWN 1971) und andererseits die zusätzliche Einbeziehung von weiteren Parametern eines Fourier-Reihen basierten Modells (TANG et al. 2012).
In einer weiteren Auswertungsvariante wurden die Bildverbände aller Aufnahmekonfigurationen auf metergenauen GNSS-Positionen gelagert und mit vier, fünf bzw. elf PP indirekt georeferenziert, um einen Vergleich aller Kombinationen zur typischerweise durchgeführten UAV-Photogrammetrie zu erhalten.

2.5 Referenzdatensätze

Um das Genauigkeitspotential beurteilen zu können, wurden verschiedene Referenzdatensätze aufgezeichnet (Abb. 1). Unter anderem wurden 27 Kontrollpunkte (KP) gleichmäßig im Feld verteilt und mittels PPK-GNSS eingemessen. Die dabei erreichte Standardabweichung betrug 1 cm (in Lage und Höhe). Ein Teil der KP wurde zusätzlich als PP für die Lagerung der Bildverbände genutzt.
Bei der Nutzung von KP für die Genauigkeitsbetrachtung können systematische Blockverformungen unter Umständen unerkannt bleiben. Höhenprofile dienen dazu, die Erkennung solcher Systematiken zu unterstützen. Es wurde ein Höhenprofil (22 Höhenpasspunkten – HP) mittels PPK-GNSS mit der gleichen Genauigkeit wie die KP gemessen, um mögliche systematische Blockverformungen detektieren zu können.
Der dritte Referenzdatensatz ist eine terrestrisch aufgenommene Laserscannerpunktwolke. Das Testgelände wurde mit dem Laserscanner RIEGL VZ 400i von 14 Standpunkten aus erfasst. Der Laserscandatensatz wurde mit RiScan Pro (Version 2.8) aufbereitet. Die einzelnen Laserscannerstandpunkte wurden mittels eines Iterative-Closest-Point (ICP) –Ansatzes (HORN 1987) in ein gemeinsames Koordinatensystem transformiert. Anschließend wurden retroreflektierende Zylinderzielmarken genutzt, welche auf den eingemessenen PP aufgestellt wurden, um die fusionierte Laserscannerpunktwolke (TLS) zu georeferenzieren. Die Punktwolke wurde in CloudCompare (CLOUDCOMPARE 2021) manuell von umliegender Vegetation bereinigt, um den Vergleich auf bebaute Flächen zu konzentrieren. Für den Vergleich der UAV-basierten Punktwolken mit den TLS-Daten wurde die Funktion Cloud-to-Cloud-Distanz verwendet.

3 Ergebnisse und Diskussion

3.1 Einfluss der Aufnahmekonfiguration

In Abb. 3 sind die RMS-Abweichungen an den im Testfeld installierten KP (Abb. 1) dargestellt. Die Resultate zeigen, dass bei einer direkten Georeferenzierung der Bildverbände Lagegenauigkeiten von bis zu 1,2 cm erreichbar sind – unabhängig vom Befliegungsmuster und bei Verwendung des Brown-Modells ohne zusätzliche Fourier-Reihen-Parameter bei der Simultankalibrierung. Lediglich die Lage-RMS-Werte der TS5 *Multi-Oriented* Variante sind mit 2,2 cm etwa doppelt so groß. Dabei muss beachtet werden, dass TS5 die Aufnahmekonfiguration ist, welche nur aus Bildern des vierten Teilfluges bestand, dessen Positionierungsgenauigkeit 2 cm (versus 1 cm für die ersten drei Teilflüge) betrug. Die Lagegenauigkeiten liegen im Bereich der GSD und im Genauigkeitsniveau der PPK-Bildpositionen.

Die Betrachtung der Höhengenauigkeit zeigt größere Unterschiede zwischen den Aufnahme- und Passpunktkonfigurationen auf. Das ausschließlich aus Nadiraufnahmen bestehende TS1 *Regular 50* besitzt mit 19,0 cm eine hohe Abweichung der Kontrollpunkte in der vertikalen Komponente. Durch das Hinzufügen von Bildern aus einer Kreuzbefliegung (TS2 *Kreuz 50*) oder durch Aufnahmen in unterschiedlichen Flughöhen (TS3 *Regular 50/60*) kann der Höhenfehler auf bis zu 4,1 cm reduziert werden. Den größten Einfluss auf die Höhengenauigkeit zeigen die Kombinationen von Nadir- und Schrägaufnahmen aus Kreuzbefliegungen in unterschiedlichen Höhen. Diese beträgt auch ohne Passpunkte 1,0 cm (TS4 *Nadir 50/ Kreuz Oblique 70)* bzw. 1,7 cm (TS5 *Multi-Oriented*). Bei diesen Szenarien zeigt auch der Vergleich mit dem Höhenprofil die niedrigsten Abstände zwischen den HP und UAV-basierten Punktwolken (zwischen -1,2 und 3,0 cm; Abb. 4).

3.2 Einfluss der Passpunktkonfiguration

Die Betrachtung der HP zeigt, dass der Höhenfehler für TS1 (~20 cm) bis TS3 (~4 cm) gleichmäßig im 3D-Modell gegeben ist, wenn keine PP genutzt werden (Abb. 4). Bei einer Lagerung der Bildverbände auf einen in der Blockmitte befindlichen PP kann der vertikale Offset bei TS1 bis TS3 deutlich minimiert werden. Die RMS-Abweichungen der Höhenkomponente der KP betragen dann noch 2,0 bis 2,6 cm. Bei den Konfigurationen mit konvergenten Bildaufnahmen kann kein zusätzlicher Genauigkeitsgewinn durch einen PP in der Blockmitte beobachtet werden. Werden vier PP in den Blockecken des Aufnahmegebiets in die BBA integriert, wird die Höhengenauigkeit bei TS1 *Regular 50* weiter verbessert (1,4 cm). Weitere Verbesserungen durch fünf bzw. elf PP sind im Vergleich zu vier PP kaum messbar. Der Mehraufwand bei der Einmessung weiterer PP ist somit kritisch zu beurteilen.

3.3 Einfluss des Verzeichnungsmodells

Werden neben den Parametern des Brown-Modells zusätzliche Parameter (Fourier-Reihen) in der Simultankalibrierung berücksichtigt, werden die Genauigkeiten stark verbessert. Die Lage- und Höhengenauigkeit liegt bereits bei einer Standardbefliegung (TS1 *Regular 50*) ohne PP bei etwa 1,3 bzw. 1,1 cm. Zusätzliche Nadiraufnahmen in einer Kreuzbefliegung oder von einer anderen Flughöhe führen zu einer Verschlechterung der Ergebnisse um wenige Millimeter. Die Kombination von Nadir- und Schrägaufnahmen (TS4 *Nadir 50/ Kreuz Oblique 70*) verbessert die Höhengenauigkeit auf 0,9 cm.

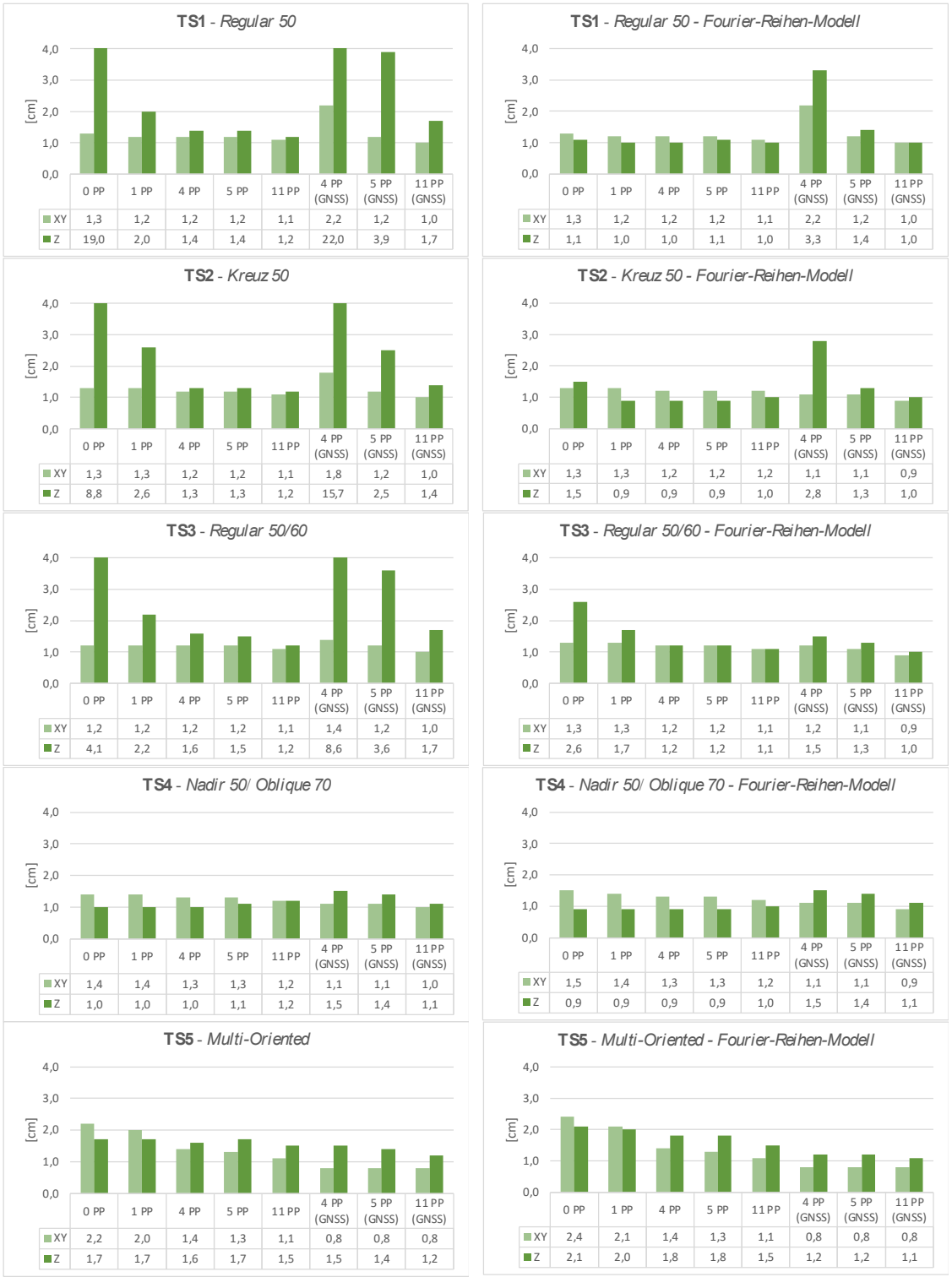


Abb. 3: Dargestellt sind die RMS-Abweichungen [cm] an den Kontrollpunkten. Um die Ergebnisse visuell besser darzustellen, wurde die Y-Achse auf 4,0 cm begrenzt – einige RMS-Werte der Z-Koordinate übersteigen diese deutlich.

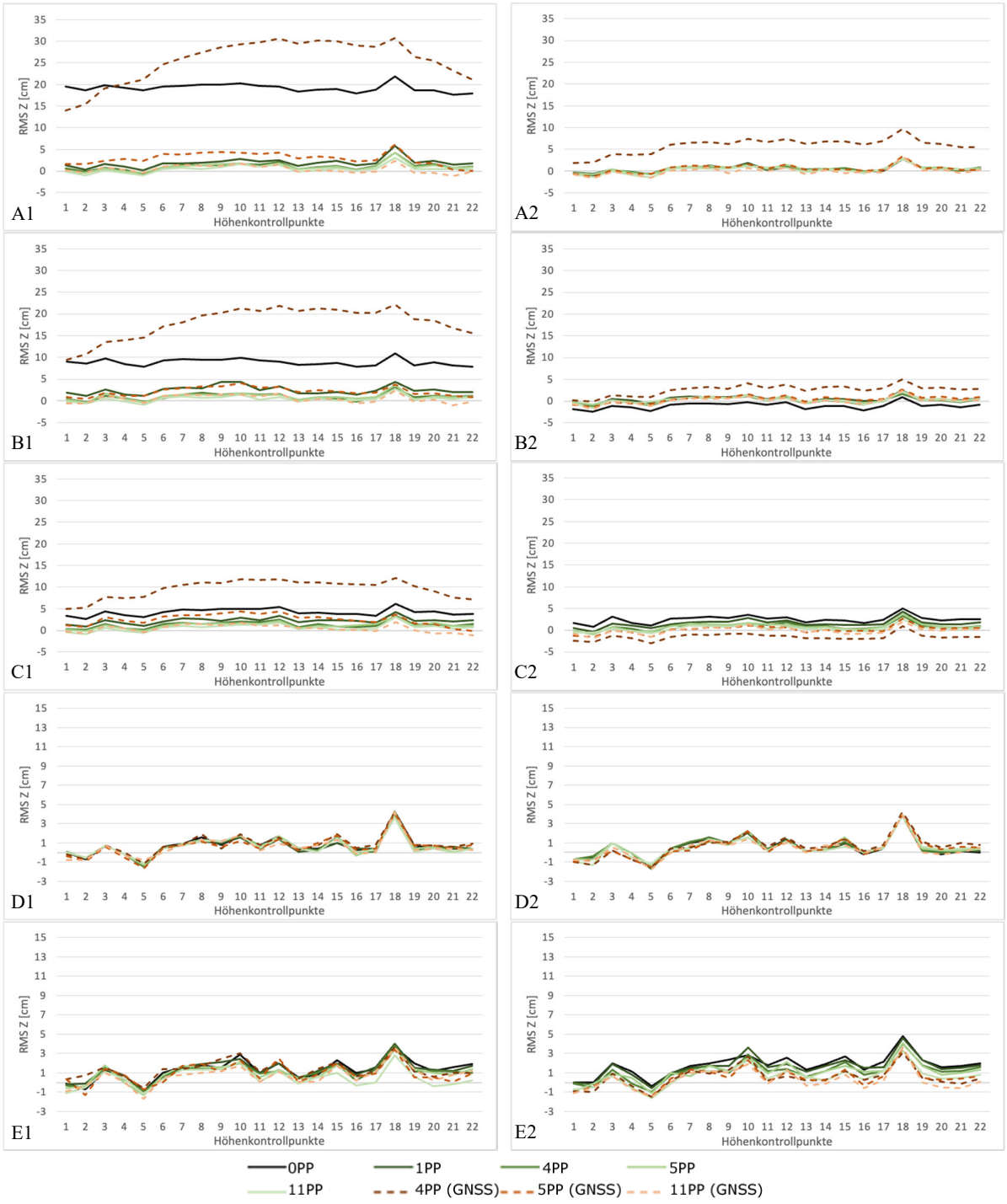


Abb. 4: Dargestellt sind die RMS-Werte der Z-Koordinaten an den Höhenkontrollpunkten des Höhenprofils für jede Aufnahmekonfiguration. A) *Regular* 50, B) *Kreuz 50*, C) *Regular 50/60*, D) *Regular 50/ Oblique 70*, E) *Multi-Oriented*; Simultankalibrierung 1) ohne zusätzliche Parameter, 2) mit zusätzlichen Parametern. Zur besseren Visualisierung unterscheidet sich die Skalierung der Y-Achse für D und E. Anmerkung: KP 18 ist ein im Schatten einer Laterne befindlicher Ausreißer.

3.4 Einfluss der Kamerapositionsgenauigkeit

Die Untersuchungsergebnisse der indirekten Georeferenzierung der Bildverbände lassen für diese Studie zwei Schlussfolgerungen zu:

i) Bestehen die Flugmuster ausschließlich aus Nadiraufnahmen (TS1-TS3), sind elf PP notwendig, um ähnliche Genauigkeiten zu erreichen, wie bei der direkten Georeferenzierung und der zusätzlichen Nutzung eines PP. Wenn zusätzlich das Fourier-Reihen-Modell berücksichtigt wird, sind fünf PP ausreichend.

ii) Werden die Aufnahmekonfigurationen um Schrägaufnahmen aus einer Kreuzbefliegung in unterschiedlichen Flughöhen ergänzt (TS4), müssen vier PP in die BBA integriert werden, um ähnliche Genauigkeiten zu erreichen, wie bei der direkten Georeferenzierung.

Bei der indirekten Georeferenzierung und einer Lagerung des Bildverbandes auf nur vier PP in den Blockecken sind typische Kuppelfehler erkennbar (Abb. 4). Erst durch die Einführung weiterer PP oder Schrägaufnahmen kann die Wölbung im 3D-Modell minimiert werden.

3.5 Auswirkung auf die Bestimmung der Parameter der inneren Orientierung

In diesem Abschnitt werden die in der Simultankalibrierung bestimmten Parameter der inneren Orientierung (Brennweite, Hauptpunktversatz) diskutiert. Während die Lage des Hauptpunktes bei allen Aufnahmekonfigurationen zufällige Variationen unter 1 Pixel aufweisen, zeigen die Brennweiten größere Abweichungen (Abb. 5). Wenn die Kamerapositionen mit schlechteren Genauigkeiten (indirekte Georeferenzierung) in die BBA eingehen, weist die Brennweite Unterschiede von bis zu 45 Pixel zwischen den Aufnahme- und Passpunktkonfigurationen auf. Des Weiteren ist ein Zusammenhang zwischen der berechneten Brennweite und der berechneten Höhe des Projektionszentrums gegeben; mit zunehmender Brennweit nimmt die Höhe zu und vice versa.

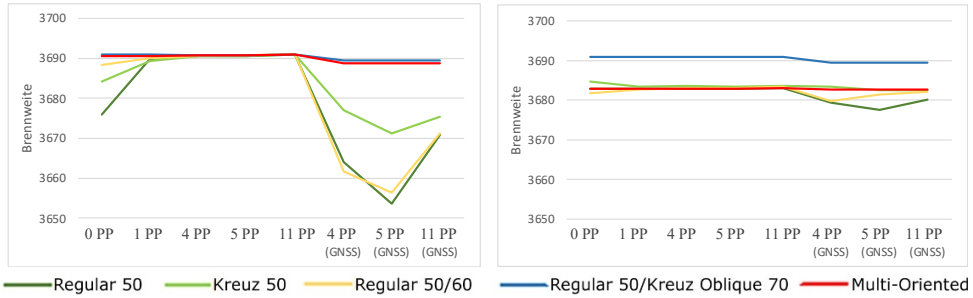


Abb. 5: Brennweiten nach Simultankalibrierung basierend auf Brown-Modell (links) und Brown-Modell + Fourier-Reihen-Modell (rechts)

Bei der direkten Georeferenzierung zeigen die Aufnahmekonfigurationen ohne PP ebenfalls größere Schwankungen in der Brennweitenbestimmungen. Wird jedoch mindestens ein PP bei der direkten Georeferenzierung verwendet, besitzt die Brennweite von TS2 bis TS5 geringe Unterschiede. Wenn zusätzlich die Parameter des Fourier-Reihen-Modells gebraucht werden, variiert die Brennweite nur noch sehr wenig zwischen den verschiedenen Passpunktkonfigurationen.

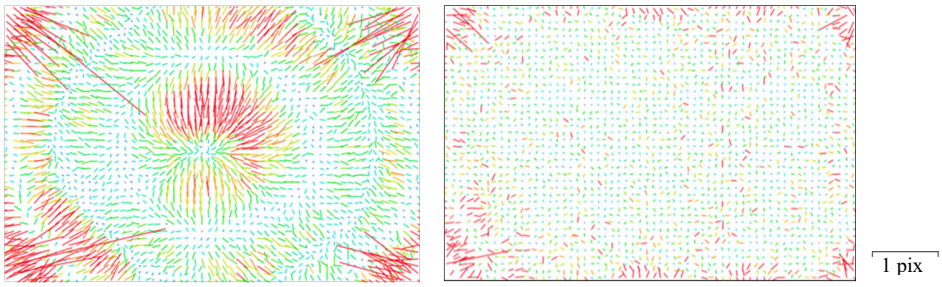


Abb. 6: Residuen der Verzeichnungen nach Simultankalibrierung basierend auf Brown-Modell (links) und Brown-Modell + Fourier-Reihen-Modell (rechts) für das TS1 *Regular 50*. Skalierungsfaktor 1000.

In Abb. 6 sind die Residuenmuster nach der Bestimmung der Verzeichnungsparameter des DJI Kameramodells der P4RTK dargestellt. Durch die zusätzlichen Korrekturparameter des Fourier-Reihen-Modells können systematische Abweichungen im Bildraum stark minimiert werden. Die Verbesserung im Bildraum führt auch zu einer höheren Genauigkeit der 3D-Modelle, das anhand der Abweichungen der KP sowie der HP erkennbar ist.

3.6 Betrachtung der räumlichen Fehlerverteilung

Der Cloud-to-Cloud-Distanz Vergleich zwischen den bildbasierten und TLS -Punktwolken bestätigt die Beobachtung dieser Studie, dass bei einer direkten Georeferenzierung der Bilddatensätze keine globalen systematischen Blockdeformationen in Form von Kuppel- oder Schalenfehlern auftreten (Abb. 7A). Dies wird insbesondere an den befestigten Flächen (Bahnsteig, Asphaltstraße, Fußgängerrampe) deutlich. Die Aufnahmekonfiguration mit dem geringsten Aufwand in der Datenakquise (TS1 *Regular 50* ohne PP) erzielt einen RMS Punktabstand von 1,7 cm, wenn die zusätzlichen Parameter des Fourier-Reihen-Modells berücksichtigt werden. Bei der komplexeren Aufnahmekonfiguration TS4 *Regular 50/ Oblique 70* (keine PP) unter Berücksichtigung nur des Brown-Modells ist der RMS Punktabstand ähnlich (1,8 cm). Die indirekt georeferenzierten UAV-Daten besitzen einen großen systematischen Fehler von stellenweise mehr als 30 cm für TS1 *Regular 50* und bei Verwendung von lediglich vier PP (Abb. 7B). Der RMS Wert der Punktabstände zwischen TLS- und UAV-basierter Punktwolke beträgt in diesem Fall 18,1 cm. Wird eine komplexere Aufnahmekonfiguration (z.B. TS4 *Regular 50/ Oblique 70*) geflogen, verbessert sich die Höhengenauigkeit deutlich und der RMS Wert beträgt nur noch 1,9 cm.

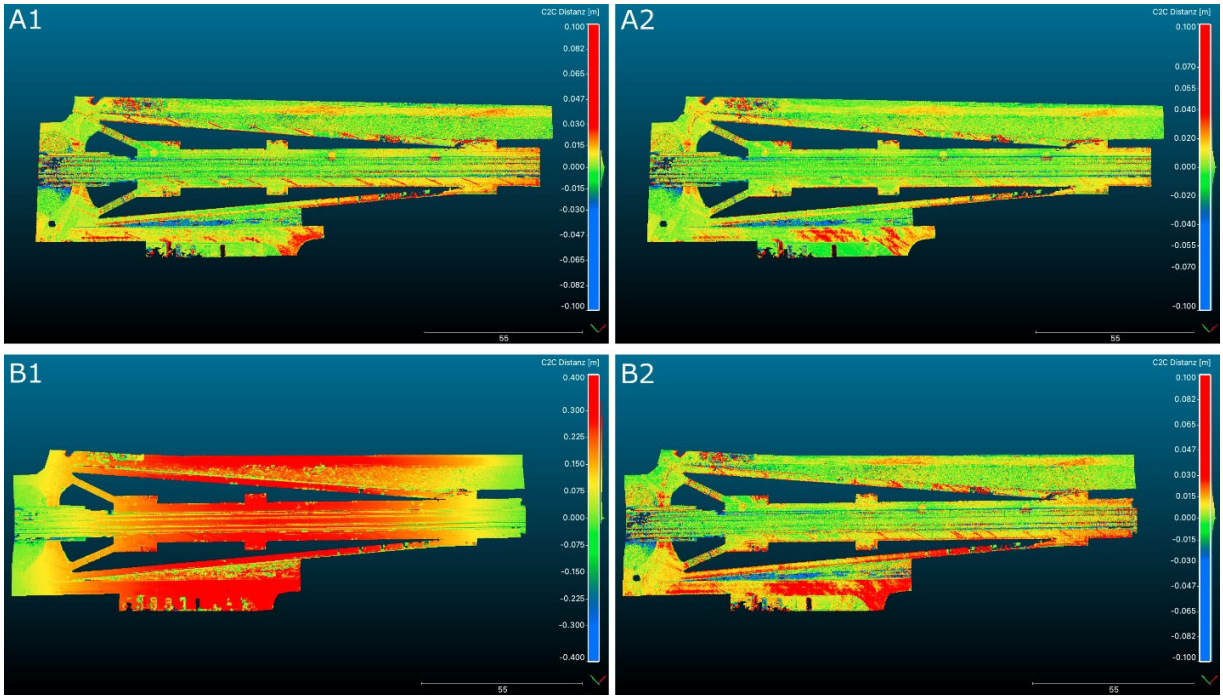


Abb. 7: PW-Vergleich (C2C Distanz) direkt (A) und indirekt (B) georeferenzierter UAV-basierter Punktwolken mit dem TLS-Datensatz. Vergleich TLS mit TS1 *Regular 50 – Fourier-Reihen-Modell* (A1) und TS4 *Regular 50/ Oblique 70* (A2), TS1 und TS4 ohne PP; PW-Vergleich TLS mit TS1 *Regular 50* GNSS (B1) und TS4 *Regular 50/ Oblique 70* GNSS (B2), TS1 und TS4 GNSS mit 4 PP

Der Vergleich mit den TLS-Daten erlaubt die Betrachtung lokaler Einflüsse auf die Genauigkeiten. Es wird erkennbar, dass sich Schattenformationen negativ auf die bildbasierten Punktwolken auswirken. Die Schatten von Bäumen, Masten oder Geländern sorgen für lokale Artefakte mit Abweichungen von bis zu 5 cm (Abb. 8). Diese Fehler treten unabhängig vom Befliegungsmuster und der Passpunktverteilung auf.

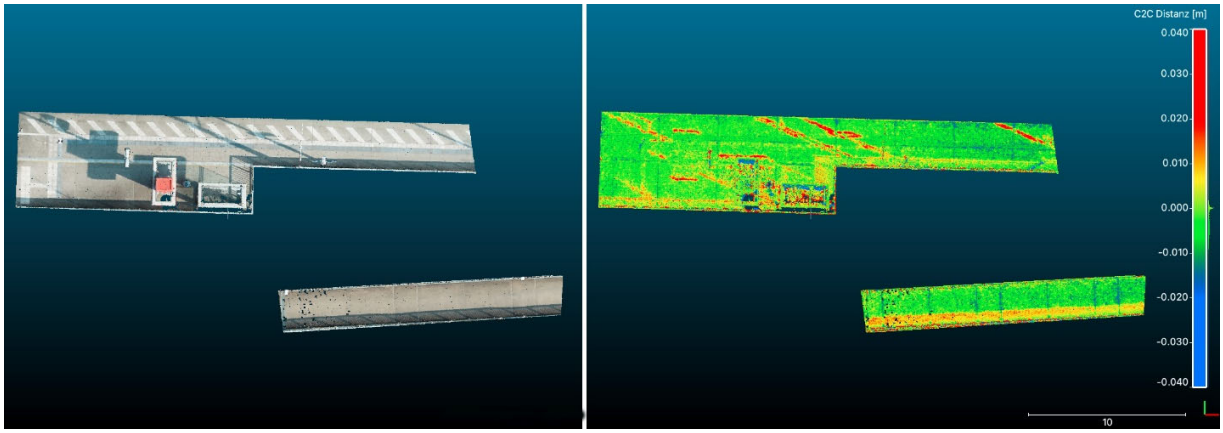


Abb. 8: Artefakte in den UAV-Punktwolken durch Schatten (Vergleich TLS – TS1 *Regular 50* - 0 PP)

4 Fazit und Ausblick

In dieser Studie wurde die Auswirkung fünf verschiedener Bildaufnahmekonfigurationen, fünf verschiedener Passpunktkonfigurationen, unterschiedlicher Genauigkeiten der Kameraposition (direkte und indirekte Georeferenzierung) und zwei verschiedener Verzeichnungsmodelle auf die Genauigkeit eines mittels UAV-Photogrammetrie berechneten 3D Modells untersucht. Es konnte gezeigt werden, dass durch eine direkte Georeferenzierung der Bildverbände Lage- und Höhengenauigkeiten von bis zu 1 cm möglich sind. Setzt sich der Bilddatensatz ausschließlich aus Nadiraufnahmen in einer festen Flughöhe zusammen und es wird kein PP verwendet, entsteht ein konstanter Höhenoffset. Dieser kann durch das Hinzufügen eines einzigen PP, durch Schrägaufnahmen in einer anderen Flughöhe oder durch das Anwenden eines Fourier-Reihen-Modells der Verzeichnung in der BBA minimiert bzw. eliminiert werden. Dadurch kann auf eine umfangreiche und aufwendige Passpunktvermessung verzichtet werden. Dennoch ist es unabhängig von der Aufnahmekonfiguration empfehlenswert, eine kleine Anzahl an PP aufzunehmen, um sie als KP zur unabhängigen Beurteilung der erreichten Genauigkeit nutzen zu können. Die Erkenntnisse dieser Studie sind bedeutend für UAV-Einsätze mit einer hohen Wiederkehrzeit (z.B. Baustellenüberwachung oder Flächenmonitoring) oder in unwegsamen Gelände (z.B. Hangrutschung oder Bergbau).

5 Literaturverzeichnis

BROWN, D.C., 1971: Close-Range camera calibration. Photogrammetric Engineering, **37**, 855-866.

GERKE, M. & PRZYBILLA, H.-J., 2016: Accuracy Analysis of Photogrammetric UAV Image Blocks: Influence of Onboard RTK-GNSS and Cross Flight Patterns. PFG – Photogrammetrie – Fernerkundung – Geoinformation, Heft 1, Stuttgart, E. Schweizerbart'sche Verlagsbuchhandlung, 17-30.

CLOUDCOMPARE, 2021: GPL software, version 2.6.2, retrieved from http://www.cloudcompare.org/.

HASTED, H., LUHMANN, T., PRZYBILLA, H.-J. & ROFALLSKI, R., 2021: Evaluation of the interior orientation modelling for cameras with aspheric lenses and image pre-processing with special emphasis to SFM reconstruction. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **IXLIII**-B2-2021, 17-24, https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-17-2021.

HORN, B., 1987: Closed-Form Solution of Absolute Orientation Using Unit Quaternions. JOSA A, **4**(4), 629-642.

LINDSTAEDT, M. & KERSTEN, T., 2018: Zur Bedeutung von Passpunkten bei der Aerotriangulation UAV-basierter Bildflüge. UAV 2018 - Vermessung mit unbemannten Flugsystemen, Schriftenreihe des DVW, Band **89**, Beiträge zum 169. DVW-Seminar am 19. und 20. Februar 2018 in Hamburg, Wißner-Verlag, Augsburg, pp. 81-101.

PRZYBILLA, H.-J. & BÄUMKER, M., 2020: Untersuchungen zur Qualität des Realtime Kinematic GNSS Systems der DJI Phantom 4 RTK. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **29**, 47-61.

PRZYBILLA, H.-J., LINDSTAEDT, M. & KERSTEN, T., 2019: Untersuchungen zur Qualität bildbasierter Punktwolken aus UAV-Bildflügen. In: Luhmann, T., Schumacher C. (Hrsg.): Photogrammetrie, Laserscanning, Optische 3D-Messtechnik. Beiträge der Oldenburger 3D-Tage 2020, Berlin, Wichmann Verlag, 236-248.

PRZYBILLA, H.-J., BÄUMKER, M., LUHMANN, T., HASTEDT, H. & EILERS, M., 2020: Interaction between direct georeferencing, control point configuration and camera self-calibration for RTK-based UAV photogrammetry. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **IXLIII**-B2-2021, 485-492, https://doi.org/10.5194/isprs-archives-XLIII-B1-2020-485-20.

SANZ-ABLANEDO, E., CHANDLER, J.-H., BALLESTEROS-PÉREZ, P. & RODRÍGUEZ-PÉREZ, J.-R., 2020: Reducing systematic dome errors in digital elevation models through better UAV flight design. Earth Surface Processes and Landforms, **45**, 2134-2147, https://doi.org/10.1002/esp.4871.

SAPOS, 2021: Satellitenpositionierungsdienst SAPOS der sächsischen Vermessungsverwaltung. http://www.landesvermessung.sachsen.de/sapos/, letzter Zugriff am 09.04.2021.

TADDIA, Y., STECCHI, F. & PELLEGRINELLI, A. 2019: Using DJI Phantom 4 RTK drone for topographic mapping of coastal areas. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XLII**-2/W13, 625-630, https://doi.org/10.5194/isprs-archives-XLII-2-W13-625-201.

TAKASU, T., 2021: RTKLIB: An Open Source Program Package for RTK-GPS. http://gpspp.sakura.ne.jp/rtklib/rtklib.htm, letzter Zugriff: 09.05.2021.

TANG, R., FRITSCH, D. & CRAMER, M., 2012: New rigorous and flexible Fourier self-calibration models for airborne camera calibration. ISPRS Journal of Photogrammetry and Remote Sensing, **71**, 76-85, https://doi.org/10.1016/j.isprsjprs.2012.05.004.

Machbarkeitsstudie zur Prädiktion von Waldbränden aus Sentinel-2 Zeitreihen

JANIK STEIER[1], LINA E. BUDDE[1] & DOROTA IWASZCZUK[1]

Zusammenfassung: Eine frühzeitige Vorhersage von Waldbränden ist entscheidend, um die negativen Auswirkungen und Gefahren für Ökosysteme und Menschen zu minimieren. Herkömmliche Waldbrandprädiktionssysteme, wie das weitverbreitete Canadian Forest Fire Danger Rating System, generieren ihre benötigten Informationen in der Regel durch lokale Messstationen auf der Erdoberfläche. Dies führt zu hohen operativen Kosten und macht es daher kaum möglich diese Systeme flächendeckend zu unterhalten. In der vorliegenden Studie wird anhand von multispektralen Sentinel-2 Zeitreihen sowie verschiedener Maschine Learning Modelle, eine Waldbrandprädiktion vorgenommen und damit eine kostengünstige und flächendeckende Alternative zu herkömmlichen Waldbrandprädiktionssystemen bereitgestellt. Für die Evaluierung der erarbeiteten Methodik wird ein Waldbrandereignis im Mai 2021 in Griechenland untersucht. Das leistungsstärkste implementierte Modell prädiziert 81 % des tatsächlichen Waldbrandgebiets korrekt voraus.

1 Einleitung

Waldbrände sind ein global auftretendes Phänomen und in vielen Regionen ein Bestandteil der natürlichen Prozesse in einem Ökosystem. Ökosysteme können von Waldbränden profitieren, da sie tote Vegetation beseitigen, eingeschlossene Nährstoffe freisetzen und neues Vegetationswachstum unterstützen.
Dennoch treten Waldbrände häufig in Gebieten auf in denen sie eine Gefahr für die natürliche Vegetation und den Lebensraum von Tieren und Menschen darstellen.
Desweitern haben Waldbrände aufgrund der Freisetzung von Treibhausgasen einen signifikanten Anteil am Klimawandel. So lassen sich 15 % der globalen Treibhausgase auf Waldbrände zurückführen. Vor allem bei der Freisetzung von Kohlenstoffmonoxid und Methan, mit einem Anteil von 32 % und 10 % an den globalen Emissionen, spielen Waldbrände eine wichtige Rolle. Zudem erhöht ein wärmeres Klima das Waldbrandrisiko sowie die Ausbreitung und die Auftrittsfrequenz von Waldbränden. Das wiederum bedingt eine noch höhere Freisetzung von Treibhausgasen (HIRSCHBERGER 2016).
Eine frühzeitige Vorhersage von Waldbränden ist daher entscheidend, um die negativen Auswirkungen und Gefahren für Ökosysteme und Menschen zu minimieren.
Das weitverbreitetste Waldbrandprädiktionssystem ist das Canadian Forest Fire Danger Rating System (CFFDRS), welches von vielen Organisationen wie z.B. dem European Forest Fire Information System oder dem New Zealand Forest Service übernommen wird.
Herkömmliche Waldbrandprädiktionssysteme mit einem umfassenden Waldbrandvorhersagemodell arbeiten mit einer Vielzahl an wetter- und bodenbezogenen Variablen und den daraus

[1] Technische Universität Darmstadt, Fachbereich Bau- und Umweltingenieurwissenschaften, Fachgebiet Fernerkundung und Bildanalyse, Franziska-Braun-Straße 7, D-64287 Darmstadt, E-Mail: [janik.steier, lina.budde, dorota.iwaszczuk]@tu-darmstadt.de

entstehenden komplexen Beziehungen. So verwendet das CFFDRS unter anderem Informationen über das Entzündungsrisiko, das Wetter, die Topographie und die Vegetation, um das Waldbrandauftreten und das Brandverhalten vorherzusagen. Diese Art von Vorhersagemodellen leidet allerdings unter zwei wesentlichen Nachteilen. Zum einen müssen oftmals vereinfachte Annahmen getroffen werden, um die komplexen Beziehungen annähernd zu beschreiben. Zum anderen werden viele Informationen und Daten für die Modelle mit Instrumenten und lokalen Messstationen am Erdboden gewonnen. Das führt zu hohen operativen Kosten und zu einer Datengewinnung in einem beschränkten spezifischen Gebiet. Vor allem die hohen operativen Kosten können Entwicklungsländer wie beispielsweise Indonesien abschrecken, in flächendeckende Systeme zur Waldbrandprädiktion zu investieren (YANG et al. 2021).
Eine globale Prädiktionskarte des Waldbrandrisikos bietet das Global Wildfire Information System (GWIS) an. Dabei arbeitet das System auf Grundlage von meteorologischen Vorhersagedaten und stellt unter anderem den Candian Forest Fire Weather Index, welcher eine Subkomponente des CFFDRS ist, dar. Der Vorhersagezeitraum liegt bei GWIS zwischen einem und neun Tagen. Das Waldbrandrisiko wird in einer räumlichen Auflösung von etwa 8 km visualisiert.
Für die meisten Waldbrandrisikokarten werden statische Informationen wie Topographie, Vegetationsdichte und -feuchtigkeit aus satellitengestützten Fernerkundungsdaten bereits eingesetzt (MICHAEL et al. 2020). Ebenfalls sind Machine Learning Algorithmen in der Waldbrandprädiktion etabliert (ABID 2021). Dennoch ist die Prädiktion von Waldbränden ausschließlich aus satellitengestützten Fernerkundungsdaten und Machine Learning Modellen ein unbekannteres und neueres Forschungsfeld.
Die veröffentlichte Studie von YANG et al. (2021) erwies erst kürzlich, dass mittels neuronaler Netze und lediglich multispektraler Landsat 7 Daten eine effektive Prädiktion von Waldbränden möglich ist.
Daher wird in der vorliegenden Studie untersucht, ob auf Basis von lediglich multispektralen Sentinel-2 Daten in Kombination mit Machine Learning Modellen eine effektive, kostengünstige und flächendeckende Alternative zur Waldbrandprädiktion geschaffen werden kann.
Bei der Beantwortung dieser Forschungsfrage stehen, neben der eigentlichen Vorhersage, auch eine räumlich hohe Auflösung (< 30 m) der Daten und die globale Anwendbarkeit der Methodik im Vordergrund. Dadurch lassen sich aus den Prädiktionsergebnissen hochauflösende und global verfügbare Waldbrandrisikokarten ableiten.

2 Methodik

In dieser Studie werden verschiedene Machine Learning Modelle zur Vorhersage eines binären numerischen Wertes, der das Auftreten bzw. das Nichtauftreten eines Waldbrandereignisses, repräsentiert, verglichen.
Die Vorhersage des Waldbrandes wird anhand der zeitlichen Entwicklung bzw. der Zeitreihen zweier Vegetationsindizes vor dem Brandereignis getroffen. Hierzu wird ein eigener Eingangsdatensatz, welcher die dynamische Entwicklung der Vegetationsindizes repräsentiert, sowie ein eigener Referenzdatensatz, der ein einzelnes Waldbrandgebiet abbildet, generiert.

Zur Prädiktion des Waldbrandgebiets werden die folgenden drei überwachten Machine Learning Modelle eingesetzt: eine logistischen Regressionsanalyse und zwei künstliche neuronale Netze. Die künstlichen neuronale Netze sind ein Fully Connected Network, mit einer verborgenen Schicht und ein Long Short Term Memory Network, welches eine spezielle Form des Recurrent Neural Networks ist.
Der untersuchte Waldbrand fand im Mai 2021 in der Region Attika und Peloponnes in Griechenland statt. Das gesamte Untersuchungsgebiet umfasst eine Fläche von 386 km², davon sind ca. 51,3 km² Waldbrandgebiet.

2.1 Generierung der Eingangs- und Referenzdaten

2.1.1 Referenzdaten

Die Referenzdaten (engl.: Ground-Truth) enthalten Informationen über die wahren Begebenheiten anhand derer die Machine Learnig Modelle in der Trainingsphase lernen und die vorhergesagten Ergebnisse der Modelle in der Testphase verglichen werden. Die im Referenzdatensatz enthaltenen Daten entsprechen der Definition des Labels in den Machine Learning Modellen. In dieser Studie enthält der Referenzdatensatz Informationen über das tatsächliche Waldbrandgebiet im Untersuchungsraum.
Die Kartierung der verbrannten Fläche und des Verbrennungsgrades des zu untersuchenden Waldbrandereignisses wird anhand des Indizes Normalized Burn Ratio (NBR) bzw. der Differenz von multitemporal aufgenommenen NBRs (dNBR) vorgenommen:
Der Index des NBR (1) macht sich die Eigenschaft zu Nutze, dass gesunde Vegetation eine hohe Reflektion der elektromagnetischen Strahlung im nahinfraroten Wellenlängenbereich (NIR) und eine geringe Reflektion im kurzwelligen infraroten Wellenlängenbereich (SWIR) aufweist, wohingegen kürzlich verbrannte Flächen sich genau umgekehrt verhalten. Sie besitzen eine geringe Reflektion im NIR und eine hohe Reflektion im SWIR. (SERCO ITALIA SPA 2017; UN-SPIDER o.D.)

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} = \frac{B\,8 - B\,12}{B\,8 + B\,12} = \frac{842\,nm - 2190\,nm}{842\,nm + 2190\,nm} \quad (1)^2$$

Die Identifikation einer kürzlich verbrannten Fläche und die Unterscheidung derselben von kahlem Boden und nicht bewachsenen Gebieten sowie die Bestimmung des Verbrennungsgrades geschehen anhand der dNBR. Die dNBR ist die Differenz zwischen der NBR vor dem Waldbrand und der NBR nach dem Waldbrandereignis (2). Der Wertebereich der dNBR liegt in der Regel zwischen -0,5 und 1,3, wobei ein hoher dNBR-Wert ein Indikator einer schweren Verbrennung ist und eine negativer Wert auf das Nachwachsen der Pflanzen nach dem Brand hindeuten könnte (SERCO ITALIA SPA 2017, UN-SPIDER o.D.).

$$dNBR = NBR_{vor\ Brandereignis} - NBR_{nach\ Brandereignis} \quad (2)$$

[2] In der Formel 1,3 und 4 entspricht die Variable B dem verwendeten multispektralen Kanal bzw. Band der Sentinel-2 Satelliten zur Berechnung des jeweiligen Index. Die Wellenlänge der Variable B ist im letzten Term der Formel beschrieben.

Anhand einer schwellenwertbasierten Klassifikation der dNBR wird das Untersuchungsgebiet in ein Waldbrandgebiet (dNBR >= 0,27) und ein vom Waldbrand nicht betroffenes Gebiet (dNBR < 0,27) binär unterteilt. Eine dNBR >= 0,27 entspricht per Definition von UN-SPIDER (o.D.) einem moderaten bis leichtem Verbrennungsgrad.
Die verwendeten Bänder, für die Berechnung der NBR, der Sentinel-2 Satelliten liegen in einer maximalen räumliche Auflösung von 20 m vor. Somit weist der Referenzdatensatz ebenfalls eine räumliche Auflösung von 20 m auf.

2.1.2 Eingangsdaten

Für die Prädiktion des Waldbrandereignisses werden in dieser Studie Zeitreihen von Aufnahmen des Untersuchungsgebiets als Eingangsdaten für die Machine Learning Modelle verwendet. Die Eingangsdaten enthalten Informationen über die zeitlichen Verläufe der Vegetationsdichte und -gesundheit sowie der Vegetationsfeuchtigkeit im Untersuchungszeitraum. Zur Bestimmung dieser Informationen werden zwei Indizes gewählt: der Normalized Difference Vegetation Index (NDVI) und der Normalized Difference Infrared Index (NDII).
Der NDVI (3) wird in dieser Studie als Indikator für den Gesundheitszustand und die Dichte der Vegetation interpretiert. In Anbetracht eines Waldbrandes ist die Vegetationsdichte auch als hypothetische Brennstoffdichte zu werten.

$$NDVI = \frac{NIR - Rot}{NIR + Rot} = \frac{B\,8 - B\,4}{B\,8 + B\,4} = \frac{842\,nm - 665\,nm}{842\,nm + 665\,nm} \tag{3}$$

Der NDII (4) ist ein Indikator für den Feuchtigkeitszustand der Vegetation.

$$NDII = \frac{NIR - SWIR}{NIR + SWIR} = \frac{B\,8 - B\,11}{B\,8 + B\,11} = \frac{842\,nm - 1610\,nm}{842\,nm + 1610\,nm} \tag{4}$$

Die Nutzung von Zeitreihen für die Prädiktion von Waldbränden fußt auf der Annahme, dass sich durch die Betrachtung der Indizes NDVI und NDII über einen Beobachtungszeitraum bzw. die Inklusion dynamischer Informationen die Prädiktion von Waldbrandrisiken verbessert. Die dynamische Betrachtung der Indizes enthält Informationen über die Vegetationsstadien sowie den Trockenheitszustand der Vegetation über die Zeit. Die Vegetationsstadien sind ebenfalls ein Indikator für die verfügbare Menge und Typ des Brennstoffs für den Waldbrand (MICHAEL et al. 2020).
Die Eingangsdaten umfassen Informationen über die Vegetation des Untersuchungsgebiets für die Dauer von etwa einem Jahr vor dem Waldbrandereignis. Die Aufnahmezeitpunkte sind so gewählt, dass für jede Woche in dem Jahr vor dem Waldbrandereignis jeweils ein Indexwert in den Eingangsdatensatz einfließt. Der erste Zeitpunkt in dem Eingangsdatensatz liegt neun Tage vor dem Brandereignis. Somit generieren die Machine Learning Modelle eine theoretische Prädiktion der Labels „Waldbrandgebiet“ und „kein Waldbrandgebiet“ für neun Tage in der Zukunft für das Untersuchungsgebiet.

2.2 Machine Learning Modelle

Für die Vorhersage des Waldbrandgebiets werden eine logistischen Regressionsanalyse (LR), ein Fully Connected Network (FCN) mit einen Hidden Layer und ein Long Short Term Memory

Network (LSTM) anhand des Eingangs- und Referenzdatensatzes trainiert und getestet. Das Ausgabesignal der Modelle ist in dieser Studie als Vorhersage über das Auftreten eines Waldbrandes in einem Gebiet bzw. in einem Pixel zum Zeitpunkt des Waldbrandereignisses (WE) zu interpretieren.
Die logistische Regressionsanalyse ist ein „klassisches“ Model für die Anwendung der binären Klassifikation und wird als linearer Klassifikator bezeichnet. Das bedeutet, dass die logistische Regressionsanalyse sich besonders für die Vorhersage von linear trennbaren Klassen, in diesem Fall der Klassen „Waldbrandgebiet“ und „kein Waldbrandgebiet“, eignet. Ein neuronales Netzwerk, wie das in dieser Arbeit verwendete Fully Connected Network, besteht aus einer Serie von vollständig verbundenen Schichten. Die logistische Regressionsanalyse kann im Zusammenhang mit neuronalen Netzen als Teilmenge bzw. als eine Schicht des neuronalen Netzwerkes verstanden werden. Es ist weitverbreitet eine logistische Sigmoid-Funktion (zu der auch die logistische Funktion gehört) als Aktivierungsfunktion in den versteckten Schichten der neuronalen Netzwerke zu verwenden. Aktivierungsfunktionen definieren den Wertebereich der Ausgabe der jeweiligen Schicht. Neuronale Netzwerke haben den Vorteil gegenüber der logistischen Regressionsanalyse auch komplexere und nicht-lineare Zusammenhänge zu erkennen (MAHAJAN 2020; RASCHKA O.D.; RASHID 2017).
Die Modelle der logistischen Regressionsanalyse und des Fully Connected Network erhalten während der Prozessierung (Training- und Testphase) der Modelle als Input ein Array, in dem das NDVI- und die NDII-Werte je Pixel chronologisch angeordnet sind. Für einen Pixel i ergibt sich so ein 1D-Array mit 100 Input-Werten bzw. Features. Den Modellen ist weder bewusst, dass ein zeitlicher Zusammenhang zwischen den Input-Werten besteht, noch können sie zwischen der Featurekategorie „NDVI“ und „NDII“ unterscheiden. Die Prozessierung der Prädiktion anhand der LR und des FCNs wird in der Abb. 1 veranschaulicht.

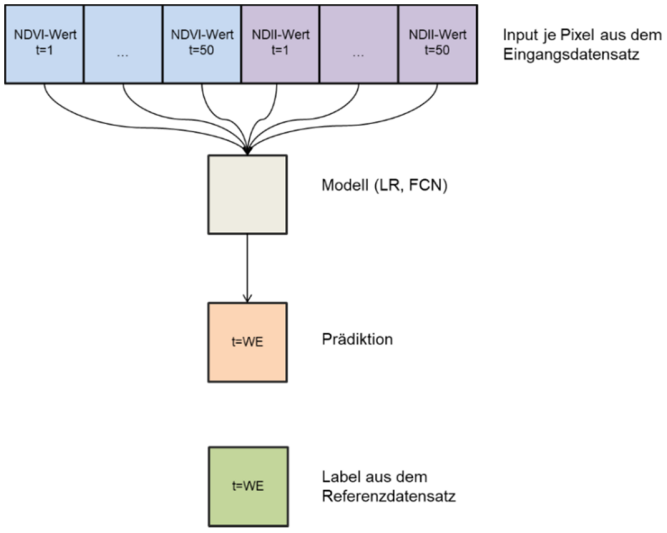


Abb. 1: Prozessierung der Prädiktion anhand der LR und des FCNs

Ein Long Short Term Memory Network (LSTM) wiederum besteht aus einzelnen „Wiederholungsmodulen“, die in einer kettenförmigen Struktur verbunden sind. In jedem Modul befinden

sich vier einzelne neuronale Netzwerkschichten, die miteinander interagieren, und wiederum sigmoidale Aktivierungsfunktionen beinhalten. Das LSTM lernt einen temporalen Zusammenhang aus historischen Daten herzustellen und eignet sich daher für die Prädiktion von Zeitreihen (OLAH 2015; YANG et al. 2021).
Im Vergleich zu der Prädiktion unter Anwendung der LR und des FCNs erhält das LSTM als Input in der Trainings- und Testphase ein 2D-Array je Pixel, welches 50 Zeilen und 2 Spalten aufweist. Die 50 Zeilen repräsentieren die 50 Aufnahmezeitpunkte bzw. die Zeitschritte in der Zeitreihe. Der Ablauf der Prädiktion mit Hilfe des LSTMs wird anhand der Abb. 2 veranschaulicht.

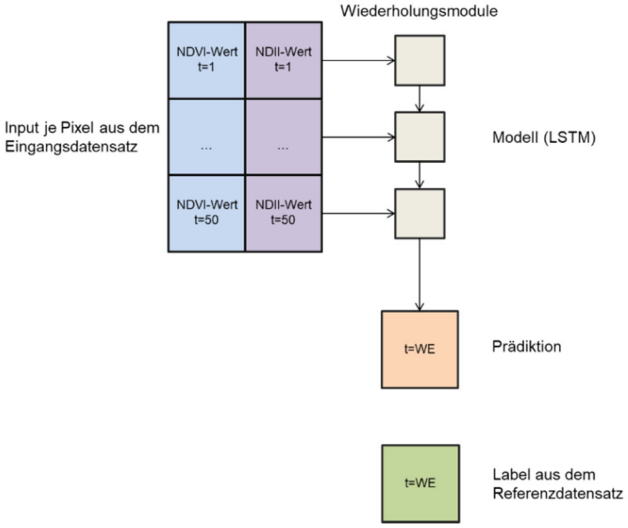


Abb. 2: Prozessierung der Prädiktion anhand des LSTMs

3 Referenz- und Eingangsdatensatz

Der generierte Referenzdatensatz besteht aus einer Aufnahme mit 1181 × 817 Pixeln, wobei jeder Pixel die binäre Information 1 (=Waldbrandgebiet) oder 0 (=kein Waldbrandgebiet) enthält. Jeder Pixel weist ein räumliche Auflösung von 20 m auf.
Der Eingangsdatensatz enthält insgesamt 50 Aufnahmen (eine Aufnahme pro Beobachtungszeitpunkt). Jede Aufnahme besteht, wie der Referenzdatensatz, aus 1181 × 817 Pixeln. Dadurch lässt sich für jeden Pixel eine Zeitreihe erstellen, die den dynamischen Verlauf der NDVI- und NDII-Werte vor dem Waldbrandereignis repräsentiert. Die Pixel weisen ebenfalls eine räumliche Auflösung von 20 m auf.
In der Abb. 3 wird der zeitliche Verlauf der NDVI- und NDII-Werte im Untersuchungsgebiet während des beobachteten Zeitraums grafisch dargestellt. In den Diagrammen wird zwischen dem durchschnittlichen NDVI- und NDII-Werten des Waldbrandgebiets und des Gebiets, das nicht vom Waldbrand betroffen war, unterschieden.

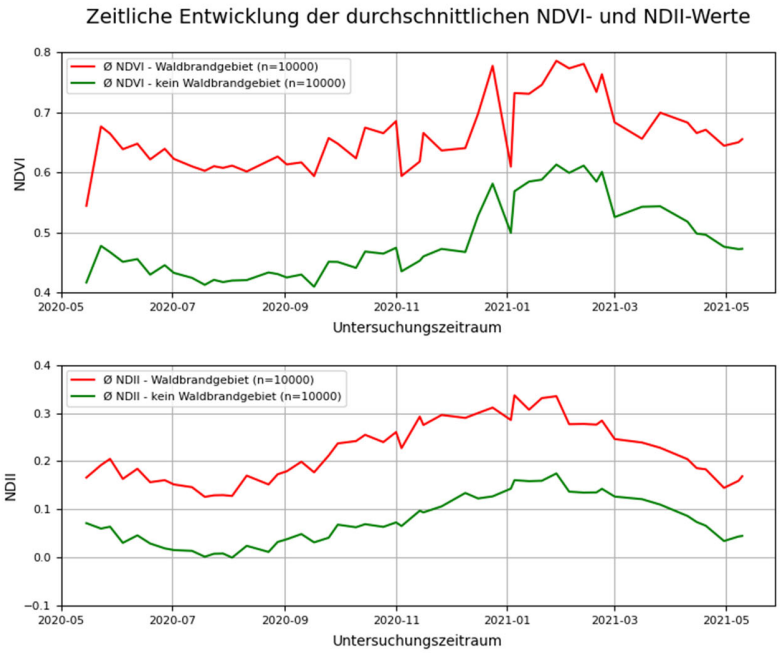


Abb. 3: Zeitliche Entwicklung der durchschnittlichen NDVI- und NDII-Werte im Untersuchungsgebiet

Die Diagramme ermöglichen einen Einblick in den jahreszeitlichen Verlauf der Vegetationsstadien bzw. des Gesundheitszustands der Vegetation und Vegetationsfeuchtigkeit. Hierbei ist zu beobachten, dass die generelle zeitliche Dynamik der NDVI-Werte und der NDII-Werte im Untersuchungszeitraum phasengleich verläuft. Jedoch ist ein deutlicher Versatz auf der Ordinate zwischen den NDVI- und NDII-Werten des Waldbrandgebiets und des nicht betroffenen Gebiets festzustellen.

4 Ergebnisse

4.1 Ergebnisse der Machine Learning Modelle

Für die Bewertung der Leistung der Machine Learning Modelle werden die Leistungskennzahlen (Metriken) Accuracy, Area Under the Curve (AUC), F1-Score und der Recall bestimmt. Vor allem der Recall (True Positive Rate) ist bei der Vorhersage einer Naturkatastrophe von hoher Wichtigkeit. Der Recall gibt den Anteil der tatsächlich positiven Klasse, die korrekt von den Modellen identifiziert werden, wieder. Auf diese Studie bezogen, beantwortet der Recall die wesentliche Fragestellung: „Wie hoch ist der Anteil der tatsächlich von einem Waldbrand betroffenen Gebiete, die als solche korrekt prädiziert wurden?“. Im Umkehrschluss impliziert ein geringer Recall eine hohe False Negative Rate. Eine hohe False Negative Rate wiederum bedeutet, dass viele Gebiete, die tatsächlich von einem Waldbrandereignis betroffen sind, nicht als solche vom Modell prognostiziert werden. Das kann bei der Vorhersage von Waldbränden verheerende Folgen haben.
Die erzielten Metrikergebnisse der Modelle, welche anhand eines Testdatensatzes berechnet werden, sind in Tab. 1 aufgeführt.

Tab. 1: Aufführung der leistungsstärksten Metrikergebnisse der verwendeten Machine Learning Modelle

Modell	Accuracy	AUC	F1-Score	Recall
LR	0,87	0,91	0,64	0,58
FCN	0,88	0,92	0,68	0,67
LSTM	0,92	0,97	0,81	0,81

Die Accuracy bzw. das Verhältnis der Anzahl der korrekt ausgegeben Klassifizierungen des Modells zu der Gesamtanzahl an getätigten Klassifizierungen erreicht bei allen drei Modellen mindestens 87 %. Dies ist auf die sehr präzise Prädiktion (>= 93 %) aller Modelle des vom Waldbrand nicht betroffenen Gebiets zurückzuführen. Die AUC, als Metrik zur Abschätzung der generellen Leistungsfähigkeit eines Klassifikators zu verstehen, liegt bei den verwendeten Modellen ebenfalls in einem hohen Wertebereich von über 0,9. Wesentliche Unterschiede zwischen den Prädiktionsergebnissen zwischen den Modellen werden beim Recall sichtbar. Die logistische Regressionsanalyse und das Fully Connected Network sagen lediglich 58 % bzw. 67 % des tatsächlichen Waldbrandgebiets korrekt voraus bzw. es werden 42 % und 33 % der Klasse „Waldbrandgebiet“ fälschlicherweise als „kein Waldbrandgebiet“ prädiziert.
Das LSTM erreicht für den generierten Eingangs- und Referenzdatensatz für alle berechneten Metriken die leistungsstärksten Ergebnisse. Das vom Waldbrand nicht betroffene Gebiet wird zu 95 % richtig bestimmt und 81 % des tatsächlichen Waldbrandgebiets werden korrekt vom LSTM prognostiziert (Abb. 4).

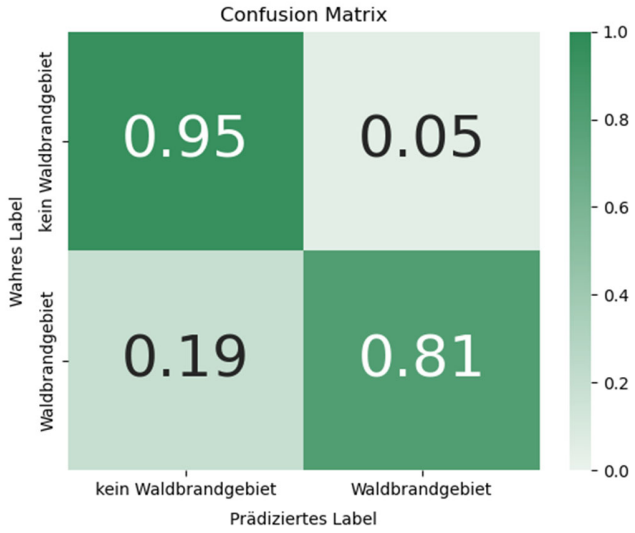


Abb. 4: Confusion Matrix in % des LSTMs

4.2 Binäre Waldbrandprädiktionskarte

Anhand der Prädiktionsergebnisse des LSTMs wird eine binäre Waldbrandprädiktionskarte für das Untersuchungsgebiet erstellt. Dafür wird die ursprüngliche Methodik angepasst. Für eine vollständige Rekonstruktion der ursprünglichen Koordinaten der prädizierten Labels wird für die Testphase des Modells der vollständige Eingangs- und Referenzdatensatz verwendet. Die Vorhersage des Modells geschieht somit hauptsächlich auf Basis von Daten, mit denen es auch trai-

niert wird. Dadurch generiert das Modell zwar genauere Prädiktionsergebnisse, verfehlt allerdings die Aufgabe des Machine Learning Modells eine Prädiktion anhand eines ihm unbekannten Datensatz vorzunehmen.
Der erzielte Recall des LSTMs liegt unter Anwendung dieser Methodik bei 84 %. Dies spiegelt sich in der in Abb. 5 visualisierten binären Waldbrandprädiktionskarte (räumliche Auflösung von 20 m) im Vergleich zum Referenzdatensatz, der das tatsächliche Waldbrandgebiet repräsentiert, wider (Abb. 6).

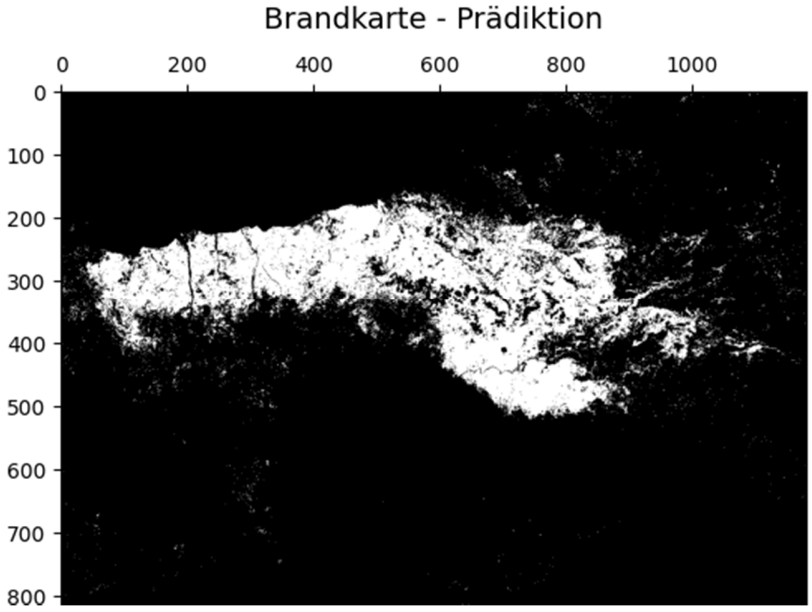


Abb. 5: Binäre Darstellung der Prädiktionsergebnisse. Weiße Fläche entspricht dem prädizierten Waldbrandgebiet.

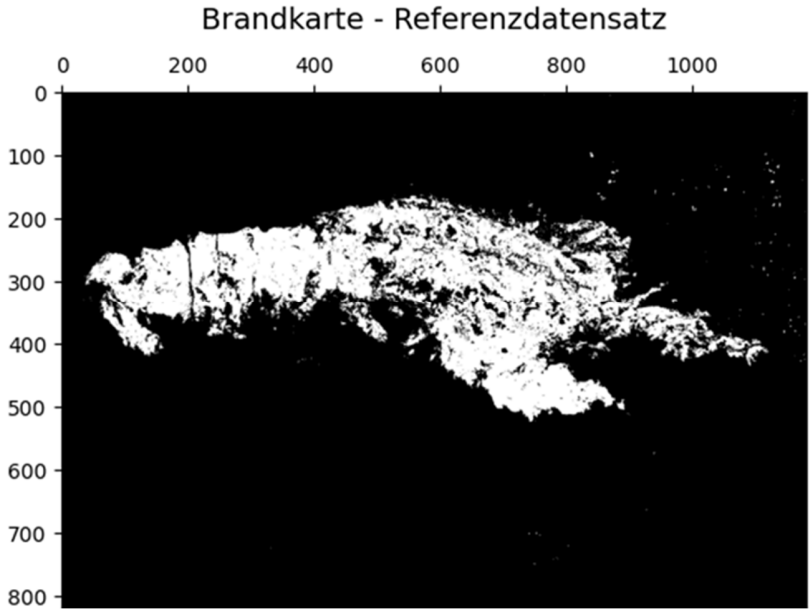


Abb. 6: Binäre Darstellung des Referenzdatensatzes. Weiße Fläche entspricht dem prädizierten Waldbrandgebiet.

5 Diskussion der Ergebnisse

Unter den drei verwendeten Machine Learning Modellen erwies das Long Short Term Memory Network die genausten Prädiktionsergebnisse und bestätigt damit seine Anwendbarkeit auf Zeitreihen. Das LSTM erzielt für die Prädiktion des untersuchten Waldbrandereignisses für einen Testdatensatz ein AUC-Wert von 0,97 und einen Recall von 0,81. Die präzisen Ergebnisse lassen sich vermutlich dadurch begründen, dass eine klare Trennbarkeit der Features (NDVI- und NDII-Werte), die den Klassen „Waldbrandgebiet“ und „keine Waldbrandgebiet“ zugeordnet werden, über den gesamten Beobachtungszeitraum vorliegt (Abb. 3). Diese eindeutige Separation der Features scheint einen positiven Einfluss auf die Prädiktionsfähigkeit der Modelle zu haben. Die Trennbarkeit der NDVI- und NDII-Werte zwischen dem Waldbrandgebiet und dem Gebiet, das nicht vom Waldbrand betroffen ist, lässt sich anhand der unterschiedlichen Vegetationsbedingungen im Untersuchungsgebiet erläutern.
Vergleichbare Studien von MICHAEL et al. (2020) und YANG et al. (2021), die ebenfalls auf Grundlage von satellitengestützten Fernerkundungsdaten (in YANG et al. (2021) mit Multispektral-Daten von Landsat 7, in MICHAEL et al. (2020) mit Zeitreihen von Vegetationsindizes und Topographiedaten) und Machine Learning Modellen versuchen Waldbrände zu prognostizieren, erzielen mit ihren leistungsstärksten Modellen einen AUC-Werte von 0,81 (YANG et al. (2021)) und 0,8 (MICHAEL et al. (2020)). Zwar ist ein direkter Vergleich der vorliegenden Studie zu dieser Arbeit aufgrund der unterschiedlichen methodischen Ansätzen nicht zulässig, dennoch kann der sehr hohe erreichte AUC-Wert des LSTMs in dieser Studie als vielversprechender Ansatz für tiefergehende Bestrebungen in der Waldbrandprädiktion gewertet werden.

6 Fazit & Ausblick

In der vorliegenden Studie wird eine Methodik vorgestellt, um aus Zeitreihen, die aus multispektralen satellitengestützten Fernerkundungsdaten der Sentinel-2 Mission generiert werden, sowie der Implementierung verschiedener Machine Learning Modelle eine Waldbrandprädiktion anhand eines Waldbrandereignisses vorzunehmen.
Das leistungsstärkste LSTM Modell sagt für das untersuchte Waldbrandereignis 81% der Labels der Klasse „Waldbrandgebiet“ anhand des eigens erstellten Referenz- und Eingangsdatensatz korrekt voraus. Zudem kann durch die Verwendung eigener Datensätze eine Waldbrandprädiktion in einer räumlichen Auflösung von 20 m bzw. eine hochauflösende Waldbrandprädiktionskarte (Abb. 5) generiert werden.
Ebenfalls kann durch die Beobachtung der zeitlichen Entwicklung der Vegetationsindizes in den Zeitreihen, der Einfluss des Zustandes und die räumliche Verteilung der Vegetation auf das Waldbrandrisiko verdeutlicht werden. Zudem wird eruiert, dass ein Potential besteht eine Waldbrandprädiktion vorzunehmen, ohne auf weitere meteorologische und topographische Daten zugreifen zu müssen.
Die erarbeitete Methodik lässt sich aufgrund der globalen Landflächenabdeckung und der kostenlosen Bereitstellung der Sentinel-2 Produkte sowie der freien Verfügbarkeit von Software-Bibliotheken für maschinelles Lernen und Deep Learning wie z.B. Scikit-learn und Keras einfach reproduzieren und für Untersuchungsgebiete weltweit einsetzen.

Dennoch muss die Allgemeingültigkeit der Ergebnisse der verwendeten Methodik validiert werden. So müsste unter anderem die Methodik auf weitere Waldbrandereignisse und Untersuchungsgebiete angewendet, evaluiert und automatisiert werden. In Hinblick auf den Referenzdatensatz müsste validiert und erprobt werden, ob eine ausreichend präzise Prädiktion für einen weiter als neun Tage in der Zukunft gelegenen Zeitraum möglich ist.

7 Literaturverzeichnis

ABID, F., 2021: A Survey of Machine Learning Algorithms Based Forest Fires Prediction and Detection Systems. Fire Technology, **57**, 559-590, https://doi.org/10.1007/s10694-020-01056-z.

MAHAJAN, P., 2021: Fully Connected vs Convolutional Neural Networks. https://medium.com/swlh/fully-connected-vs-convolutional-neural-networks-813ca7bc6ee5, letzter Zugriff am 15.09.2021.

MICHAEL, Y., HELMAN, D., GLICKMAN, O., GABAY, D., BRENNER, S. & LENSKY, I., 2020: Science of the Total Environment: Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series. Science of The Total Environment, **764**, 142844, https://doi.org/10.1016/j.scitotenv.2020.142844.

OLAH, C., 2015: Understanding LSTM Networks. Colah's Blog. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

RASCHKA, S., o.D.: What is the relation between Logistic Regression and Neural Networks and when to use which? https://sebastianraschka.com/faq/docs/logisticregr-neuralnet.html, letzter Zugriff am 17.10.2021.

RASHID, T., 2017: Neuronale Netze selbst programmieren - Ein verständlicher Einstieg mit Python. O'Reilly.

SERCO ITALIA SPA, 2017: Burned Area Mapping with Sentinel-2 (SNAP). RUS Lectures, https://rus-copernicus.eu/portal/the-rus-library/learn-by-yourself/

UN-SPIDER, o.D.: Normalized Burn Ratio (NBR). UN-SPIDER Knowledge Portal, https://un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio, letzter Zugriff am 06.10.2021.

HIRSCHBERGER, P., 2016: Forests Ablaze: Causes and Effects of Global Forest Fires. Winter, S., von Laer, Y., Köberich, T., Eds., WWF Deutschland: Berlin, Germany, https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF-Study-Forests-Ablaze.pdf.

YANG, S., LUPASCU, M. & MEEL, K.S., 2021: Predicting Forest Fire Using Remote Sensing Data And Machine Learning. 35th AAAI Conference on Artificial Intelligence, arXiv preprint arXiv:2101.01975.

Beschreibung der geometrischen Qualität eines Digitalen Geländemodells des Wasserlaufs

ROBERT WEIß[1] & SILKE MECHERNICH[1]

Zusammenfassung: Digitale Geländemodelle werden vielfältig erstellt und eingesetzt. Je nach Anwendungsfall werden Geländemodelle unterschiedlich modelliert und den Modellen liegen unterschiedliche Daten zugrunde. Für die Wasserstraßen- und Schifffahrtsverwaltung des Bundes (WSV) werden Digitale Geländemodelle des Wasserlaufs (DGM-W) erstellt. Diese zeichnen sich dadurch aus, dass neben der Topografie an Land auch der Gewässerboden detailliert dargestellt wird. Neben der Topografie als solche sind auch Angaben zur Qualität entscheidend. Eine Beschreibung der Qualität der DGM-W ist vielfältig möglich. Neben den Eigenschaften des DGM-W (TIN, GRID, Rasterweite, ...) und der damit verbundenen Modellierungsvorschrift, sind auch die Eigenschaften der eingehenden 3D-Messpunkte entscheidend. Dabei ist neben der Sensorart auch die Unsicherheit der eingehenden 3D-Messpunkte, die Punktdichte, Punktverteilung und Footprintgröße von großem Interesse. Weitere mögliche Qualitätsmerkmale sind (i) Abweichungen der Modellhöhen gegenüber Kontrollpunkten und -flächen, (ii) die Unsicherheiten, die sich aus den Ausgleichungen ergeben sowie (iii) statistisch aufbereitete Abweichungen der 3D-Punktwolke gegenüber der modellierten Oberfläche. In Summe ergeben sich damit vielfältige Parameter, mit denen die geometrische Qualität eines DGM-W beschrieben werden kann.

1 Einleitung und Motivation

Für eine Vielzahl verschiedenster Anwendungen werden Digitale Geländemodelle (DGM) benötigt. In der Regel sind DGMs auf trockene Flächen beschränkt bzw. Wasserflächen wie etwa Seen oder Flüsse werden durch einfache Interpolationen geschlossen.

In Deutschland ist die Wasserstraßen- und Schifffahrtverwaltung des Bunds (WSV) für die Sicherheit und Leichtigkeit des Verkehrs auf Bundeswasserstraßen zuständig. Dies beinhaltet z.B. auch das Vorhalten einer Fahrrinne mit entsprechenden Tiefen in Abhängigkeit vom Abflussgeschehen. Dementsprechend werden Geländemodelle benötigt, die neben trocken gefallenen Flächen an Land auch präzise Informationen zu der Topografie unterhalb der Wasseroberfläche inklusive aller hydraulisch relevanten Bauwerke beinhalten. Diese Modelle werden als Digitalen Geländemodelle des Wasserlaufs (DGM-W) bezeichnet. Zudem sind DGM-Ws ein wesentliches Basisprodukt zur Bearbeitung gewässerkundlicher, wasserwirtschaftlicher, wasserbaulicher und ökologischer Aufgabenstellungen an Bundeswasserstraßen.

Die Bundesanstalt für Gewässerkunde (BfG) ist eine dem Deutschen Bundesministeriums für Digitales und Verkehr (BMDV) nachgeordnete wissenschaftliche Oberbehörde, deren Aufgabe unter anderem darin besteht, die WSV fachtechnisch zu beraten. In der Vergangenheit wurden die DGM-

[1] Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, D-56068 Koblenz, Deutschland, E-Mail: [weiss, mechernich]@bafg.de

Ws primär durch die BfG erstellt. Zukünftig wird die WSV auch eigenständige DGM-Ws generieren, wobei die BfG Ihre Kompetenzen zukünftig in Form von fachlichen Beratungen und vor allem Weiterentwicklungen einbringen wird.
Neben der reinen Geometrie der modellierten Topografie ist die Qualität der Geländemodelle ein wichtiges Kriterium, das für die weiteren Anwendungen relevant ist. Die Beschreibung der Qualität von DGM-W ist relativ komplex, da bei einem DGM-W die Daten verschiedenster Sensoren und weiterführender Informationen fusioniert werden. Darüber hinaus bedingen auch die Modellierungsvorschriften die Qualität eines DGM-W.
Im Folgenden soll beschrieben werden, was unter einem DGM-W zu verstehen ist, wie es in groben Zügen erstellt wird und damit einher gehen auch Angaben, wie die Qualität eines DGM-W beschrieben werden kann.

2 Modellverständnis

Die DGM-Ws als solche werden vielfältig genutzt, aber es werden dabei relativ wenig Aussagen zur Qualität gemacht. Für eine Beschreibung der Qualität sind verschiedenste Aspekte zu berücksichtigen. An erster Stelle steht dabei das Verständnis darüber, was ein DGM-W eigentlich darstellt bzw. repräsentiert.
Ein Modell soll einerseits die Wirklichkeit approximieren, andererseits aber auch die Realität soweit vereinfachten, damit eine Nutzung für verschiedenskalige Anwendungen möglich ist. Im Fall eines Geländemodells und damit auch eines DGM-W erfolgt die Approximation der Realität als eine Erfassung und Modellierung der Topografie oberhalb und unterhalb der Gewässeroberfläche. Die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) definiert ein Digitales Geländemodell als ein numerisches, auf ein regelmäßiges Gitter reduziertes Modell der Geländehöhen und -formen (ADV 2022). Im Allgemein werden diese als GRIDs oder Rastermodelle bezeichnet. Davon abweichend können DGM auch durch unregelmäßige verteilte Geländepunkte repräsentiert werden. In der Regel sind die Geländepunkte hierbei noch durch Dreiecksvermaschungen mit einander verbunden und werden als Triangulated Irregular Network (TIN) bezeichnet (Abb. 1). Üblicherweise sind morphologische Strukturelemente (z. B. Geländekanten, markante Höhenpunkte) in TINs detailliert abgebildet. Ein DGM beschreibt folglich die Erdoberfläche, frei von Vegetation oder Gebäuden. Sofern diese beiden Elemente in einem Modell enthalten sind, wird von einem Digitalen Oberflächenmodell (DOM) gesprochen (Abb. 2).

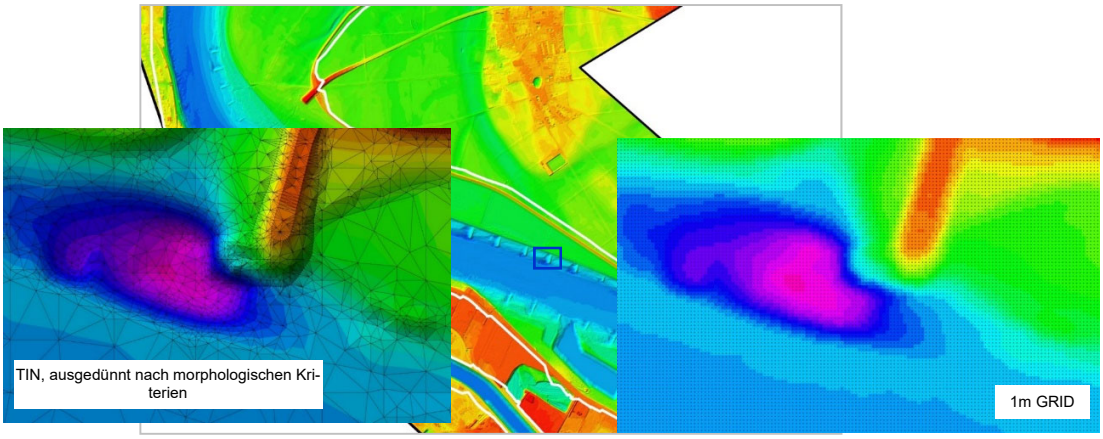


Abb. 1: DGM-W am Niederrhein mit TIN und GRID Darstellung

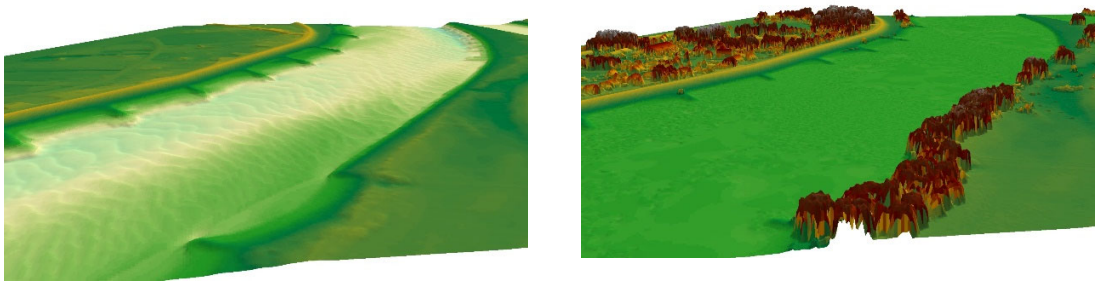


Abb. 2: DGM-W (links) und DOM (rechts) der Elbe bei Geesthacht, El-km 587

Die Modelle werden durch regelmäßig oder unregelmäßig verteilte Modellstützpunkte realisiert, wobei die Modellstützpunkte aus 3D-Punktwolken abgeleitet werden, die wiederum auf unterschiedlichen Sensoren bzw. Erfassungstechniken basieren. Neben der 3D-Punktwolke als solche ist auch die Modellierungsvorschrift von entscheidender Bedeutung. Ein DGM-W kann somit als eine Menge von Basisdatensätzen zzgl. einer Interpretationsvorschrift für jeden Basisdatensatz verstanden werden (SELLERHOFF 2011).

3 Ausgangsdaten einer DGM-W Modellierung

Die DGM-Ws basieren im Gegensatz zu einem konventionellen DGM neben konventionellen Airborne Laserscanning (ALS) Beobachtungen auch auf hydroakustisch erfassten 3D-Messpunkten, auf Ersatzmodellen, sowie z.T. auch auf bathymetrischen ALS-Beobachtungen. Hierbei ist zu beachten, dass den Messmethoden gänzlich unterschiedliche physikalische Messgrößen zugrunde liegen und somit die Eigenschaften der erfassten 3D-Messpunkte unterschiedlich sind. Üblicherweise werden neben den eigentlichen ALS-Daten auch Luftbilder erfasst. Diese erreichen nicht die Qualität von Produkten einer originären Luftbildbefliegung. In der Regel kommen hierbei Mittelformatkameras zu Einsatz. Die Ergebnisse dienen primär der Klassifizierung und Plausibilisierung. Insofern fließen Luftbildinformationen nur indirekt in eine DGM-W Modellierung ein.

3.1 Hydroakustisch erfasste 3D-Messpunkte

Hydroakustisch erfasste 3D-Messpunkte basieren auf unterschiedlichen Echolotsystemen. In der heutigen Zeit kommen in der Mehrzahl Fächerecholotsysteme (Multi Beam Echo Sounder - MBES) zum Einsatz. Diese bestehen aus einem bis drei Schallwandlern (Transducern), die entsprechend getriggert Einzelmessungen mit einem bestimmten zeitlichen Abstand (Lotfolge) ausführen. Eine Messung eines Transducers wird im Allgemeinen als „Ping“ bezeichnet. Bei einem Ping wird ein, quer zur Fahrtrichtung ungerichteter, Schallimpuls mit einer Frequenz in der Größenordnung von 200-600 kHz in das Wasser emittiert. Die Schallsignale werden am Gewässerboden reflektiert und die Echos im Transducer empfangen. Die Fächerlotsysteme sind dabei in der Lage, das Rückkehrsignal richtungsspezifisch aufzuteilen bzw. das Rückkehrsignal in einzelne, winkelabhängige Signale (Beamwinkel) aufzuteilen. Jedes Einzelsignal wird als „Beam“ bezeichnet, ist über einen Beamwinkel und eine Streckeninformation auf Basis von Laufzeitmessungen charakterisiert und weist etwa eine Beamdivergenz von 0,5° – 1,5° auf. Diese ist abhängig von dem Fächerlot und den eingestellten Messparametern. Aus jedem Beam lässt sich folglich eine Koordinate im System des Transducers ableiten, wobei moderne Fächerlotsysteme den aufgespannten Fächer in ca. 500 - 1000 einzelne Beamwinkel aufteilen. Die üblicherweise in der WSV eingesetzten Transducer sind der Lage, einen Fächer mit einer Spannweite von 120°-140° aufzuspannen. Einer Kombination von zwei Transducern wird aus „dual head“ System bezeichnet. In diesem Fall werden die Transducer geneigt installiert, damit ein max. Beamwinkel von 70°-75° relativ zur Schiffshochachse erfasst werden kann.
Alternativ kommen noch sogenannte Einstrahlecholote (Single Beam Echo Sounder - SBES) zum Einsatz. Im Gegensatz zu den Fächerloten wird hier jedem emittierten Ping ein Echo und damit auch nur eine Koordinate in der Z-Achse des Systems zugeordnet. Die Beamdivergenz entsprechender Systeme ist abhängig von der Schallfrequenz sowie dem Transducer und liegt in der Größenordnung mehrerer Grad. Eine getriggerte Kombination von mehreren SBES Systemen wird als „Multi Channel Echo Sounder“ (MCES) – System bezeichnet. Die einzelnen Transducer werden hierbei üblicherweise in einem Abstand verbaut, der in dem geplanten Messrevier zu einer flächenhaften Verteilung der Beobachtungen führt (z.B. 0,5 m). Die prinzipielle Messanordnung ist in Abb. 3 dargestellt. In Abb. 4 sind die resultierenden Punktmuster am Boden dargestellt.

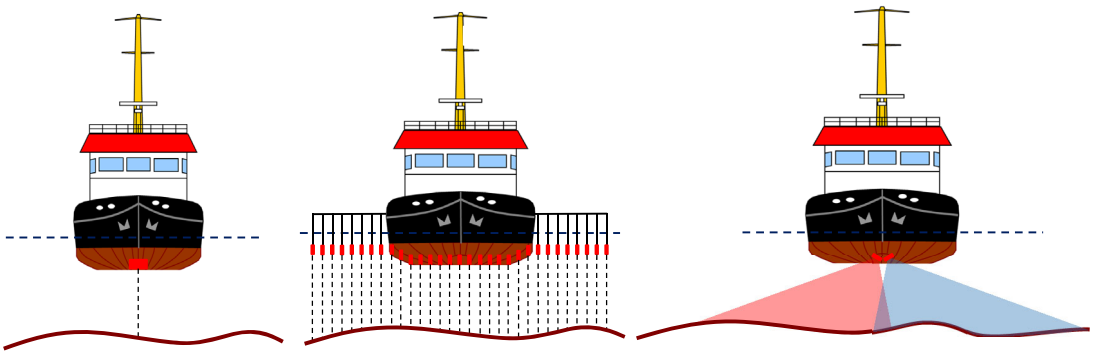


Abb. 3: SBES (links), MCES (Mitte) und (dual-head) MBES (rechts) - Systeme

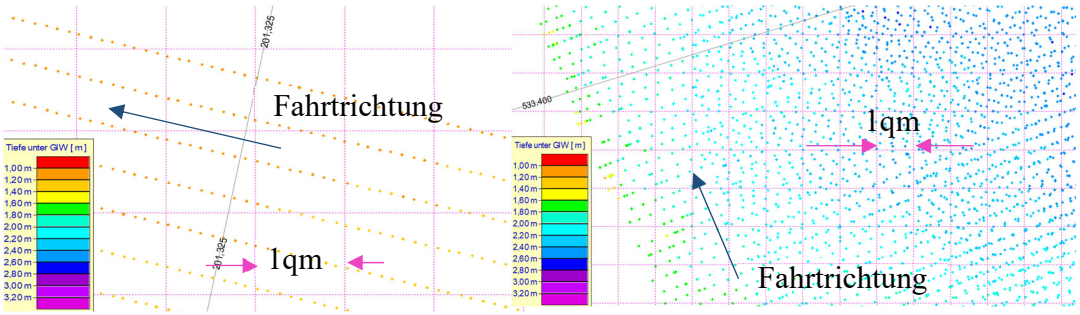


Abb. 4: Erfassungsmuster eines MCES System (links) und einem MBES System (rechts)

Die Lotfolgen liegen in der Größenordnung von 10 Hz-20 Hz und sind neben dem Lotsystem auch von der Wassertiefe und den Systemeinstellungen abhängig. Bei einer Geschwindigkeit von 5 m/s über Grund ergeben sich beispielsweise Abstände zwischen von 0,25 m – 0,5 m.

3.2 ALS-basierte 3D-Punktwolken

Der Standard für das Vorland und trocken gefallene Gebiete ist eine konventionelle ALS Erfassung. Dabei wird mit nahem Infrarot (NIR) Laserlicht die Erdoberfläche abgetastet. Im Gegensatz zu hydroakustischen Verfahren wird dabei jedem emittierten Laserpuls genau ein Rückkehrimpuls zugeordnet. Dieser Rückkehrimpuls wird in der Regel in einzelne Echos aufgeteilt, wobei jedes Echo einen Reflektionsschwerpunkt entlang des Signalwegs darstellt. Die ALS Technik ist in der Lage innerhalb bestimmter Grenzen die Vegetation zu penetrieren bzw. durch diese hindurch zu messen.

Neben der konventionellen ALS Technik wird in jüngster Vergangenheit zunehmend die bathymetrische ALS Technik eingesetzt. Im Gegensatz zu NIR Technik kommt hierbei grünes Laserlicht zum Einsatz, welches Wasser durchdringt. Ein Teil des emittierten Laserpulses wird an der Wasseroberfläche reflektiert. Der nicht reflektierte Teil dringt in den Wasserkörper ein und wird anteilig am Gewässerboden reflektiert. Sofern das verbleibende Signal am Gewässerboden reflektiert wird und dieses Echo im empfangenen Rückkehrsignal detektierbar ist, lässt sich die Position des Gewässerbodens ableiten.

Für ein neues DGM-W der Elbe wurde beispielsweise 2018 eine flächendeckende Befliegung der Elbe zwischen der Grenze zur Tschechische Republik und dem Elb-km 587 mit einem bathymetrischen ALS-System durchgeführt (KUEHNE 2021). Diese fand während des anhaltenden Niedrigwassers statt. Zum Zeitpunkt der Befliegung war das Wasser sehr klar bzw. die Gewässertrübung sehr gering, so dass große Bereiche des Gewässerbodens erfasst wurden.

3.3 Footprintgröße und Punktdichten

Allen Erfassungssystemen ist gemein, dass die 3D-Messpunkte zwar punktuell erfasst werden, die abtastenden Sensoren allerdings eine mehr oder weniger großen Footprint haben. In der Regel werden die Koordinaten der 3D-Messpunkte in cm oder mm-Auflösung gespeichert. Die Auflösung ist an dieser Stelle nicht mit der Genauigkeit oder repräsentierten Fläche (Footprint) zu verwechseln. Bei einem Fächerlotsystem mit einer Beamdivergenz von 1° ergibt sich bei einer Wassertiefe von 3 m im Nadir eine Footprintgröße von gut 5 cm. Bedingt durch die schräge Projektion

steigt dieser mit dem Beamwinkel sehr stark an. Bei einem Beamwinkel von 70° ergibt sich bei gleicher Wassertiefe ein Footprintdurchmesser von rund 45 cm. Prinzipiell ähnlich verhält es sich bei ALS Sensoren. Bei konventionellen (NIR) ALS-Systemen liegen die Beamdivergenzen in der Größenordnung von einigen 1/10 mrad (RIEGL VQ780, RIEGL 2022). Bei einer Flughöhe von 1000 m ü.G. und einer Beamdivergenz von 0,18 mrad ergeben sich Footprintgrößen von 0,18 m (Nadir) bis 0,24 m (30° Scanwinkel). Wesentlich größere Beamdivergenzen von etwa 0,8 mrad bis 3 mrad weisen topobathymetrische Scanner auf. Bei einer typischen Flughöhe von 400 m, einem Scanwinkel von 21° sowie einer Beamdivergenz von 0,8 mrad ergeben sich Footprintdurchmesser in der Größenordnung von 0,3 m.
Die Punktdichten variieren je nach Vorgabe und Messkonfiguration. Im Bereich der Wasserstraßen und Schifffahrtsverwaltung des Bundes sind Punktdichten von etwa 20-30 Pkt./m² unter Wasser die Regel, wobei die Punktverteilung stark vom Erfassungssystem abhängt.
Bei ALS-Systemen werden bei großräumigen Projekten i.d.R. geringere Punktdichten erreicht. Aktuelle Befliegungen fordern Punktdichten von mind. 8 Pkt./m², wobei sich dieser Wert auf Gelände/Oberflächenpunkte bezieht bzw. es werden nur die „Last Pulse“-Punkte betrachtet.

3.4 Unsicherheiten der 3D-Messpunkte

Für die Unsicherheitsbetrachtung von Messwerten steht mit dem „Guide to the Expression of Uncertainty in Measurement“ (GUM) ein international anerkannter Leitfaden zur Ableitung von kombinierten Unsicherheitsangaben (z.B. „Total Propagated Uncertainty“ (TPU) bzw. Kombinierte Standardunsicherheiten (KSU)) zur Verfügung. Der GUM beschreibt, wie aus den Unsicherheiten der Eingangsgrößen kombinierte Standardunsicherheiten berechnet werden (MIEKE 2020).
Sowohl bei der hydroakustischen Datenerfassung, als auch bei der ALS-Datenerfassung kommen abtastende Sensoren zum Einsatz. Beide Verfahren basieren darauf, dass Koordinaten der Geländeoberfläche oberhalb und unterhalb der Wasseroberfläche im Bezugssystem des Sensors erfasst werden. Dem folgt eine Rotation/Translation unter Nutzung der Einbaukoordinaten und Einbauwinkel in das Koordinatensystem des Sensorträgers (Peilschiff oder Luftfahrzeug) und eine weitere Rotation/Translation unter Nutzung der erfassten Orientierungswinkel und Positionen der Ortungssensoren in das Zielsystem (z.B. ETRS89/DREF91.R16).
Wird das mathematische Modell aufgestellt, ergeben sich die Koordinaten als Funktion der eigentlichen Sensormesswerte, der geometrischen Beziehungen innerhalb des Sensorträgers sowie der äußeren Orientierung und Positionierung. Die geometrischen Beziehungen innerhalb des Sensorträgers beinhalten in erster Linie die Hebelarme und Einbauwinkel der Sensoren. Nach GUM sind die Sensitivitätsfaktoren der jeweiligen Eingangsgrößen zu ermitteln und zusammen mit deren Standardabweichungen zu Gesamtunsicherheiten in Lage und Höhe zu fusionieren (MIEKE 2020).
Bei den Standardabweichungen ist zu beachten, dass neben den direkten Unsicherheitsangaben (z.B. Unsicherheit einer Winkelmessung) auch indirekte Einflüsse wie etwa Signalverzögerungen, Uhrenfehler oder die Unsicherheit von Nullablagen zu beachten sind.
Die Verteilung der Unsicherheiten ist dabei abhängig vom Abstand zum Nadir. Die Unsicherheiten eines hydroakustischen Standard-Messsystems sind exemplarisch in Abb. 5 dargestellt, wobei die Werte auf eine Wassertiefe von 4 m und einem max. Beamwinkel von 75° gegenüber dem Nadir

bezogen sind. Weitere wesentliche Parameter sind die Unsicherheiten der Einbaukoordinaten auf dem Sensorträger von 5 mm, Nullablagen der Einbauwinkel von 0,02°, die Unsicherheiten der Orientierungswinkel von 0,05° und Unsicherheiten der kinematischen GNSS-Koordinaten von 0,02 m / 0,04 m in Lage / Höhe. Alle Sensoren sind bei diesem Beispiel an einem Mast montiert.

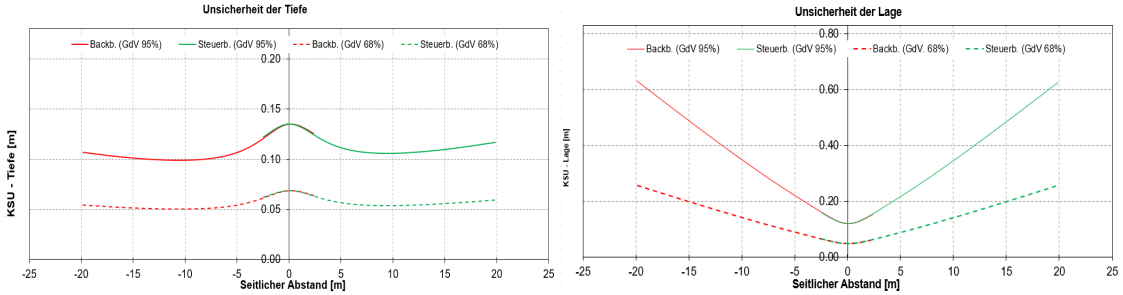


Abb. 5: Kombinierte Standardunsicherheiten (KSU) der Tiefe (links) und Lage (rechts) eines Fächerlotmesssystems

In vielen hydrografischen Softwaresystemen ist eine individuelle Berechnung der TPU/KSU auf Basis der tatsächlich installierten Sensoren, Geometrien und individuell anliegenden Parametern möglich bzw. realisiert. Im Ergebnis können zu jeder Koordinate gehörende Unsicherheitsangaben mit einem entsprechenden Grad des Vertrauens mit ausgegeben werden.

4 Geländemodellierung

4.1 Allgemeines

Auf Basis der 3D-Messpunkte erfolgt die Modellierung von Modellstützpunkten. Diese können als ein regelmäßiges Raster von Modellstützpunkten (GRID) oder als unregelmäßiges Stützpunktfeld mit Verknüpfungsinformationen (TIN) definiert werden, wobei bei einer TIN Modellierung auch häufig Bruchkannten eingesetzt werden.
Mit einer Modellierungsvorschrift wird definiert, um welche Modelle es sich handeln soll (TIN oder GRID) und wie die Modelle aus den 3D-Punktwolken abgeleitet werden.

4.2 TIN-Modellierung

Wird ein TIN gefordert, müssen seitens der Modellierung Vorgaben gemacht werden, welche Eigenschaften die Dreiecke haben sollen. Dabei ist vor allem neben der max. Kantenlänge und der Vorgabe einer Delaunay-Vermaschung entscheidend, bis zu welchem Abstand 3D-Messpunkte oberhalb oder unterhalb der Modelloberfläche liegen dürfen, ohne eine weitere Vermaschung zu generieren. Außerdem sind bei einer Ausdünnung von 3D-Messpunkten bei der DGM-W Modellierung hydraulische Kriterien, wie etwa hydraulisch relevante Bruchkannten zu berücksichtigen.
In vielen Fällen werden aus TINs durch eine Interpolation GRID-Stützpunkte abgeleitet. Dabei ist aber zu beachten, dass der tatsächliche Informationsgehalt des GRIDs immer auf dem zugrunde liegenden TIN basiert. Eine hohe Auflösung spiegelt bei diesem Ansatz nicht immer eine hohe Informationsdichte wieder.

4.3 Direkte GRID-Modellierung

Im hydrografischen Bereich der WSV ist es üblich, aus den 3D-Punktwolken direkt GRIDs mit einem definierten Abstand der GRID-Punkte (Rasterweite) abzuleiten. Hierzu werden für jeden Stützpunkt im Rahmen einer vermittelten Ausgleichung Parameter für Flächenpolynome verschiedenen Grades nach dem Verfahren der Minimierung der kleinsten Quadrate geschätzt (LORENZ et al. 2021). Die so ermittelte Polynomfläche passt sich damit bestmöglich in die 3D-Punktwolke ein und an der Position des Modellstützpunkts entspricht die Höhe der Polynomfläche (Abb. 6) der Höhe des Modellstützpunktes.

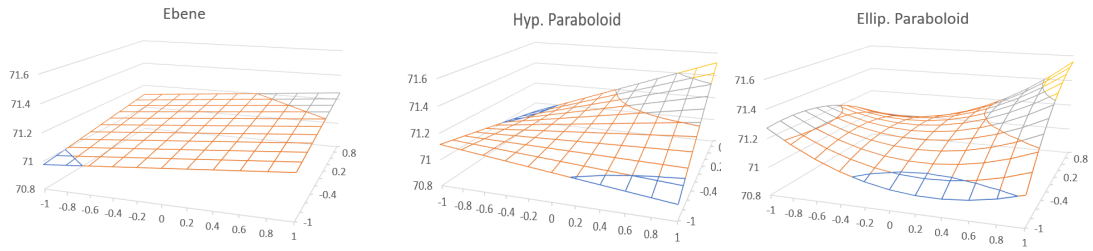


Abb. 6: Polynomfunktionen unterschiedlichen Grades (links: Ebene (3 Unbekannte), Mitte: hyperbolisches Paraboloid (4 Unbekannte), Rechts: Elliptisches Paraboloid (6 Unbekannte))

Grundlage für die Parameterschätzung je Modellstützpunkt ist eine definierte Punktschar. Für die Punktselektion können Einflussradien oder rasterbasierte Methoden zum Einsatz kommen. Punktselektion bzw. definierte Umkreisradien müssen gegenüber der Rasterweite so abgestimmt sein, dass einerseits keine wesentlichen Informationen verloren gehen und andererseits genügend Beobachtungen für eine sichere Modellierung vorhanden sind. Alternativ zum Umkreisradien ist es auch möglich, alle Punkte innerhalb einer definierten Rasterzelle bzw. je nach Vorgabe auch die umliegenden Rasterzellen zu nutzen.
Entscheidend ist in beiden Fällen, dass die Anzahl der eingehenden Beobachtungen wesentlich größer als die Anzahl der Unbekannten ist. Damit geht einher, dass die für dieses Verfahren benötige Punktdichte relativ groß sein muss bzw. wenige Pkt./m² nicht ausreichend sind.
Für die Gewichtung im Rahmen der Ausgleichung werden unterschiedliche Ansätze genutzt. Häufig kommen abstandsgewichtete Ansätze zum Einsatz. Das Gewicht des jeweiligen 3D-Messpunktes verhält sich dabei indirekt proportional zum Abstand des Punktes zum Modellstützpunkt. Alternativ sind auch Gewichtungen auf Basis von TPU möglich. Exemplarisch sind als Ergebnisse in (Abb. 7) GRIDs mit unterschiedlichen Rasterweiten dargestellt. Es zeigt sich, dass der Informationsgehalt mit zunehmender Rasterweite abnimmt.

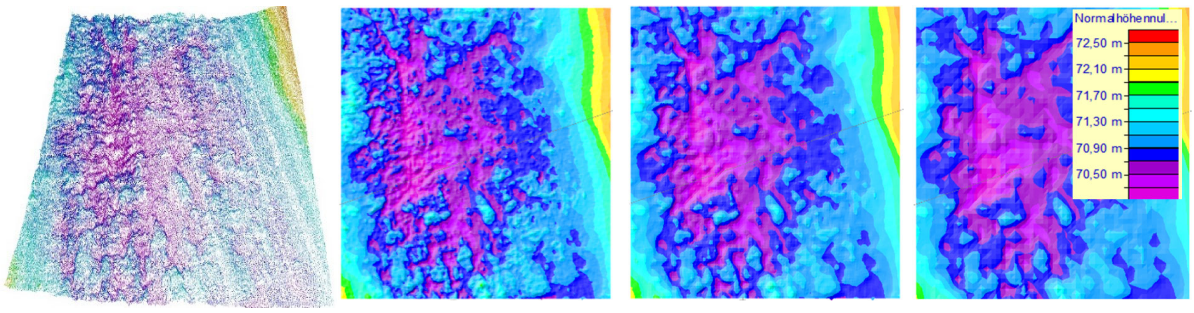


Abb. 7: v.l.n.r: 3D-Punktwolke, GRID mit 0.5 m, 1.0 m, 2.0 m Rasterweite

5 Betrachtung der geometrischen Qualität

5.1 Modell vs. Kontrollpunkte

Zur Abschätzung der resultierenden Genauigkeit ist es möglich, auf Basis der 3D-Punktwolke an den exakten Positionen der Kontrollpunkte einen Modellstützpunkt nach der für das DGM-W gültigen Modellierungsvorschrift zu modellieren. Die resultierende Modellstützpunkthöhe kann dann der Höhe des Kontrollpunktes gegenübergestellt werden.

Exemplarisch sind die Ergebnisse dieser Prüfung für eine Befliegung im Bereich des Jungferngrunds am Rhein mit dem Leica SPL100 ALS System in Abb. 8 dargestellt. Es zeigt sich, dass die Abweichungen der aus der 3D-Punktwolke modellierten Modellstützpunkte zu den Höhen der Kontrollpunkte in der Größenordnung weniger Zentimeter liegen. Von 12 Kontrollpunkten weichen in diesem Fall 11 Punkte um weniger als 3 cm von der Kontrollpunkthöhe ab. Dabei ist zu beachten, dass trotz der Einhaltung aller vermessungstechnischen Grundsätze und Mehrfachbestimmung die Unsicherheit der GNSS-Höhenbestimmung von RTK Punkten in der Größenordnung von 2 -3 cm (ZSS, 2022) bei einem Grad des Vertrauens von 68,3% (BW, 2022) liegt. Somit ist nicht sicher nachweisbar, ob die Ursache der Abweichungen auf Seiten des ALS-Systems oder auf Seiten der Kontrollpunkte zu suchen sind.

Sofern an der Position des Kontrollpunktes kein Modellstützpunkt modelliert wird, ist es dennoch möglich, die vertikale Abweichung der Modelloberfläche zum Kontrollpunkt zu ermitteln und weiter auszuwerten.

Abb. 8: Höhenkontrollpunkte und Abweichung der Modellergebnisse zu den Kontrollpunkten

5.2 Ergebnisse der Ausgleichung

Im Fall der unter Abschnitt 4.3 beschriebenen Geländemodellierung werden die Höhen der Modellstützpunkte über eine Ausgleichung bestimmt. Bei jeder Ausgleichung nach dem Gauß-Markov-Modell mit vermittelnden Beobachtungen werden die Unbekannten so geschätzt, dass die gewichtete Quadratsumme der Verbesserungen minimiert wird (Gl.1 - 2, NEITZEL, 2010). Aus den Verbesserungen selbst lässt sich der Varianzfaktor σ^2_0 und in Verbindung mit der invertierten Normalgleichungsmatrix dann die Unsicherheit der Unbekannten $\sigma_{k,0}$ ableiten. Dieser ist abhängig von den Freiheitsgraden und damit von der Anzahl der Beobachtungen.

$$v = Adx - l \qquad \text{Gl. 1}$$

$$dx = N^{-1}(A^T Pl) \qquad \text{Gl. 2}$$

$$\sigma_0^2 = \frac{v^T Pv}{r} \qquad \text{Gl. 3}$$

$$\sigma_{k,0}^2 = \sigma_0^2 Q_{xx[0,0]} \text{ mit } Q_{xx} = N^{-1} = (A^T PA)^{-1} \qquad \text{Gl. 4}$$

Das Ausgleichungsmodell geht von unkorrelierten Beobachtungen aus, die frei von systematischen Effekten sind. Da die Beobachtungen in aller Regel mit einem Messsystem erfasst wurden, ist diese Bedingung im strengen Sinne nicht erfüllt. Damit geht einher, dass die $\sigma_{k,0}$ (Gl. 3 - 4) Werte in der Regel zu kleine Beträge annehmen. Bei einer flächenhaften Visualisierung der Werte wird aber dennoch deutlich, in welchen räumlichen Bereichen die Verbesserungen größere Beträge annehmen. In diesen Fällen passt sich die geschätzte Polynomfläche schlechter in die 3D-Punktwolke ein (Abb. 9).

5.3 Modell vs. 3D-Punktwolke

Es ist weiterhin möglich, die 3D-Punktwolke der Modelloberfläche gegenüber zu stellen. Die Modelloberfläche besteht im Fall eines TIN aus vermaschten Dreiecksflächen. Im Fall eines GRIDs könne diese trianguliert werden, wodurch ebenfalls vermaschte Dreiecksflächen entstehen. Die Höhenabweichungen eines Punktes gegenüber einem Dreieck lassen sich mathematisch über die Hessische Normalform bestimmen. Alternativ können die Modellhöhen im Fall eines unvermaschten GRIDs durch bilineare Interpolationen abgeleitet werden. Es ist damit möglich, jedem 3D-Messpunkt eine Höhenabweichung zur Modelloberfläche zuzuordnen und diese Angaben statistisch auszuwerten. Es können rasterbasierte Mittel- oder Medianwerte abgeleitet werden, wobei diese aber um Null schwanken sollten. Die rasterbasierten Standardabweichungen der Höhenabweichungen zeigen eine Streuung der Beobachtungen rund um die Modelloberfläche an. In Abb. 9 sind neben einem Modellbeispiel auch die Unsicherheit der Unbekannten $\sigma_{k,0}$ und die Abgaben zur Höhenabweichung gegenüber der Modelloberfläche dargestellt. Aus der Abbildung geht hervor, in welchen Bereichen sich die Beobachtungen gut über Flächenpolynome approximieren lassen und welche Bereiche eher schlecht modellierbar sind.

In dem Rechenbeispiel der Abb. 9 weisen 86,5% eine Unsicherheit der Unbekannten von <10 mm auf. In Bezug auf die Abweichung der Messwerte zu der Modelloberfläche liegen 86,4% aller Messpunkte in einem Intervall von 10 cm. In einem Intervall von 4 cm liegen 56,1% der Messpunkte.

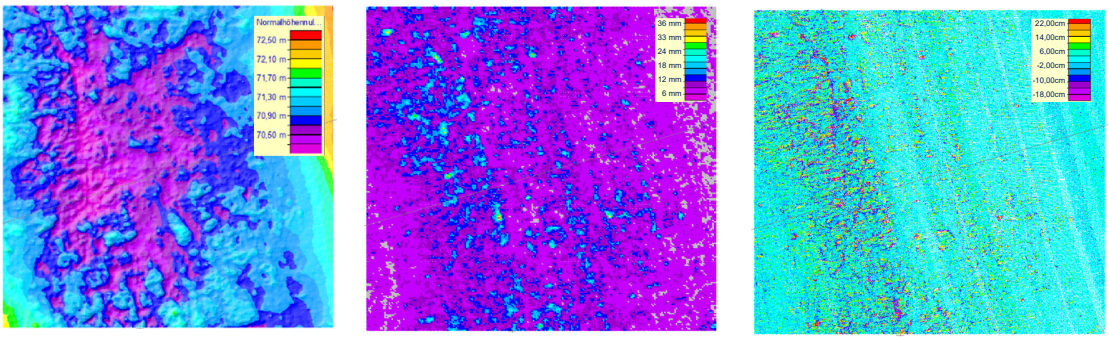


Abb. 9: Modellergebnis (links), Fehler der Unbekannten (Mitte) und Abweichung der Messwerte zur Modelloberfläche (rechts)

Aus den Ergebnissen lassen sich neben Aussagen zur geometrischen Qualität auch Rückschlüsse auf weitere Eigenschaften der Geländeoberfläche zu. Insbesondere Informationen über die Rauigkeit und Beschaffenheit des Bodens sind hierbei von großem Interesse. So lassen sich beispielsweise mit den genannten Parametern strukturierte Bereiche von gleichmäßigen Bereichen unterscheiden.

6 Fazit

Für die Beschreibung der Qualität digitaler Geländemodelle des Wasserlaufs sind vielfältige Parameter nutzbar. Neben den üblichen Metadaten, wie etwa Erfassungszeitraum oder räumliche Auflösung sind insbesondere die dem Modell zugrunde liegenden 3D-Messdaten entscheidend.
Der Detaillierungsgrad steht dabei in direktem Zusammenhang mit der Auflösung des Modells. Ein hoch aufgelöstes Modell mit einem geringen Abstand der Modellstützpunkte untereinander sollte einen wesentlich höheren Detaillierungsgrad aufweisen, als dies bei einem niedrig aufgelösten Modell der Fall ist. Der Informationsgehalt steigt folglich mit der Auflösung und damit einher geht auch eine Steigerung der benötigen Speichermenge bzw. der für eine Bearbeitung nötigen Ressourcen. Der Informationsdichte sind durch die eingesetzten Erfassungssensoren Grenzen gesetzt. Jedes der eingesetzten Sensorsysteme weist unvermeidliche Unsicherheiten auf, kann nur eine definierte Anzahl von Messungen innerhalb einer Zeiteinheit ausführen und darüber hinaus wird die räumliche Auflösung bzw. der Detaillierungsgrad durch den Footprint beschränkt.
Neben einer Beschreibung, welche Sensorik an welcher Stelle zum Einsatz kam, sind auch die damit verbundenen systemspezifischen Parameter, wie etwa Punktdichte, Footprintgröße und Unsicherheiten der 3D-Messpunkte in Lage und Höhe entscheidend. Diese Informationen lassen indirekt Rückschlüsse auf die Qualität des Modells zu.
Weiterhin ist die Modellierungsvorschrift von entscheidender Bedeutung. Diese gibt vor, wie ein Modell erstellt wird und damit auch welche Eigenschaften das Modell hat. Je nach Modellierungsvorschrift werden Informationen ausgedünnt und damit einhergehend die Details mehr oder weniger stark reduziert.

Für eine Beurteilung der geometrischen Qualität ergeben sich verschiedene Möglichkeiten. Einerseits ist es möglich, Modellstützpunkte an der Position von Kontrollpunkten nach der Modellierungsvorschrift zu modellieren und diese den Kontrollpunkten gegenüber zu stellen. Hierbei ist zu beachten, dass auch die Kontrollpunkte Unsicherheiten aufweisen und nicht fehlerfrei sind.
Im Falle einer Ausgleichung ist es zusätzlich möglich, die sich aus der Ausgleichung ergebenden Unsicherheiten der Unbekannten zu visualisieren und zur Qualitätsbeurteilung zu nutzen. Dabei ist zu beachten, dass eine vermittelnde Ausgleichung von unkorrelierten Beobachtungen ausgeht, was streng genommen nicht der Fall ist. Insofern sind die geschätzten Unsicherheiten der Unbekannten i.d.R. zu klein.
Letztendlich kann die ermittelte Modelloberfläche den Beobachtungen gegenübergestellt werden. Die vertikalen Abweichungen zwischen Modelloberfläche und Messwerte können statistisch analysiert und Stützpunktbezogen ausgegeben werden.

7 Literaturverzeichnis

ADV, 2022: Homepage der AdV zu den Produkten des Amtlichen Vermessungswesens in Deutschland. https://www.adv-online.de/AdV-Produkte/Geotopographie/Digitale-Gelaendemodelle/, letzter Zugriff am 13.01.2022.

BW, 2022: Metainformationssystem BW, SAPOS-Daten. https://metadaten.geoportal-bw.de/geonetwork/srv/api/records/944f98d7-99b5-6240-1229-c284b792c489, letzter Zugriff am 13.01.2022.

KUEHNE, E., 2021: New Opportunities for Capturing the Topography of the River Elbe by Airborne Hydromapping in a Low Discharge Period 2018. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **89**, 177-182. https://doi.org/10.1007/s41064-021-00151-2.

LORENZ, F. ARTZ, T., BRÜGGEMANN, T., REICH, J., WEIß, R. & WINTERSCHEID, A., 2021: Simulation-based Evaluation of Hydrographic Data Analysis for Dune Tracking on the River Rhine. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, **89**, 111-120, https://doi.org/10.1007/s41064-021-00145-0.

MIEKE, S., 2020: Berechnung der Messunsicherheit nach GUM Kurzfassung in 20 min. Vortrag – PTB-Institut Berlin, https://www.ptb.de/cms/fileadmin/internet/fachabteilungen/abteilung_8/8.4_mathematische_modellierung/277_PTB_SEMINAR/VORTRAEGE/11_Mieke_-_Berechnung_der_Messunsicherheit_nach_GUM__Kurzfassung_in_20.pdf, letzter Zugriff am 13.01.2022.

NEITZEL, F., 2010: Ausgleichungsrechnung – Modellbildung, Auswertung, Qualitätsbeurteilung, Qualitätsmanagement geodätischer Mess- und Auswerteverfahren, Beiträge zum 93. DVW-Seminar am 10. und 11. Juni 2010 in Hannover. DVW-Schriftenreihe, Band **61**, Seite 95-127.

RIEGL, 2022: Data Sheet – RIEGL VQ780-II. http://www.riegl.com/uploads/tx_pxpriegldownloads/RIEGL_VQ-780II_Datasheet_2021-06-29.pdf, letzter Zugriff am 13.01.2022.

SELLERHOFF, F., 2011: BFG-Kolloquium „Zeitgemäße Erfassung und Bereitstellung von Geobasisdaten für die WSV, 6.4.2011 in Koblenz – Vortrag „Aspekte der DGM-W-Modellierung in tidebeeinflussten Gebieten“. https://www.bafg.de/DE/05_Wissen/02_Veranst/2011/2011_04_06_sellerhoff.pdf?__blob=publicationFile, letzter Zugriff am 13.01.2022.

ZSS, 2022: Zentrale Stelle SAPOS (ZSS), Dienste – SAPOS HEPS, https://zentrale-stelle-sapos.de/dienste-anwendungen-alternativ/, letzter Zugriff am 13.01.2022.

Comparison of different 2D and 3D Sensors and Algorithms for Indoor SLAM on a low-cost Robotic Platform

WEI ZHANG[1], DAVID SKUDDIS[1], PHILIPP J. SCHNEIDER[1] & NORBERT HAALA[1]

Abstract: Mobile robots are becoming a fairly important part of people's lives. Whether they are service robots that assist people in daily life, such as robot vacuums or robots in industry. Simultaneous localization and mapping (SLAM) is one of the most fundamental capabilities to perceive the surroundings and keep track of the robot's position while constructing a map incrementally. SLAM-based surveying equipment is also increasingly used for areas without GNSS availability e.g. mining or indoor cartography. For this purpose, there is a wide range of products from different manufacturers. In practice, depending on the application requirements, different sensors are deployed for this task. Furthermore, with the rapid development of this field in recent years, more new methods have emerged and pushed the boundaries of sensor performance. We noticed a lack of widespread discussion and consensus on which sensor or algorithm is more suitable for a low-cost indoor robotic platform. Therefore, this work aims to compare different low-cost environmental sensors and different advanced algorithms for each of the sensors with the ultimate goal of being able to help make decisions when it comes to choosing sensors and algorithms for a specific robot application. In order to achieve a fair comparison, third-party unbiased reference data is needed. For this purpose, we utilize a wide-angle camera mounted on the ceiling and ArUco marker to achieve a bird's view tracking of the robot's poses serving as reference data. We compare the results of different sensors and algorithms quantitatively against the reference trajectory. In addition to trajectory comparison, another product of the SLAM method is the constructed 2D and 3D maps, which are compared and analyzed qualitatively.

1 Introduction

Mobile robots are widely used in various application domains. They can replace human beings in industry, agriculture and service industries to a certain extent. In addition, they can also be employed in many dangerous environments, such as emergency rescue, space exploration, construction exploration, etc.. In the research of intelligent mobile robots, many technologies are included, such as SLAM, path planning, and navigation. Among them, SLAM technology has always been a research focus in this field.

In an unknown environment, the task of SLAM methods is to obtain an accurate map of the surroundings and localize precisely the robot's position within the environment. Measurements of different 2D and 3D sensors can be used as input for SLAM. As a low-cost robotic platform, we choose the affordable 2D Lidar and 3D depth camera for comparison. With decades of development of SLAM technology, numerous algorithms for each sensor have been developed. For the 2D Lidar, the Matlab Lidar SLAM and ICP graph SLAM methods are selected. As for visual SLAM methods, there are distinctions between keypoint-based, direct, and dense methods.

[1] University of Stuttgart, Institute for Photogrammetry, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, Germany, E-Mail: [Wei.Zhang, David.Skuddis, Philipp.Schneider, Norbert.Haala]@ifp.uni-stuttgart.de

For this reason, one representative method for each paradigm is selected, which are the ORB-SLAM, the Stereo-DSO, and the DROID-SLAM methods.
For the experiment, a low-cost robotic platform is assembled, consisting of a 2D Lidar and 3D depth camera. Additionally, to provide a reference for comparison, an ArUco marker is appended on top of the platform as illustrated in Fig. 1. We employed a wide-view GoPro camera on the room's roof to keep tracking the position and orientation of the robot. The experimental results show that the recent deep learning-based DROID-SLAM method performs best with an ATE error of 2.9 cm. Nevertheless, thanks to the high precision of direct distance measurements, the 2D Lidar-based SLAM provides a more consistent 2D occupancy map. Besides, the Lidar map covers more spaces because of the greater measurement range. By contrast, the resulting 3D map of the visual system contaminates more clutter due to the insufficient accuracy of depth estimate.

2 Methods

This section first explains the low-cost robotic platform, which has a modular structure consisting of three subsystems. Then each applied sensor is introduced. After that, the SLAM algorithms of each sensor modality are elaborated in detail.

2.1 Plattform

Fig. 1 presents an overview of the used low-cost robotic platform and the module diagram consisting of three subsystems, which are the Raspberry-Pi centric robot control and 2D Lidar scanning subsystem, the 3D stereo camera subsystem with the Jetson Xavier NX board as the computing unit, and the reference data provider using a GoPro camera and ArUco marker fixed on top of the robot body. Tab. 1 presents the hardware specification of the two central processing units of 2D Lidar and 3D camera systems.

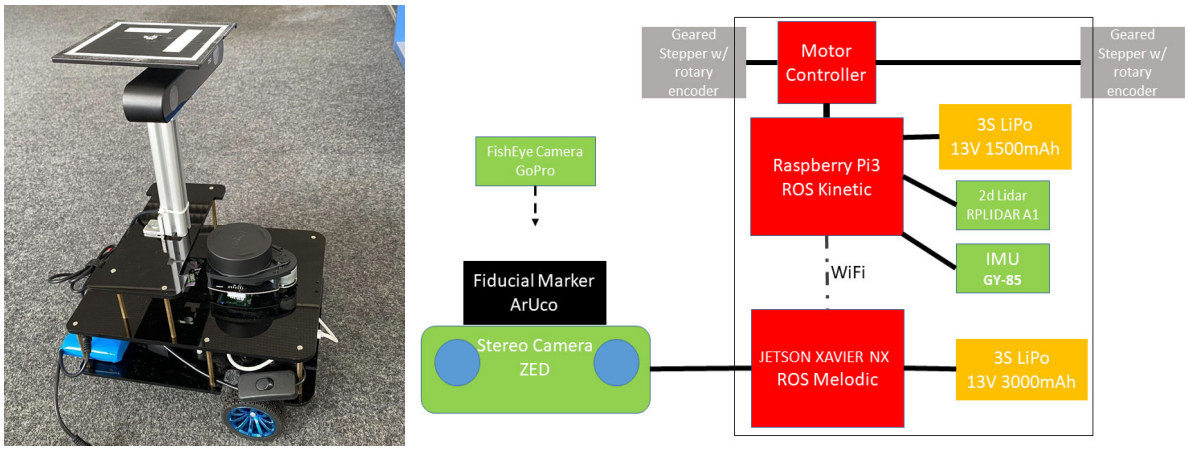


Fig. 1: The low-cost robotic platform and the module diagram

Tab. 1: Hardware specification of the two processing units on the low-cost robotic platform

	Raspberry Pi 3	**Jetson Xavier Nx**
Processor	Broadcom ARMv8 4-cores	Nvidia carmel ARMv8 6-cores
GPU	VideoCore IV	Nvidia vota 384 cores/48 tensor-cores
RAM	1GB	8GB
Sensor	RPLidar, rotary encoder	Zed2 camera
OS	Ubuntu 16.04	Ubuntu 18.04
ROS	ROS Kinetic	ROS Melodic

2.2 Sensors

This section introduces three different sensors applied in our platform, whose measurements are used to estimate the robot's trajectory.

2.2.1 Wheel Encoder

The wheel encoder is equipped on each stepper motor of the two front wheels, and it measures the angular velocities of the left and right wheels. With the available wheel radius and the baseline length between the wheels, the moving and steering speed of the robot platform can be derived, and then the 2D movements can be integrated.

2.2.2 2D Lidar

The 2D Lidar of the model *RpLidar A1* transmits laser beams to the 360-degree surrounding. With an angular resolution of 1 degree, one full spin takes exactly a point cloud of 360 points. The spin frequency is adjusted to 5 Hz so it can sample up to 3600 points per second. Two different Lidar SLAM algorithms are applied to the 2D Lidar data, and the estimated trajectories are compared against the reference.

2.2.3 Camera

The ZED2 stereo camera with a baseline length of 12 cm is applied to capture the color view and 3D information of the environment. The camera is connected to *the Jetson Xavier NX* board, on which runs the Ubuntu 18.04 operating system. Based on the ROS wrapper provided by the manufacturer, the left and right images and stereo depth estimates by ZED SDK are recorded in a rosbag file. The image resolution is set to 672x376 and recorded at 15 Hz. We choose the representative methods to verify the state-of-the-art performance of current visual SLAM methods.

2.3 SLAM Algorithms

Different SLAM algorithms are selected to produce the state-of-the-art performance of each modality. For 2D Lidar, we use Lidar SLAM implementation in the Matlab navigation toolbox and a self-implemented ICP-graph SLAM. According to (TAKETOMI et al. 2017), for visual SLAM, based on whether the image data is directly used for tracking or indirectly via feature extraction, there are two mainstream types of visual SLAM algorithms: the feature-based indirect and the

photo-consistency based direct method. Furthermore, with the recent advances from the application of deep learning, a new visual SLAM paradigm emerges and is represented by the work DROID-SLAM, which we include for comparison.

Tab. 2: Different characteristics of the evaluated Lidar and visual SLAM methods

Method	Sensor	Matching	Loop closure	Map
Matlab-Lidar-SLAM	2D Lidar	Ceres optimization	Submap alignment	2D occupancy
ICP-graph-SLAM	2D Lidar	ICP	ICP close distance	2D point cloud
ORB-SLAM	Stereo	Keypoint-based	Bag of Words	Sparse point cloud
Stereo-DSO	Stereo	Photometric-based	None	Semi-dense point cloud
DROID-SLAM	Mono,RGB-D	Optical-flow-based	Visual view overlap	Dense point cloud

2.3.1 Matlab Lidar SLAM

One of the Lidar-SLAM algorithms investigated in this work is the one provided by Matlab. It is an implementation of the Google Cartographer (HESS et al. 2016). The Google Cartographer is a 2D Lidar SLAM system. The system combines local and global scan matching methods. Newly generated scans are matched locally against a submap using the Ceres solver (AGARWAL & MIERLE, 2010). Many individual submaps are created and stored as map representations. To find loop closures, submaps are compared using a branch-and-bound approach. After a loop is detected, global nonlinear optimization of the residues is performed using the Ceres solver. Since Google Cartographer is available as an open-source project, it is widely used and very popular.

2.3.2 Lidar ICP Graph SLAM

The ICP Graph SLAM is, as the name suggests, a simple self-implemented ICP-based 2D Graph SLAM algorithm. As for map representation, single downsampled scans are stored as keyframes. Newly acquired scans are first downsampled. Then they are matched with the last keyframe using the classical ICP algorithm (BESL & MCKAY 1992). Loop Closure candidates are selected based on the distance to the current keyframe. Old keyframes that are within a certain distance of the current keyframe are compared to the current keyframe using the ICP algorithm. If the residual falls below a threshold, the loop is closed. If a loop closure is found, a global pose graph optimization is performed using the g2o framework (KÜMMERLE et al. 2011). The method was created to provide a second Lidar-based option when comparing the algorithms. Two variants of the method were investigated: one with the raw Lidar data and one with the de-skewed Lidar data. De-skewing means the correction of distortion of the points induced by movement of the scanner. In our case, only rotation rates are taken into account. Velocities can be neglected due to the low vehicle speed. For this purpose the rotation rate of the robot was determined using ICP from two consecutive scans.

2.3.3 Stereo-ORB-SLAM

ORB-SLAM (CAMPOS et al. 2021) is a feature-based method, which extracts ORB features for tracking and creates a sparse point cloud as the map. It was first introduced to run in a monocular

mode. The following extensions include a stereo mode, which amongst others provides an absolute scale for the resulting map. Due to the availability of a depth map from the stereo images, the system can initialize faster without the need of moving the camera for the triangulation of initial map points. Thanks to the loop closure detection module and global pose optimization (bundle adjustment), it is a complete SLAM system widely used in robotic applications.

2.3.4 Stereo-DSO

Stereo Direct Sparse Odometry (DSO) is a direct method (WANG et al. 2017), which minimizes the photometric consistency error for the pixels with sufficient intensity gradient. Within a sliding window, the selected pixels are tracked across multiple frames. While camera poses are tracked, the 3D locations of the pixels are iteratively estimated. Although there is no loop closure and global optimization, it still produces a good result with a much denser map than the ORB-SLAM result. Like ORB-SLAM, a stereo setting can provide initial depth estimation at the initial phase and thus facilitate the system initialization.

2.3.5 DROID-SLAM

DROID-SLAM (TEED & DENG 2021) is a dense optical-flow-based method, which minimizes the reprojection error with the optical flow prediction by a pre-trained neural network as the reprojection targets. The neural network extracts multi-level feature maps from color images. To predict an update of optical flow, the current flow is firstly computed based on the current pose and depth estimates, i.e. the dense projection of the pixel array between two frames with overlap view. With that, the local features at the source and reference frame can be looked up from the extracted feature maps. The local features from both frames are processed via CNN and GRU subnetwork to predict an updated optical flow. For efficiency, the keyframes are selected based on the mean motion of the optical flow, and the depth maps are down-sampled to 1/8 resolution as estimation variables. The poses and low-resolution depth maps of keyframes are jointly optimized via bundle adjustment with the Gauss-Newton algorithm. In the following experiment, the DROID-SLAM will be tested with monocular and RGB-D settings. For the former setting, only the left camera image is fed into the system. For RGB-D, the depth estimate produced by ZED SDK is used as additional depth prior to the bundle adjustment.

2.4 Ground Truth

In order to compare the different trajectories to each other a ground truth is needed. We opted to use a wide-angle camera on the ceiling to track a squared fiducial marker (ROMERO-RAMIREZ et al. 2018), mounted on top of the robot. In order to undistort the fisheye camera's images, a rectification to a rectilinear image was applied (SCARAMUZZA et al. 2006). Since all marker positions are on the same plane, a homography (HARTLEY & ZISSERMAN 2004) was applied to remove the projective distortion from the perspective linear image of this plane into a rectangular image (Fig. 2). These rectangular images are used to detect the fiducial marker's pixel coordinates and heading. By placing the rover on top of known points, a transformation from pixel coordinates to object coordinates can be computed.

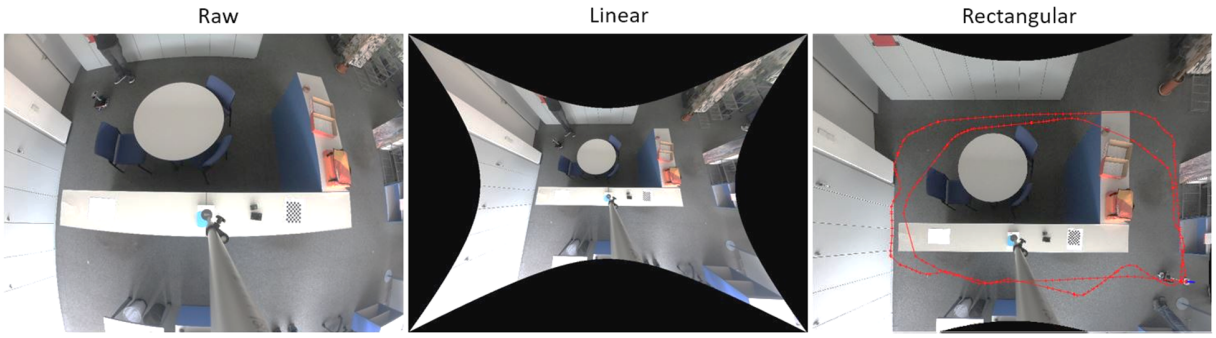


Fig.2: Workflow of reference data generation. The raw distorted images are transformed into rectilinear ones, using a fisheye camera calibration model. A homography was used to transform the marker's plane into a rectangular plane that allows measuring the rover's position in pixel coordinates, which can be transformed into object coordinates by placing the rover onto fixed points with known coordinates

2.5 Evaluation Metric

To evaluate the estimated trajectory $\mathbf{Q}_{1:n}$ by different methods, we compute the absolute trajectory error (ATE) (STURM et al. 2012) against the ground truth trajectory $\mathbf{P}_{1:n}$. Due to the difficulty of synchronizing the time of the ground truth system and sensor-specific system, the corresponding pose $\mathbf{Q}_i$ is found via the nearest neighbor search for each ground truth pose $\mathbf{P}_i$. Let *trans()* refers to the translational part of the relative pose, the error at a timestamp i can be computed as

$$E_i = trans(\mathbf{Q}_i^{-1} \cdot \mathbf{P}_i) \tag{1}$$

From the pose errors of all timestamps, the ATE can be derived by taking root mean square over the entire trajectory as follows:

$$ATE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}||E_i||^2} \tag{2}$$

The ATE and the maximum of E_i (Max) are used as the metrics to indicate the average and worst performance.

3 Results

The estimated trajectories are presented in Fig. 3. For clarity, each modality is plotted in a separate subplot. The trajectory estimated by the wheel odometer has the largest deviation from the ground truth, as it suffers from the accumulated error during the movement integration over time. It should also be noted that all methods except wheel odometry and Stereo-DSO can detect loop closures by design and have detected them successfully. Since the trajectory consists of two loops, this can improve the result. For 2D Lidar-based systems, the evaluated trajectories are shown in Fig. 3b) and the evaluation results are shown in Fig. 4 and Tab. 3. Comparing the two Lidar-based SLAM approaches, it is seen that the Matlab SLAM performed marginally better than the ICP Graph SLAM.

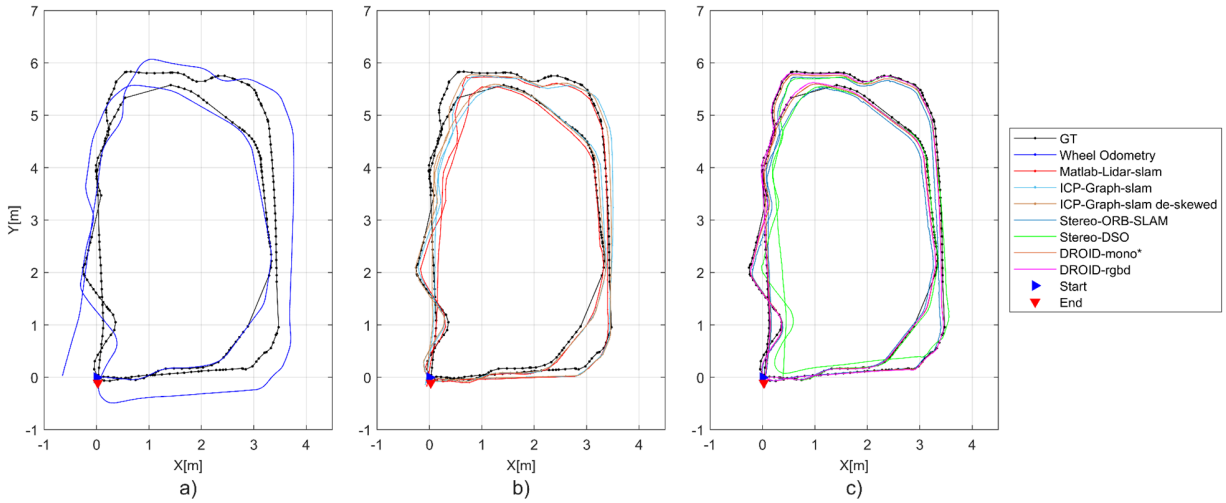


Fig. 3: Experimental results. a) shows the estimated trajectory given by the wheel odometer, b) shows the results of Lidar algorithms. c) is the estimated trajectories of visual algorithms in comparison to ground truth (GT). Note mono* method has scale ambiguity and its scale is adjusted to the GT scale.

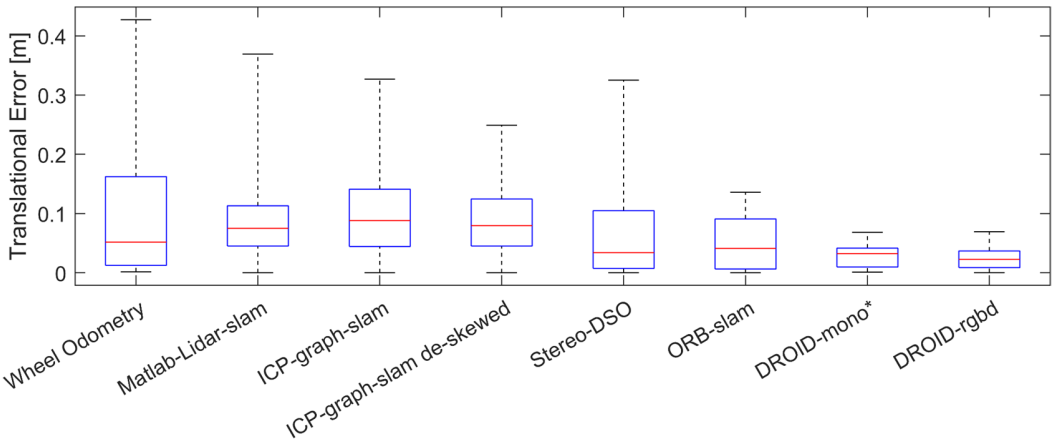


Fig. 4: Quantitative evaluation results with respect to the translational errors for different methods compared to the GT poses. The red lines mark the median in each case. The boxes mark the range between the 25th percentile and the 75th percentile.

Tab. 3: Evaluation results of different methods. Bold font indicates the best result for each metric

	Wheel	Matlab-Lidar-SLAM	ICP-graph-SLAM	ICP-graph-SLAM de-skewed	Stereo-DSO	ORB-SLAM	DROID-mono	DROID-rgbd
Data	Wheel Encoder	2D-Lidar	2D-Lidar	2D-Lidar	Stereo image	Stereo image	Mono image	Left image +depth
Loop Closure	-	+	+	+	-	+	+	+
#Pose	1663	167	125	125	2453	2554	853	853
Max Error [m]	0.567	0.369	0.322	0.249	0.325	0.136	***0.068***	0.069
ATE [m]	0.201	0.121	0.118	0.103	0.104	0.065	0.034	***0.029***

The results of the camera-based methods are shown in Fig. 3c). It can be seen that each of the camera-based methods performs better than the Lidar-based methods. For Lidar-based methods,

de-skewing also significantly improved the result of the ICP-graph-slam. Furthermore, in Fig. 3 it looks like the trajectories determined by the Lidar-based methods deviate similarly in direction and magnitude from the ground truth. This may be due to a systematic measurement error of the Lidar. Within the camera-based methods, the DROID-rgbd achieves the best results. The second best results can be achieved with the DROID-mono SLAM. It should be noted that the produced result by DROID-mono is not in metric scale, which is an inherent problem of monocular SLAM. Thus the trajectory is aligned in addition by a scale factor of 1.2 to GT. The largest deviations in the camera-based methods are observed in the stereo-DSO method. The result of the stereo DSO is however remarkable in so far as it has no loop closure functionality. Thus, the errors were continuously integrated in both loops. Nevertheless, it achieves lower ATE than the Lidar-based methods as present in Tab.3. The accuracy of the ORB-SLAM is in the midfield. Nevertheless, the magnitudes of the ATE are interesting at this point. Both the ATE of the ORB-SLAM and the maximum deviation from the trajectory are twice as large as for the determined favorite DROID-rgbd.

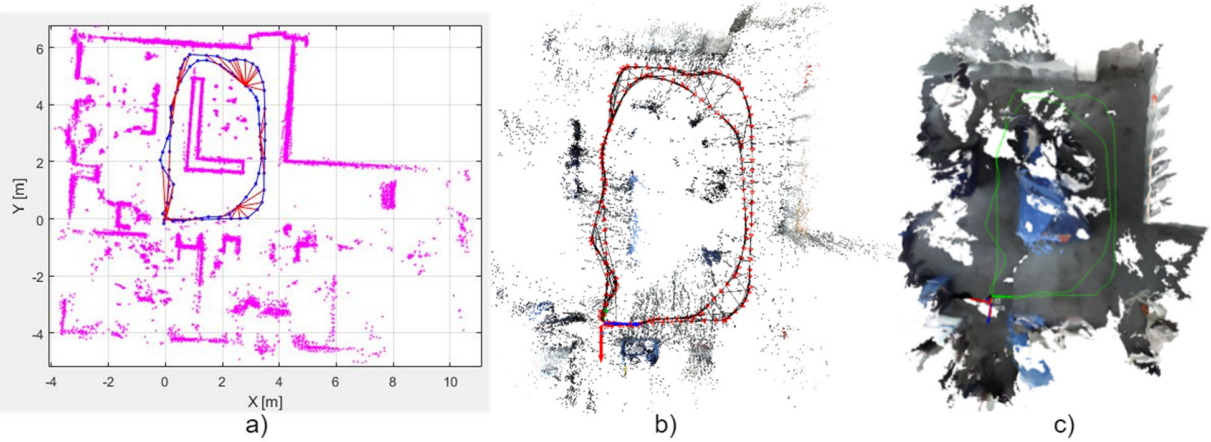


Fig. 5: Different map representations, which are 2D Lidar map, 3D point cloud map, and 3D dense volumetric map respectively

Fig. 5 shows different map representations of the applied methods. The 2D Lidar-based method produces a 2D occupancy map. As shown in Figure 5a, this 2D Lidar map is globally consistent with visually sharp edges of the room walls and furniture. By contrast, the reconstructed maps by visual SLAM methods are in 3D, whether presented as a point cloud or dense volumetric map. Compared to the Lidar map, the map reconstructed by the applied stereo camera has more clutter due to the inaccurate depth estimation. Furthermore, the space covered by the camera map is smaller than that of the Lidar map, as the camera depth range is limited by the baseline length of the stereo camera.

4 Conclusion

In our test scenario, it was shown that camera-based methods are superior to Lidar methods in the context of the used accuracy metrics. The reasons for this can be the more precise capture of the environment due to the higher resolution of the camera pixels compared to the Lidar resolution.

Nevertheless, Lidar-based methods have other advantages that do not apply in the selected scenario. These are for example the robustness in very homogeneous and featureless environments, e.g. large white walls.
It is also important to note that based on the chosen comparison scenario, no conclusions can be made about large-scale routes and the ability to detect loop closures there. Furthermore, the environment within the scenario is limited to an indoor office environment.
For future work, we believe that a more robust system can be achieved by combining the information from the 2D and 3D sensors, and also the valuable information from the inertial sensor and wheel odometer to compensate for the respective shortcomings. Moreover, towards a more intelligent robotic system, we will investigate how to derive higher-order semantic information from the constructed 2D and 3D maps.

5 Literaturverzeichnis

AGARWAL, S., MIERLE, K. & OTHERS, 2010: Ceres Solver — A Large Scale Non-linear Optimization Library. http://ceres-solver.org.

BESL, P.J. & MCKAY, N.D., 1992: A method for registration of 3-d shapes. IEEE Trans Pattern Anal Mach Intell **14**(2). 239-256, https://doi.org/10.1117/12.57955.

CAMPOS, C., ELVIRA, R., RODRIGUEZ, J.J.G., MONTIEL, J.M. & TARDOS, J., 2021: ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. IEEE Transactions on Robotics, **37**(6), 1874-1890, https://doi.org/10.1109/TRO.2021.3075644.

HARTLEY, R., ZISSERMAN, A., 2004: Projective Geometry and Transformations of 2D. Multiple View Geometry in Computer Vision, Cambridge University Press, 34-36, https://doi.org/10.1017/CBO9780511811685.005.

HESS, W., KOHLER, D., RAPP, H. & ANDOR, D., 2016: Real-Time Loop Closure in 2D LIDAR SLAM. IEEE International Conference on Robotics and Automation (ICRA), 1271-1278, https://doi.org/10.1109/ICRA.2016.7487258.

KÜMMERLE, R., GRISETTI, G., STRASDAT, H., KONOLIGE, K. & BURGARD, W., 2011: G2o: A general framework for graph optimization. IEEE International Conference on Robotics and Automation, 3607-3613, https://doi.org/10.1109/ICRA.2011.5979949.

STURM, J., ENGELHARD, N., Burgard, W. & Cremers, D., 2012: A benchmark for the evaluation of RGB-D SLAM systems. IEEE/RSJ International Conference on Intelligent Robots and Systems, 573-580, https://doi.org/10.1109/IROS.2012.6385773.

ROMERO-RAMIREZ, F.J., MUÑOZ-SALINAS, R. & MEDINA-CARNICER, R., 2018: Speeded up detection of squared fiducial markers. Image and Vision Computing, **76**, 38-47, https://doi.org/10.1016/j.imavis.2018.05.004.

SCARAMUZZA, D., MARTINELLI A. & SIEGWART R., 2006: A Toolbox for Easy Calibrating Omnidirectional Cameras. IEEE International Conference on Intelligent Robots and Systems, S.5695-5701, https://doi.org/10.1109/IROS.2006.282372.

TAKETOMI, T., UCHIYAMA, H. & IKEDA, S., 2017: Visual SLAM algorithms: a survey from 2010 to 2016. IPSJ Transaction on Computer Vision and Applications, **9**(16), https://doi.org/10.1186/s41074-017-0027-2.

TEED, Z. & DENG, J., 2021: DROID-SLAM: Deep Visual SLAM for Monocular, Stereo, and RGB-D Cameras. ArXiv preprint arXiv:2108.10869.

WANG, R., SCHWÖRER, M. & CREMERS, D., 2017: Stereo DSO: Large-Scale Direct Sparse Visual Odometry with Stereo Cameras. IEEE Inter. Conference on Computer Vision, 3903-3911, https://openaccess.thecvf.com/content_ICCV_2017/papers/Wang_Stereo_DSO_Large-Scale_ICCV_2017_paper.pdf

Point Surfel Transformer Network for Semantic Segmentation of Large-Scale ALS Point Clouds

XINLONG ZHANG[1], RUIHANG XUE[1], MICHAEL KÖLLE[1] & UWE SÖRGEL[1]

Abstract: Automated semantic segmentation of point clouds plays an important role in 3D scene perception. The definition of relevant features is usually key for segmentation and classification, with automated workflows presenting the main challenges. Point Transformer networks based on self-attention operator, which is invariant to permutation of the input elements, can describe the essential attributes of disordered point clouds well. However, the application to large-scale Airborne Laser Scanning (ALS) point cloud scenes is not trivial. Naive Point Transformer lacks the ability to describe local features, therefore most of the established methods focus on small simple scenes. In this work, the Point-Surfel Transformer (PS-Transformer) network based on not only point features but also local surfel features to strengthen the local perception, is proposed. Our approach is evaluated on the Hessigheim High-Resolution 3D Point Cloud (H3D) Benchmark and achieves state-of-the-art 89.19% overall accuracy. Furthermore, our proposed PS-Transformer approach outperforms outperforms naive point transformer by a large margin of 4.64 percentage points.

1 Introduction

Driven by the fast development of depth sensors and 3D scanners, the automatic processing of point clouds has played an indispensable role in the fields of mapping geographic information, autonomous driving and robotics (CHEN et al. 2020). As a key step of understanding 3D scenes, semantic segmentation of point clouds has attracted extensive attention from researchers.

Current methods can be generally grouped into two categories: the models based on handcrafted features and the models based on deep learning. In the first category, manually designed features require expensive calculation and suffer from poor generalization (LECUN et al. 2015). On the contrary, the data-driven features extraction based on deep learning does not need hand-designed feature extractors, and can automatically learn better representations of objects (LECUN et al. 2015). Compared with the regular grid structure of image data, the disordered distribution of point clouds makes 3D semantic segmentation a challenging task.

In recent years, a large number of models based on deep learning have been proposed for 3D semantic segmentation. Inspired by 2D convolutional neural networks (CNNs), VoxNet (MATURANA et al. 2015) voxelizes point clouds to make the data structure suitable for 3D CNNs. But the sparsity of point clouds leads to the low efficiency of voxel grid arrangement and the high computational burden. To alleviate this problem, the sparse convolutional network (LIU et al. 2015) and its 3D application (VERDOJA et al. 2017) operate only on voxels that are not empty. However, such methods only depend on the voxel boundary and ignore the local geometric structure. PointNet++ (QI

[1] University of Stuttgart, Institute for Photogrammetry,
Geschwister-Scholl-Str. 24D, D-70174 Stuttgart,
E-Mail: [Xinlong.Zhang, Ruihang.Xue, Michael.Kölle, Uwe.Sörgel]@ifp.uni-stuttgart.de

et al. 2017) effectively solves the problem of extracting local features by combining sampling-grouping layer and PointNet (QI et al. 2017) layer. Due to the network's pooling operator, features in each individual dimension have the same weight. The self-attention based Point Transformer (ZHAO et al. 2021) weights each element adaptively. Due to this set operation the network is invariant to permutation of the input elements (JADERBERG et al. 2015), which is consistent with the disordered distribution of point clouds. Nevertheless, the relationship between points in sets is not taken into consideration.

In complex large-scale point clouds, naive Point Transformer usually fails to extract sufficiently effective semantic features and performs poorly, because of weak local features. Surface element (surfel) features on the other hand (PFISTER et al. 2000) comprise normal, curvature, scale and others, which can provide the attributes of the local approximate surface of each point (WEINMANN et al. 2013). This kind of geometric adjacency information can strengthen the network's perception of local regions.

In this work, the Point-Surfel Transformer (PS-Transformer) network, based on not only point features but also surfel features, is proposed. First, the local surfel features are generated from point coordinates. Then, the surfel features are mapped to the original point feature space to obtain the point-surfel features. Subsequently, the enhanced features are fed to the PS-Transformer network to generate point-wise feature vectors in an encoder-decoder architecture. Finally, a multi-layer perceptron (MLP) maps each feature vector to the final logits, which are used to predict the class probability of each point.

2 Point Surfel Transformer Network

2.1 Surfel Features Extraction

Surfel features of each individual point p are extracted from the raw XYZ coordinates of point clouds, which can be expressed by one 6-tuple

$$\mathrm{F}_p = (n_{px}, n_{py}, n_{pz}, d_p, f_p, r_p) \tag{1}$$

where n_{px}, n_{py}, n_{pz} are the parameters of the normal vector, d_p is the distance from the origin to the fitted plane of point p, f_p is change of curvature, and r_p is the residual from point p to its fitted surface. Normals distinguish flat and inclined surfaces, d_p is the association with global information, f_p represents the local surface variation and r_p describes the local roughness.

These local features of each point $p(x_p, y_p, z_p)$ are calculated as follows. Firstly, the covariance matrix $M_{3\times3}$ is constructed via the Eq. (2).

$$M_{3\times3} = \frac{1}{k}(p_i - \overline{p}) \cdot (p_i - \overline{p})^T \tag{2}$$

Where p_i is the K Nearest Neighbour(KNN) of point p, and $\overline{p}$ represents the mean vector of the point p and its KNNs. Then eigenvalues and eigenvectors of $M_{3\times3}$ are calculated by Singular Value Decomposition (SVD) as Eq. (3) and Eq. (4).

$$\lambda V = M_{3\times3} V \tag{3}$$

$$\begin{cases} V = (v_1^{(M)}, v_2^{(M)}, v_3^{(M)}) \\ \lambda = diag(\lambda_1^{(M)}, \lambda_2^{(M)}, \lambda_3^{(M)}), \lambda_1^{(M)} > \lambda_2^{(M)} > \lambda_3^{(M)} \end{cases} \tag{4}$$

where λ is the vector of eigenvalues of the covariance matrix and V is the set of eigenvectors of the covariance matrix. The local features of point p are calculated as Eq. (5).

$$\begin{cases} n_p(n_{px}, n_{py}, n_{pz}) = v_3^{(M)} \\ d_p = -(n_{px}x_p + n_{py}y_p + n_{pz}z_p) \\ f_p = \dfrac{\lambda_3^{(M)}}{\lambda_1^{(M)} + \lambda_2^{(M)} + \lambda_3^{(M)}} \\ r_p = \dfrac{|n_{px}x_p + n_{py}y_p + n_{pz}z_p + d_p|}{\sqrt{n_{px}^2 + n_{py}^2 + n_{pz}^2}} \end{cases} \tag{5}$$

The extracted surfel features composed of low-level descriptors contain rich local information, which will be used for the following processes.

2.2 Point Transformer Networks

The overall structure of the proposed PS-Transformer is shown in Fig. 1, where the generated surfel features are firstly concatenated with the point features (coordinates and RGB values), and then a U-Net-like (RONNEBERGER et al. 2015) point Transformer is implemented for feature fusing and learning. The U-Net-like Transformer is composed of point transformer block, transition down and transition up modules. Connecting them alternately could obtain an 8-layer network, where the first 4 layers are encoders, and the last 4 layers are decoders. In the first layer, a multilayer perceptron (MLP) is applied to fuse the point-surfel features. Besides, each feature encoder layer has a point transformer block connected by a transition down module. Moreover, each feature decoder layer has a point transformer block connected by a transition up module. Finally, the encoder-decoder structure together forms a U-Net-like network to fulfill the semantic segmentation task.

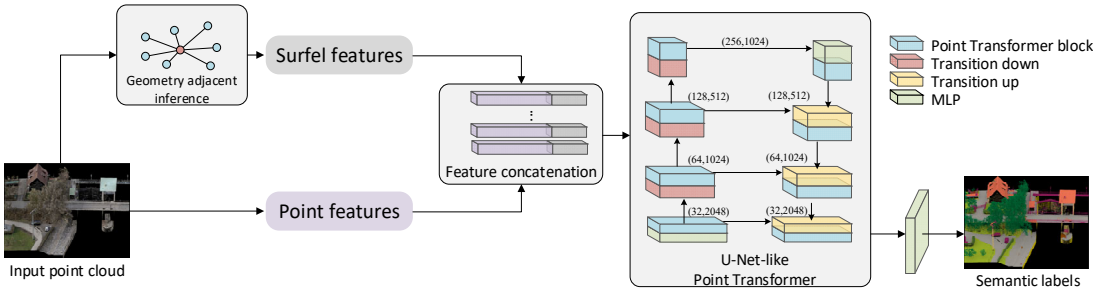


Fig. 1: Overall structure of the proposed PS-Transformer

As the core of the proposed network, the point transformer block is formed by cascading two linear mappings and a self-attention calculation. The linear mapping converts the

input-output dimension, and the self-attention estimates the internal relationship among the input points. For the input points set X, the subset $X_i \in X$ is the neighborhood of the point x_i, which is obtained by the KNN algorithm, where $i = 1, 2, ..., N$ and the number of neighborhood points is N. Hence the self-attention calculation of the point x_i is defined as:

$$y_i = \sum_{x_j \in X_i} \alpha\left(\beta\left(\varphi_k\left(x_i\right) - \varphi_q\left(x_j\right) + \text{pos}\right)\right) \text{e}\ \varphi_v\left(x_j + \text{pos}\right) \tag{6}$$

where y_i is the output feature vector, α is the softmax activation function, and β is the attention mapping function, which is implemented by a MLP, i.e. 2 linear layers and a ReLU (GLOROT et al. 2011) activation function. φ_k, φ_q and φ_v are all linear mappings for adapting to different feature dimensions, and e denotes an elementwise multiplication. pos is the positional coding, which is a linear mapping from the relative coordinate of the points:

$$\text{pos} = \varphi_p\left(cor_i - cor_j\right) \tag{7}$$

where cor_i and cor_j are respectively the 3D coordinates of the point i, j, φ_p is a MLP.

Transition down module realizes the down-sampling of the local point cloud, which is implemented by the farthest point sampling and KNNs searching. Then the feature vectors of the sampled subset are obtained by local max-pooling. Transition up module realizes the up-sampling of the local point cloud, where the number of points is increased by trilinear interpolation. Then the features of the corresponding encoder layer are added to the up-sampled new point set, which is also the U-Net-like connection. The feature decoder has a symmetrical configuration with the encoder as shown in Fig. 1, where (32,2048) represents that in this layer, the feature dimension is 32 and the output point number is 2048.

For the semantic segmentation task, the PS-Transformer network should have an output label for each point. The 4-by-4 symmetrically designed encoder and decoder could exactly guarantee the correspondence between the input point-surfel features and the output labels. At the last layer, a MLP maps the point feature to the label space y_k, and all the learnable parameters of PS-Transformer could be updated by optimizing the cross-entropy loss function:

$$\text{Loss}(\mathbf{w}) = -\sum_{i=1}^{N}\sum_{k=1}^{K} t_{ki} \ln y_k(X_i, \mathbf{w}) \tag{8}$$

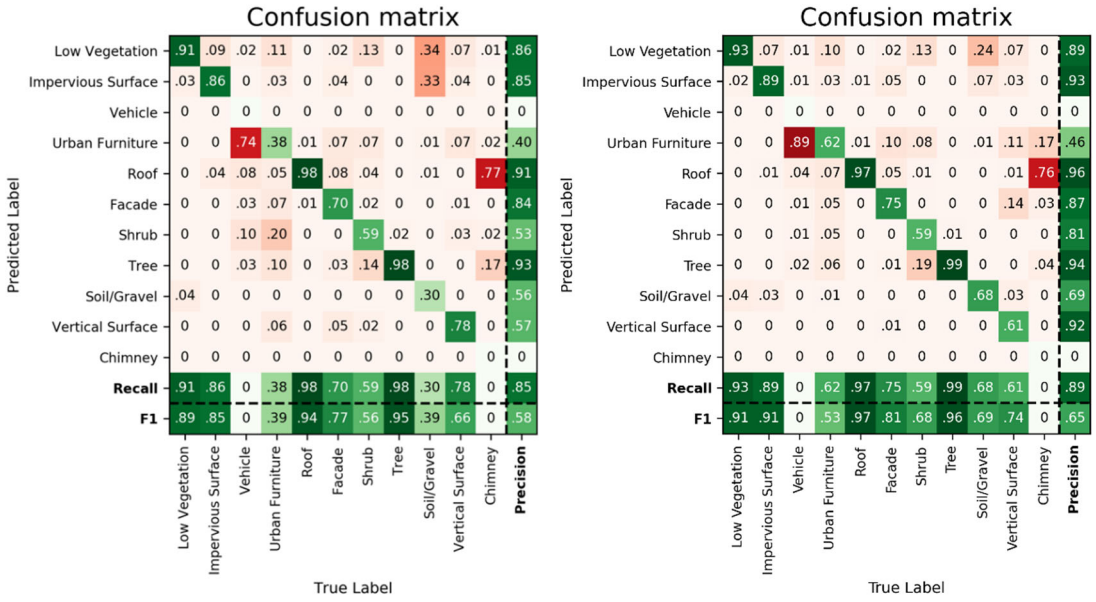
where K denotes to the number of categories, t_{ki} denotes to the one-hot truth corresponding to the ith point, and $\mathbf{w}$ denotes to the set of learnable parameters.

3 Experiments

In this section, we conduct comparative experiments to evaluate our PS-Transformer networks. The experiments are based on the Hessigheim High-Resolution 3D Point

Cloud (H3D) Benchmark (KÖLLE et al. 2021). The dataset was collected by a Riegl VUX-1LR Scanner and two oblique-looking Sony Alpha 6000 cameras integrated on a RIEGL Ricopter platform. The mean point density is 800 points/m² enriched by RGB colors and the ground sampling distance (GSD) of images is 2-3 cm. In addition, the points have been manually labeled with the following 11 classes: Low vegetation, Impervious surface, Vehicle, Urban furniture, Roof, Facade, Shrub, Tree, Soil/Gravel, Vertical surface, Chimney.

In order to reduce the computational burden, the training data and the test data are cropped into 49 splits and 22 splits, respectively. Our implementation of the PS-Transformer is realized in PyTorch. The Adam optimizer is employed in the network, and we train the network for 20 epochs with an initial learning rate of 0.0005. The segmentation results are evaluated by overall accuracy (OA), recall, and F1-score.



(a) naive Point Transfomer (b) our PS-Transfomer approch

Fig. 2: Detailed H3D test set segmentation results from the comparative experiment: (a) naive Point Transfomer (b) our PS-Transformer approach

We compare against the naive Point Transfomer to showcase the benefits of surfel features, and the results are shown in Fig. 2. Our approach achieves state-of-the-art 89.19% overall accuracy, which outperforms the naive Point Transformer (84.55%) by a large margin. Furthermore, the recall of fine-grained classes is greatly improved, the recall of urban furniture increases from 38.06% to 62.02%, and the recall of soil/gravel increases from 29.85% to 68.18%.

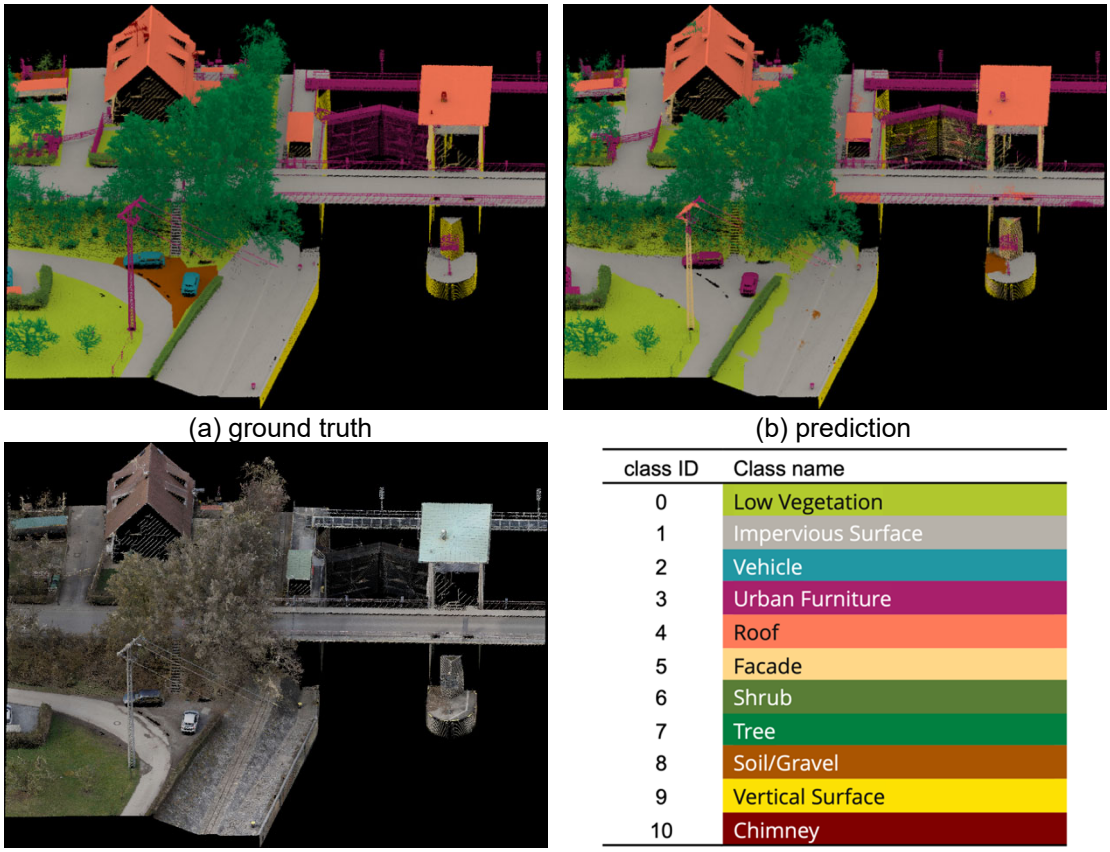


Fig. 3: Comparison of our prediction and ground truth on H3D: (a) ground truth (b) prediction (c) RGB texture (d) class catalog

Fig. 3 displays an example of our prediction, ground truth and RGB texture. Although chimneys are interpreted as roofs more frequently, they are less often classified as facades or vertical surfaces. The difficulties mainly exist between vehicles and urban furniture, and soil/gravel are often inferred as impervious surface. The ambiguities are caused by their limited inter-class distances and scarce appearances.

4 Conclusion

We have presented PS-Transformer, a surfel features enhanced network for semantic segmentation of large-scale ALS point clouds. The proposed PS-Transformer utilizes geometric adjacency information to strengthen the network's local perception, and achieves state-of-the-art 89.19% overall accuracy on the H3D Benchmark. Future work will be focused on semantic segmentation for fine-grained objects (such as vehicles), while being aimed at models in the case of imbalanced samples.

5 Acknowledgements

The research work in this paper has been funded by China Scholarship Council (No. 202008080145).

6 References

CHEN, S., LIU, B., FENG, C., VALLESPI-GONZALEZ, C. & WELLINGTON, C., 2020: 3d point cloud processing and learning for autonomous driving: Impacting map creation, localization, and perception. IEEE Signal Processing Magazine, **38**(1), 68-86, https://doi.org/10.1109/MSP.2020.2984780.

GLOROT, X., BORDES, A. & BENGIO, Y., 2011: Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 315-323.

JADERBERG, M., SIMONYAN, K. & ZISSERMAN, A., 2015: Spatial transformer networks. Advances in neural information processing systems, 28, 2017-2025.

LECUN, Y., BENGIO, Y. & HINTON, G., 2015: Deep learning. Nature, **521**(7553), 436-444, https://doi.org/10.1038/nature14539.

LIU, B., WANG, M., FOROOSH, H., TAPPEN, M. & PENSKY, M., 2015: Sparse convolutional neural networks. IEEE conference on computer vision and pattern recognition, 806-814.

MATURANA, D. & SCHERER, S., 2015: Voxnet: A 3d convolutional neural network for real-time object recognition. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 922-928, https://doi.org/10.1109/IROS.2015.7353481.

KÖLLE, M., LAUPHEIMER, D., SCHMOHL, S., HAALA, N., ROTTENSTEINER, F., WEGNER, J. D. & LEDOUX, H., 2021: The Hessigheim 3D (H3D) Benchmark on Semantic Segmentation of High-Resolution 3D Point Clouds and Textured Meshes from UAV LiDAR and Multi-View-Stereo. ISPRS Open Journal of Photogrammetry and Remote Sensing, **1**, 11.

PFISTER, H., ZWICKER, M., VAN BAAR, J. & GROSS, M., 2000: Surfels: Surface elements as rendering primitives. 27th Annual Conference on Computer Graphics and Interactive Techniques, 335-342.

QI, C. R., SU, H., MO, K. & GUIBAS, L. J., 2017: Pointnet: Deep learning on point sets for 3d classification and segmentation. IEEE conference on computer vision and pattern recognition, 652-660.

QI, C. R., YI, L., SU, H. & GUIBAS, L. J., 2017: PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. Advances in Neural Information Processing Systems, 30.

RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 234-241.

VERDOJA, F., THOMAS, D. & SUGIMOTO, A., 2017: Fast 3D point cloud segmentation using supervoxels with geometry and color for 3D scene understanding. IEEE International Conference on Multimedia and Expo (ICME), 1285-1290.

WEINMANN, M., JUTZI, B. & MALLET, C., 2013: Feature relevance assessment for the semantic interpretation of 3D point cloud data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **5**(W2), 1.

ZHAO, H., JIANG, L., JIA, J., TORR, P. H. & KOLTUN, V., 2021: Point transformer. IEEE/CVF International Conference on Computer Vision, 16259-16268.

Simulation und Auswertung eines photogrammetrischen Bildverbandes aus Krankamerabildern

SYLVIA ACKERMANN[1] & LENA JOACHIM[1]

Zusammenfassung: Zur Beschleunigung von Bauprozessen sollen autonome Transportprozesse eines Turmdrehkrans realisiert werden. Als Grundlage der Pfadplanung in der Baustellen-Umgebung wird hier ein aktuelles digitales Oberflächenmodell verwendet. Dieses wird mit Hilfe von am Ausleger des Krans montierten Kameras photogrammetrisch generiert. In dieser Arbeit wird die Zuverlässigkeit der photogrammetrischen Rekonstruktion hinsichtlich deren Eignung für die automatisierte Pfadplanung untersucht. Zuerst wird eine Simulation zur Planung von Passpunktverteilung und Aufnahmekonfiguration des Bildverbandes implementiert. Die anschließende Auswertung des realen Bildverbandes, ergibt im Baustellenbereich eine mittlere Objektpunktgenauigkeit von 2,5 cm. Der Vergleich mit Laserscanning-Referenzdaten zeigt, dass Objekte mit einer Mindestbreite von 3 cm in der Rekonstruktion enthalten sind, sofern diese sich ausreichend vom Hintergrund abheben.

1 Einleitung

Durch Bevölkerungswachstum, Ressourcenknappheit und Klimawandel ergeben sich in Zukunft zahlreiche zu bewältigende Herausforderungen für das Baugewerbe. Problematisch ist allerdings, dass die Baubranche gegenüber anderen Wirtschaftsbereichen große Rückstände in der Produktivität aufzuweisen hat (STAT. BUNDESAMT 2018). Um Bauprozesse in Zukunft beschleunigen zu können, sollen autonome Transportprozesse eines Turmdrehkrans realisiert werden. Dafür wird die automatisierte Planung eines Pfades durch die Baustellen-Umgebung benötigt. Als Grundlage dieser Pfadplanung soll ein aktuelles digitales Oberflächenmodell (DOM) der Baustelle verwendet werden. Dieses wird photogrammetrisch generiert, wofür fünf Kameras am Ausleger des Turmdrehkrans montiert werden. Eine Analyse der Zuverlässigkeit der photogrammetrischen Rekonstruktion soll deren Eignung für die automatisierte Pfadplanung zeigen. Dabei werden die Faktoren Genauigkeit und Vollständigkeit betrachtet.

2 Datengrundlage

2.1 Photogrammetrischer Bildverband

Für die Aufnahme des photogrammetrischen Bildverbandes werden fünf Kameras am Ausleger des Turmdrehkrans in einer Höhe von etwa 45 m montiert. Bei den Kameras handelt es sich jeweils um das acA4112-8gc-Modell von Basler mit einer Brennweite von 16 mm und einem 12 Megapixel CMOS-Sensor (4096 × 3000 Pixel). Die Kameras werden in Nadirsicht montiert, sodass sich eine Bodenpixelgröße von 9,7 mm ergibt. Wie Abb. 1 zu entnehmen ist, werden die Kameras in einem Abstand von 10 m angebracht, wobei sich Kamera A nah am Rotationszentrum

[1] Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: st160591@stud.uni-stuttgart.de, lena.joachim@ifp.uni-stuttgart.de

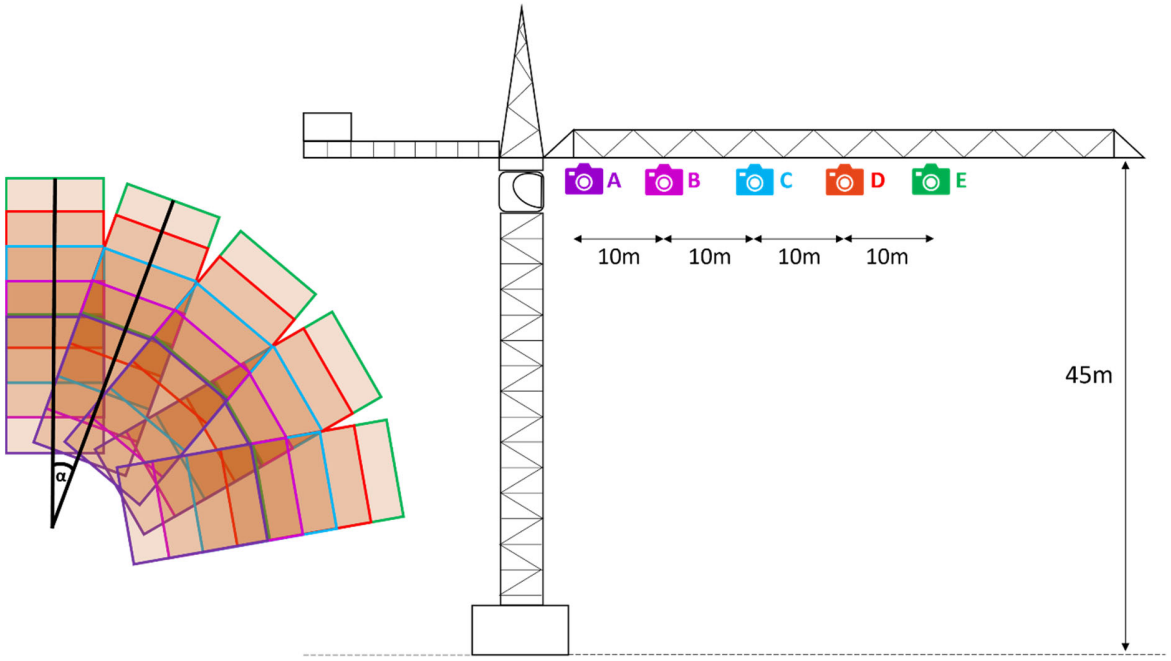


Abb. 1: Rechts: Montierung der Kameras A-E am Ausleger des Turmdrehkrans.
Links: Überdeckungen der Bilder benachbarter Kranpositionen mit inkrementellem Rotationswinkel α. Die Rechtecke einer Farbe symbolisieren die Bodenspuren der Bilder einer Kamera aus der Vogelperspektive.

des Krans befindet. Da die lange Sensorseite in Richtung des Auslegers verläuft, haben Bilder benachbarter Kameras eine Längsüberdeckung von 75%.
Die Querüberdeckung benachbarter Bilder derselben Kamera ergibt sich nach TUTTAS et al. (2016) aus Drehradius, Kameraparametern und inkrementellem Rotationswinkel α. Letzterer beschreibt den Rotationswinkel des Krans zwischen zwei Kranpositionen, an denen die Kameras A-E ausgelöst werden.

2.2 Passpunkte und Referenzdaten

Die Georeferenzierung der photogrammetrischen Punktwolke erfolgt mithilfe von Passpunkten, die tachymetrisch im UTM-Koordinatensystem eingemessen werden.
Neben der photogrammetrischen Aufnahme erfolgt auch eine Erfassung der Baustelle via Laserscanning. Zeitgleich mit der Aufnahme des photogrammetrischen Bildverbandes wird ein terrestrischer Laserscan (TLS) durchgeführt, der als Referenz der photogrammetrischen Punktwolke dient und für die Analyse der Vollständigkeit herangezogen wird.

3 Methodik

Zuerst erfolgt die Planung der Aufnahme des Bildverbandes anhand einer Simulation. Diese ermöglicht die Planung der Passpunktverteilung sowie die Festlegung des inkrementellen Rotationswinkels. Anschließend wird die Aufnahme des realen Bildverbandes und dessen photogrammetrische Auswertung im Hinblick auf Genauigkeit und Vollständigkeit durchgeführt.

3.1 Matlab-Simulation mit Schnittstelle zu Inpho UASMaster

Die in Matlab implementierte Simulation des Krankamerabildverbandes hat eine Schnittstelle zur photogrammetrischen Software UASMaster (TRIMBLE 2022). Das Schema in Abb. 2 stellt den Ablauf der Simulation dar. Als Eingabeparameter benötigt die Simulation Parameter des Bildverbandes und der Krangeometrie, sowie Angaben zu Messunsicherheiten. Außerdem wird als Input-File eine Punktwolke des Baustellenbereichs importiert, die vor der Aufnahme des Bildverbandes mit einem mobilen Laserscanner erfasst wird. Das Simulationsprogramm wählt einen Bruchteil dieser Punkte aus und betrachtet sie als Verknüpfungspunkte zwischen den einzelnen Bildern des simulierten Bildverbandes.

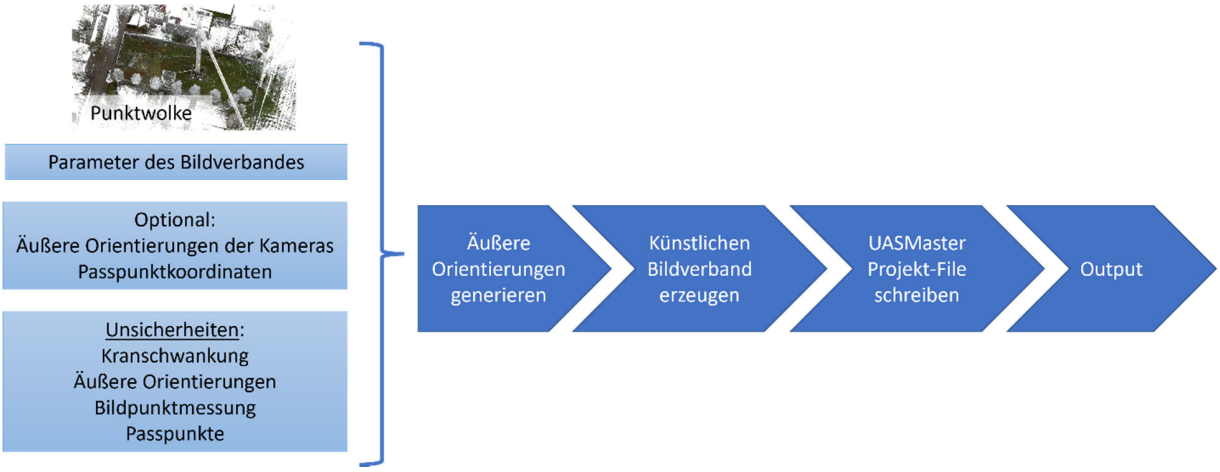


Abb. 2: Ablauf der Simulation des Krankamerabildverbandes

Anhand der Eingabeparameter berechnet die Simulation die äußeren Orientierungen der Kameras im Bildverband. Diese werden anschließend verwendet, um Objektpunktkoordinaten der Punktwolke über die Projektionsmatrix der Projektiven Geometrie in das Pixelkoordinatensystem im Bildraum zu überführen (LUHMANN et al. 2019). So ergeben sich schließlich Verknüpfungspunkte des simulierten Bildverbandes, die gemeinsam mit Informationen zu den äußeren Orientierungen der Kameras und den Passpunktkoordinaten in das UASMaster Projekt-File geschrieben werden. Mit dem Ausgabefile der Simulation kann im Anschluss eine Bündelblockausgleichung in UASMaster berechnet werden, die Ergebnisse zu den Objektpunktgenauigkeiten liefert. Anhand der Simulation werden die Objektpunktgenauigkeiten für verschiedene Passpunktverteilungen und inkrementelle Rotationswinkel des Krans untersucht. Basierend auf diesen theoretischen Voruntersuchungen wird für die reale Bildaufnahme mit den Krankameras ein inkrementeller Rotationswinkel von 5° festgelegt. Außerdem ergeben die Voruntersuchungen die in Abb. 3 dargestellte Konfiguration an 12 Pass- und 14 Kontrollpunkten.
Darüber hinaus wird die Simulation im Anschluss an die Auswertung des realen Bildverbandes mit optimierten Eingabeparametern durchgeführt. So kann auch die Objektpunktgenauigkeit außerhalb des Baustellenbereichs betrachtet werden, wo sich keine Kontrollpunkte befinden. Dazu wird die Differenz zwischen ausgeglichenen Objektpunktkoordinaten der Simulation und den tatsächlichen Koordinaten der Punktwolke betrachtet und graphisch visualisiert.

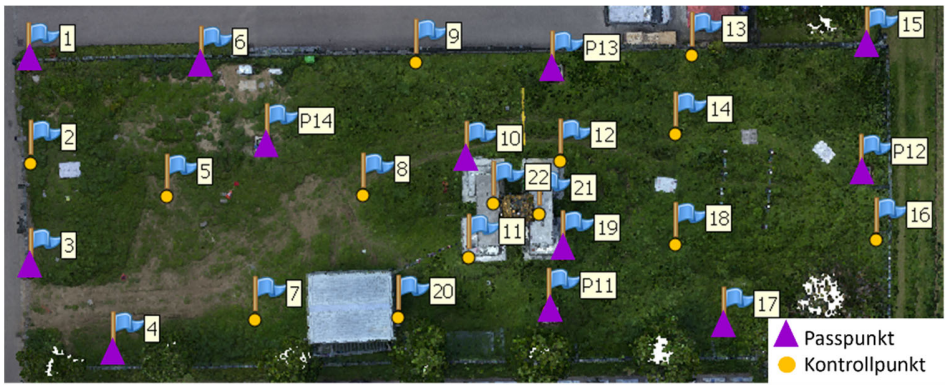


Abb. 3: Pass- und Kontrollpunktkonfiguration des Krankamerabildverbandes; Passpunkte, deren Name mit einem P beginnt, sind an einem Messpfeiler angebracht.

3.2 Auswertung des realen Bildverbandes

Im zweiten Schritt erfolgt die Aufnahme und Auswertung des realen Bildverbandes, wobei die Auswertung im Wesentlichen mit der photogrammetrischen Software Agisoft Metashape erfolgt (AGISOFT METASHAPE 2021). Betrachtet werden dabei die Genauigkeit und Vollständigkeit der photogrammetrischen Rekonstruktion.

Die **Objektpunktgenauigkeit** wird anhand der Residuen an den Kontrollpunkten bewertet. Die betrachtete Größe ist hierbei die Wurzel der mittleren Fehlerquadratsumme bzw. der root-mean-square-error (RMSE). Die Objektpunktgenauigkeit wird von verschiedenen Faktoren beeinflusst. Hier betrachtete Aspekte sind der inkrementelle Rotationswinkel, die Passpunktkonfiguration und die Kamerakalibrierung. Zur Bestimmung der Parameter der inneren Orientierung und Verzeichnung der Kamera werden drei verschiedene Kalibrierverfahren betrachtet. In Var. 1 werden alle Parameter aus einer vorherigen Testfeldkalibrierung übernommen, in Var. 2 werden die Parameter ausschließlich aus einer Simultankalibrierung bestimmt und in Var. 3 erfolgt eine Simultankalibrierung der ‚leitenden Parameter', Kamerakonstante und Bildhauptpunkt, wobei die übrigen Parameter in der Testfeldkalibrierung bestimmt werden. Die Konstellation in Variante 3 ergibt sich aus Untersuchungen von CLEDAT et al. (2020), wo diese Kombination der Parameterschätzung zu besonders hohen Objektpunktgenauigkeiten führte.

Um die **Vollständigkeit** zu beurteilen, werden TLS-Daten als Referenz verwendet. Da die Pfadplanung auf Basis eines DOM erfolgen soll, wird sowohl aus der Laserscanning- als auch aus der photogrammetrischen Punktwolke ein DOM der Rastergröße 10 cm generiert. Ein Vergleich der beiden Oberflächenmodelle wird durch Bildung der Cloud-to-Cloud-Distance (C2C) durchgeführt, die in CloudCompare implementiert ist (CLOUD COMPARE 2021).

4 Ergebnisse

Untersuchungen der Residuen an den Kontrollpunkten zeigen, dass sich im Baustellenbereich Objektpunktgenauigkeiten unter 5 cm ergeben. Diese Genauigkeit wird unter Durchführung einer Simultankalibrierung der ‚leitenden Parameter' (Var. 3) erreicht. Bei Betrachtung der Objektpunktgenauigkeiten für verschiedene inkrementelle Rotationswinkel in Tab. 1 zeigt sich, dass auch

mit steigendem Rotationswinkel hohe Objektpunktgenauigkeiten erreicht werden können, obwohl weniger Bilder zur Verfügung stehen.

Tab. 1: Objektpunktgenauigkeit in Abhängigkeit vom inkrementellen Rotationswinkel α und der Querüberdeckung, welche sich jeweils auf benachbarte Bilder der mittleren Kamera C bezieht

Rotationswinkel α	5°	10°	20°	35°	50°
Querüberdeckung	93,3 %	86,7 %	73,9 %	56,2 %	41,5 %
Anzahl der Bilder	360	180	90	52	36
RMSE [cm]	2,7	2,6	2,7	2,1	2,4

Betrachtet man zusätzlich die Entfernung der Kontrollpunkte vom Rotationszentrum des Krans, sind für hohe Objektpunktgenauigkeiten in Krannähe große inkrementelle Rotationswinkel ausreichend. Wenn auch Objekte am Baustellenrand von Interesse sind, sollte der Rotationswinkel verringert werden, um diese mit hoher Genauigkeit lokalisieren zu können.
Außerhalb des Baustellenbereichs ergeben die Untersuchungen anhand der Simulation eine Abnahme der Objektpunktgenauigkeit der photogrammetrischen Rekonstruktion mit Vergrößerung des inkrementellen Rotationswinkels (Abb. 4).

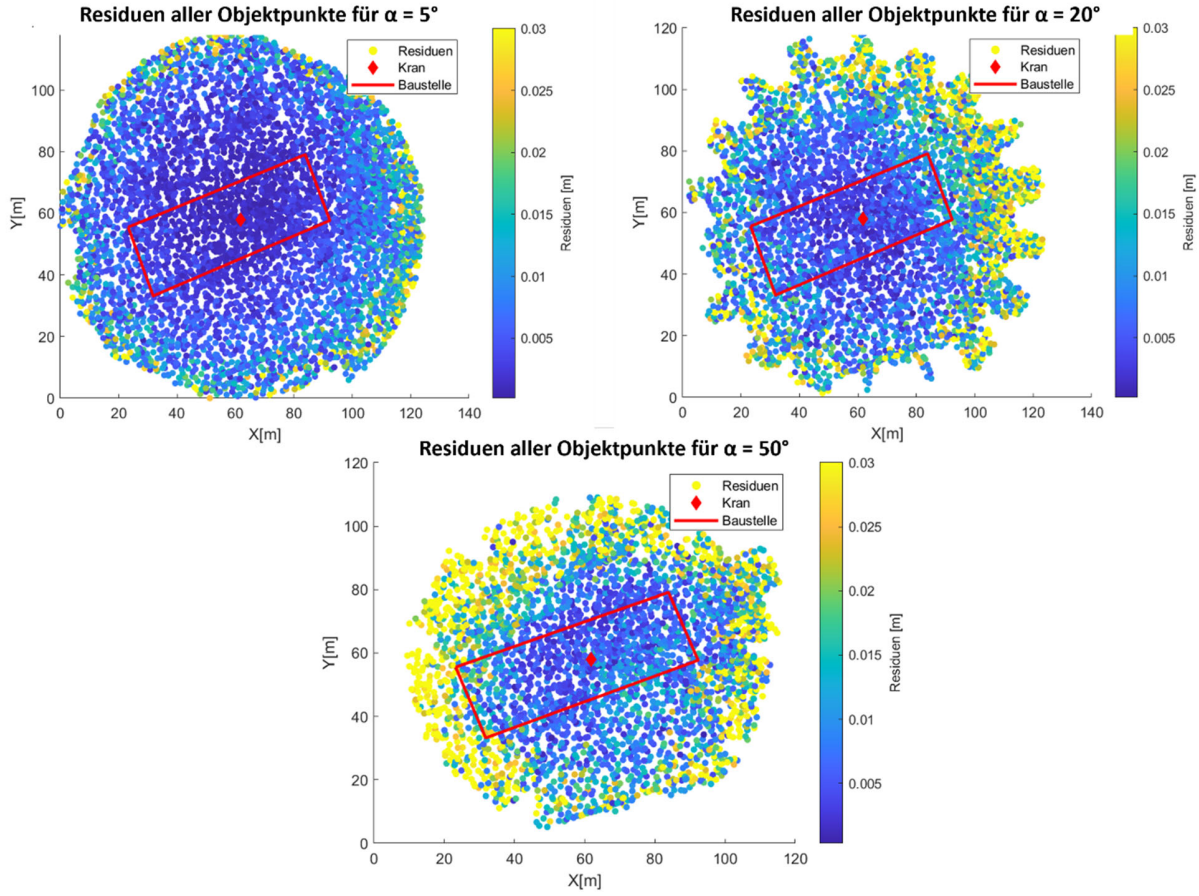


Abb. 4: Objektpunktgenauigkeiten außerhalb des Baustellenbereichs: Abweichungen zwischen den ausgeglichenen Objektpunktkoordinaten der Simulation aus UASMaster und den tatsächlichen Koordinaten der Punktwolke; dargestellt für verschiedene inkrementelle Rotationswinkel α des Krans; zur geeigneten Visualisierung wird eine Obergrenze von 3 cm festgelegt.

Untersuchungen der Vollständigkeit der photogrammetrischen Rekonstruktion zeigen eine Abnahme der Vollständigkeit mit Vergrößerung des inkrementellen Rotationswinkels α. Unabhängig vom Rotationswinkel zeigt sich in Abb. 5 die schwierige photogrammetrische Rekonstruierbarkeit von Vegetation, wie Bäumen oder der Grasfläche, die das gesamte Baustellengebiet durchzieht (rot). Die sich darauf befindenden Objekte wie Krangewichte, Holzpaletten oder Stangenketten können allerdings mit einer sehr hohen Genauigkeit von 5 cm und besser rekonstruiert werden (blau). Es zeigt sich, dass Objekte mit einer Mindestbreite von 3 cm in der photogrammetrischen Rekonstruktion enthalten sind, wenn diese sich ausreichend vom Hintergrund abheben können. Dies entspricht der Dicke des Bauzauns, an dem dieses Verhalten erkennbar wird.

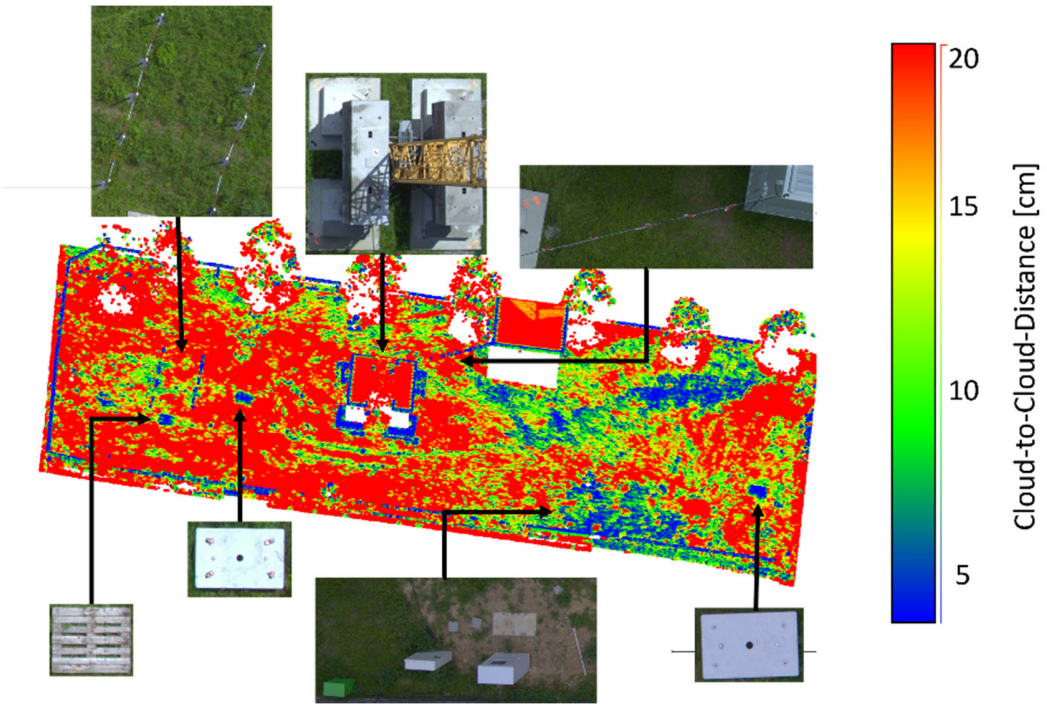


Abb. 5: Vollständigkeit der photogrammetrischen Rekonstruktion: Differenz zwischen Photogrammetrischem-DOM (α = 5°) und Laserscanning-DOM mit festgelegter Obergrenze von 20 cm.

5 Fazit & Ausblick

Die Betrachtungen der Residuen an den Kontrollpunkten zeigen, dass im Baustellenbereich eine mittlere Objektpunktgenauigkeit von 2,5 cm erreicht wird. Diese Genauigkeiten sind erreichbar, wenn die Parameter der Kameraverzeichnung mit einer Simultankalibrierung der ‚leitenden Parameter' bestimmt werden. Der Vergleich mit den Laserscanning-Referenzdaten zeigt, dass Objekte mit einer Mindestbreite von 3 cm in der photogrammetrischen Rekonstruktion enthalten sind, sofern diese sich ausreichend vom Hintergrund abheben. Untersuchungen verschiedener inkrementeller Rotationswinkel des Krans zeigen, dass für eine vollständige Rekonstruktion der Objekte kleine Rotationswinkel zu wählen sind. Die erreichten Objektpunktgenauigkeiten liegen innerhalb des Genauigkeitsbereichs, in der die Regelung des Krans arbeitet. Aufgrund der Pendelbewegung

des Hubseils ist ohnehin keine Positionierung des Kranhakens im mm-Bereich möglich. Daher bildet ein solches photogrammetrisch generiertes DOM eine geeignete Grundlage der automatisierten Pfadplanung durch die Baustellenumgebung. Um ein aktuelles Oberflächenmodell zu erhalten, muss dieses allerdings in Echtzeit generiert werden. Es wird sich zeigen, welche Qualität der Rekonstruktion unter diesen Umständen erreichbar ist.

6 Danksagung

Vielen Dank an die Firma Trimble Inpho für die Bereitstellung ihrer Software UASMaster für die Nutzung in dieser Arbeit. Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder – EXC 2120/1 – 390831618.

7 Literaturverzeichnis

AGISOFT METASHAPE, 2021: http://www.agisoft.com/downloads/installer/, Version 1.7.2.

CLEDAT, E., CUCCI, D., A. & SKALOUD, J., 2020: Camera Calibration models and methods for corridor mapping with UAVs. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., **V**-1-2020.

CLOUD COMPARE, 2021: http://www.cloudcompare.org/, Version 2.10.2

LUHMANN, T., ROBSON, S., KYLE, S. & BOEHM, J., 2019: Close-Range Photogrammetry and 3D Imaging. De Gruyter, 277-295.

STATISTISCHES BUNDESAMT 2018: Arbeitsproduktivität je Erwerbstätigenstunde im Vergleich 1991-2018, https://www.capmo.de/baulexikon/baubranche/, letzter Zugriff am 14.11.21.

TRIMBLE, 2022: Trimble Inpho UASMaster. https://de.geospatial.trimble.com/products-and-solutions/trimble-inpho-uasmaster, letzter Zugriff am 08.02.2022.

TUTTAS, S., BRAUN, A., BORMANN, A. & STILLA, U., 2016: Evaluation of acquisition strategies for image-based construction site monitoring. Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B5, 733-740, https://doi.org/10.5194/isprs-archives-XLI-B5-733-2016.

Edge-related Deep Learning for Semantic Segmentation

MONA GOEBEL[1], LINA E. BUDDE[1] & DOROTA IWASZCZUK[1]

Abstract: The identification of Land Use and Land Cover is an important task for creating maps or monitoring surface changes. The results in this work were obtained using a deep neural network called VGG19-Unet. Influenced by the increasing success of separately implementing edge-extracted inputs in a model, this work used the Canny method to create such edge extracted images. In the final model, one band was discarded from each of the original RGB and NIRRG images and replaced with an edge-extracted band, generating a closer connection between colour and edge bands. This model simultaneously learned eight classes. Segmentation results were found to be predominantly sharper when edge-extracted bands were included, and the model was more confident in its choice of classes. Lastly, improvements were achieved for objects with clear edges, such as buildings, as well as for objects with unclear edges, such as vegetation.

1 Introduction

The identification of Land Use and Land Cover (LULC) is an important task for creating maps or monitoring surface changes. Applications for such maps include environmental monitoring, agriculture, forestry and urban planning. These classification tasks combine remote sensing and computer vision (KAMPFFMEYER et al. 2016). To do this, an accurate raster map is needed in which each pixel is assigned to a specific class. This is called semantic image segmentation or pixel-level prediction and is a complex task in scene understanding (ZOU et al. 2019; LIU et al. 2019).With the introduction of Fully Convolutional Networks (FCNs) by LONG et al. (2015), this area became known and computationally possible (ZOU et al. 2019). Since then, models can also generate spatial maps and not only class labels per image (MARMANIS et al. 2017). Moreover, research advance in this field is not limited to Earth observation, but is also relevant for automatic transportation, robotics and overall automatic image analysis and sorting.

The success of model prediction is highly dependent on image pre-processing and the chosen methods and model architecture. Overall, the Deep Learning approach is more successful than methods such as support vector machines or decision trees (MA et al. 2019). Recently, researchers integrated edge detection into models to achieve sharper and cleaner class boundaries. Examples are EdgeNet introduced by CHEN & BARRON et al. (2016) including an edge-preserving mechanism, as well as models by MARMANIS et al. (2017) or ZOU et al. (2019) which use Holistically-Nested Edge Detection (HED) to improve their models.

This work investigates whether adding separately extracted edges to a deep neural network returns better results than without, even though convolutional filters within an FCN already extract edges. Furthermore, an investigation is made whether objects with clear edges, such as buildings, as well

[1] Technical University Darmstadt, Department of Civil and Environmental Engineering Sciences, Remote Sensing and Image analysis, Franziska-Braun-Str. 7, D-64287 Darmstadt,
E-Mail: [mona.goebel, lina.budde, dorota.iwaszczuk]@tu-darmstadt.de

as objects with unclear edges, such as vegetation, can be predicted better with the integrated edge extraction. For this, the focus will be on pixel-wise semantic segmentation with satellite images of urban areas using supervised Deep Learning. Training and testing of the model will be performed using the SemCity Toulouse benchmark by ROSCHER et al. (2020). The authors discuss in the published paper (ROSCHER et al. 2020) only instance segmentation and solely for buildings. In this work, all seven classes plus unclassified pixels (equals to eight classes) will be addressed simultaneously.

2 Background and Related work

2.1 Edge Detection

Detecting edges in images is essential in order to simplify images, yet preserving important features. Therefore, an errorless edge detection is crucial for the model prediction. If edges are missed, the model will not notice the importance of a specific pixel and on the contrary, if too many unimportant edges are present, then the model will not recognise the meaningful ones. Summarised, three key points of a successful edge detection can be listed: First, the result should have a low chance of missing the real edge point, while also not misclassifying non-edge points. Secondly the marked edge point should be near the centre of the real edge point. Lastly, the possibility of showing several maxima for one edge should be low, hence resulting only to one signal for one edge. (CANNY, 1986)

The edge detection function implementation in the python package scikit-image was applied in this work and is based on the paper by CANNY (1986). Some adjustments were made to the original idea, which led to the following process: At the beginning the image is smoothed with derivative of the Gaussian operator, just as in the base algorithm by CANNY (1986). Thereafter the Sobel operator in the horizontal and vertical direction is calculated, detecting high pixel intensities. The wide intensities are narrowed to a 1-pixel wide curve, by analysing the gradient directions. Possible orientations can be horizontal, vertical, diagonal and anti-diagonal. A check is then made on the focused pixel to see if it is a maximum, a minimum, or neither. By calculating the interpolation of the four directions, the result has a higher probability of being accurate than choosing the direction that is the closest. Finally, the thresholds define all pixels above the upper threshold to be edges, while the pixels below the lower threshold are background. The pixels in between are edges if they are maxima of the surrounding neighbouring pixels. (SCIKIT-IMAGE n.d.)

Just as the Laplacian of Gaussian, the Canny edge detector is categorised as Gaussian-based. It is known to be robust to noise and to be the best method, by reason of detecting edges without modifying or influencing the image features (KRISHNAN et al. 2017). The idea of including different resolutions is also covered in the Canny operator (HAFFNER & RAVAS 2014). Additionally, the Canny operator's success is extremely dependent on the choice of thresholds and width of the Gaussian σ. The latter sets the frame of neighbouring pixels that are included in the decision whether a point is an edge or not. Setting an unsuitable threshold can suppress true edges while reducing noise (HAFFNER & RAVAS 2014).

2.2 VGG and U-Net concept

The model used in this work is based on the VGG convolutional neural networks and the U-Net concept. ConvNet was originally presented by SIMONYAN & ZISSERMAN (2014) and is commonly called VGG. It won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014. Different VGG networks were adapted in many works, such as ResNet (HE et al. 2015) and DeepLab (CHEN et al. 2017).

U-Net was first presented by RONNEBERGER et al. (2015) and represents an encoder-decoder. The paper mentions the utility for pixel-wise classification for biomedical images, as well as the difference between a large training set like ImageNet and the typical small dataset in biomedical cases. The authors also aimed to reduce redundancy in training and improve the processing time of the approaches commonly used at that time. Their FCN idea was based on a balance between good localisation and using the context around a pixel (RONNEBERGER et al. 2015). Original about the U-Net architecture was the symmetry between the encoding and decoding parts. During compression within increasing depth of the model, fine details are lost. Skip connections merge outputs from encoding and decoding, thus reintroduce details and improve prediction results.

2.3 Benchmark

SemCity Toulouse was published in 2020 by a cooperation of research establishments in Germany, France and Switzerland (ROSCHER et al. 2020). The data focuses on the densely populated city of Toulouse in France, where each building is labelled separately. Taken from the Worldview-2 satellite in April 2011, the 8-band image was tiled into 16 smaller images, each covering an area of about 3 km². As the paper states, tiles number 3 and 7 are for training, while 4 and 8 for testing. The benchmark includes multispectral imagery, panchromatic imagery as well as corresponding ground truths in instance and semantic segmentation (ROSCHER et al. 2020).

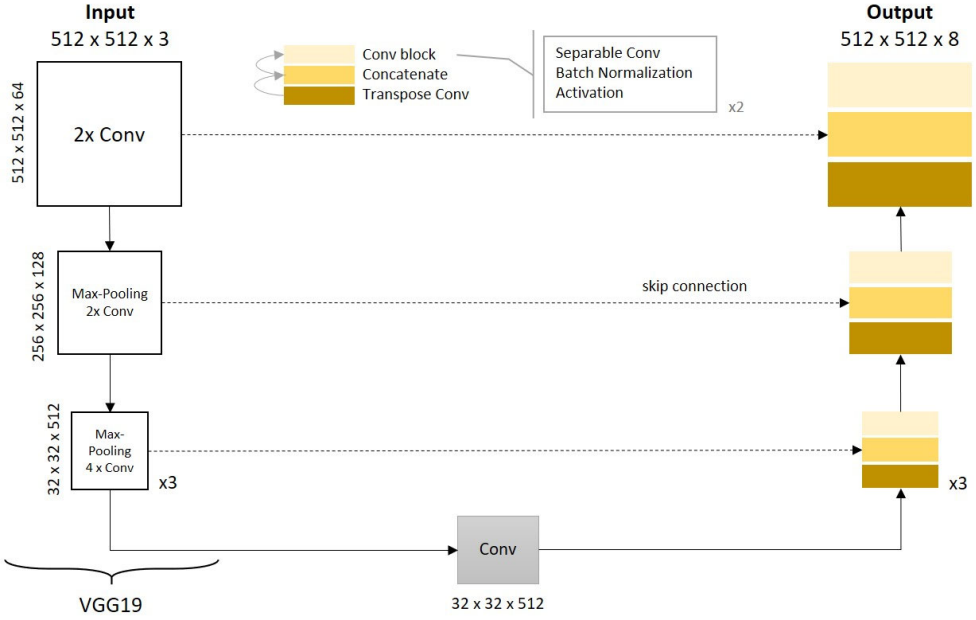


Fig. 1: VGG19-Unet architecture, shown without dropout layers. Conv: convolution, AAA x BBB x C: image size of the layer output, “x3”: sequence is run three times

3 Method

The following section describes the model architecture, model uncertainty, edge extraction, input and finally metrics. The experimental process included in the first phase to investigate the best combinations of layers and channels as well as settings of parameters. Only one comparison will be shown in this work, namely the different combinations of image channels given to the model. The best results were fused into a "final model". Lastly, to interpret the results of this model, a "reference model" was additionally run. The final and the reference model have the same architecture and settings. However, they differ in the input images, as the reference model only receives RGB and NIRRG images and none edge extracted images. The models were run on a computer with 128 GB RAM, an AMD Ryzen Threadripper 2970WX processor with 24 cores, as well as an NVIDIA TITAN RTX graphic card with 24GB.

3.1 Model architecture

The model used in this work is called VGG19-Unet, a VGG19-based encoder-decoder model extended into the U-Net form (Fig. 1). The model code was adapted from VGG19 Unet by Nikhil Tomar (NIKHILROXTOMAR 2021). VGG19 is implemented with Keras (keras.io) which comes with pre-trained weights on the ImageNet benchmark. VGG19 was chosen as the backbone because it is easy and not as complex as the state of the art (Xception, DeepLabV3+), as well as being widely used as a backbone. As shown in Fig. 1, the encoding part of the U-Net architecture consists of an adapted version of VGG19 (on the left side of the figure). Dimensions such as 512 × 512 × 64 represent the output dimensions of the images after each block. The decoding part was added on top of the VGG19 model. The significant difference in the decoding part is the use of depth-wise separable convolution instead of normal convolutional layer. Skip connections are implemented before each max-pooling layer, which merges the layers from the encoder and the previous layer via concatenation. When merging the layers from encoding and decoding, the images must have the same dimensions. Lastly it has an initial input size of 512 × 512 × 3 pixels. The output has the same image size but now with 8 channels instead of 3. For the reason of each channel storing the probabilistic distribution for each pixel and class (Softmax output). Dropout layers were included in the final and reference model after each convolution block with 20 % dropout rate and one dropout layer prior the output layer with a 50 % dropout rate. For the visual interpretation of the result and discussion, the highest probability for each pixel is taken and assigned to that class, resulting in a size of 512 × 512 × 1 pixels. As optimiser, the algorithm Adam by KINGMA et al. (2014) and categorical cross entropy as loss function were chosen.

3.2 Model Uncertainty

To better understand the confidence of the final model, the dropout layers were active during training and testing in order to use the Monte Carlo dropout (GAL & GHAHRAMANI 2016). Since the results differ with this method with each prediction run, it was carried out ten times. Then the median was determined, creating a more stable result. Thereafter, for each class the corresponding pixel indices were extracted from the target labels. With such a class selection the prediction includes, for instance, only the building pixels. Finally, the median and standard deviation for each class channel were calculated, based on the Softmax output.

3.3 Input and edge extraction

As input data, the images from the benchmark SemCity Toulouse were taken, showing the French city Toulouse. The utilised panfused images have a spatial resolution of 0.5 m and are either RGB (red, green, blue) or NIRRG (near-infrared, red, green) images. All images that are given annotated were used, which are images 3, 4, 7 and 8. The original size of 3504 × 3452 × 3 pixels was tiled into smaller squares in the form of 512 × 512 × 3 pixels, with an overlap of 50 %, which results in 169 tiles. In addition, the images were normalised to range from 0 to 1 instead of reaching between 0 and 255. NIRRG and RGB images were treated in the same way. Edge extraction was carried out based on either of the three RGB or NIRRG channels. The extraction was implemented via the Canny function of the feature package of scikit-image, based on CANNY (1986). The width of the Gaussian (σ) turned out to be best fitting with σ = 3.6 for most of the channels. Each tile median is used as a guide for setting the lower and upper thresholds in the Canny function.
Three edge-colour combinations were tested, see Fig. 2 for an overview. The best method was chosen for the "final model". The final model was trained with 906 randomly selected tiles from 1352 tiles, while the remaining 446 tiles were kept for testing.

3.4 Metrics

Comparing the results in numbers, metrics F-score and overall accuracy were selected. The F-score is based on precision and recall, where the first indicates if the found objects are true or false, while recall indicates whether all relevant objects were found from all the objects that should have been found. With β defining the balance between precision (p) and recall (r), the equation is as follows (CHINCHOR 1992):

$$F_\beta = \frac{(1+\beta^2)\times p \times r}{\beta^2 \times p + r}.$$

Normally β is set to 1, just as in the SemCity Toulouse paper (ROSCHER et al. 2020), further called F1. The weighted F1 value takes into account the number of pixels per class, to compensate for imbalance of classes.

4 Results and Discussion

First, models called RGB edges, NIRRG edges and mixed edges are compared and discussed. In the second part, the favourable input method is chosen for the "final model". Overview of the model inputs is given in Fig. 2.

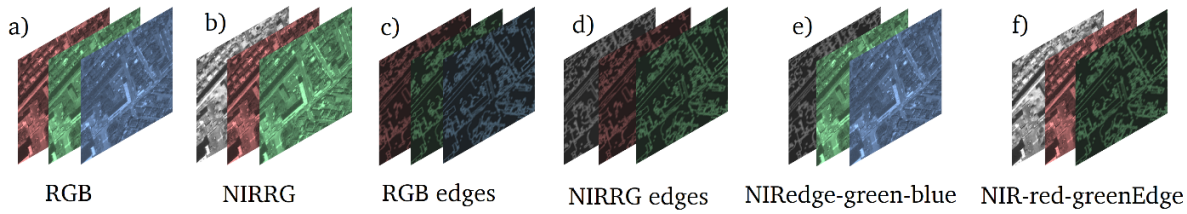


Fig. 2: Overview of model inputs. Model "RGB edges" includes input a), b) and c); "NIRRG edges" a), b) and d); model "mixed edges" and the final model include e) and f); the reference model only a) and b)

4.1 Edge input comparison

The model with RGB edges returns good results, for example in detached housing groups located in the upper left corner in each tile of Fig. 3, however the model has difficulties to distinguish between impervious surfaces and sport venues. In comparison, the model with NIRRG edges as input is challenged between pervious and impervious surfaces as well as confusing impervious surfaces and water pixel. With mixed edges the model returns overall good results. Both NIRRG and mixed edges performing less accurate in the detached housing area. In terms of numbers, the model with RGB edges reached an overall accuracy of 0.56 and a weighted F1 score of 0.55 (Tab. 1). As for the model with NIRRG edges the values are both 0.54 respectively. Lastly, the model with mixed edges achieved an overall accuracy and weighted F1 score of 0.70. This underlines that a closer connection between colour bands and edge bands improves model performance. Consequently, mixed edges were chosen as input for the final model.

4.2 Final model results

The final model needed 1 day and 4 hours and ended training after 27 epochs, whereas the reference model needed 29 epochs and 2 hours more. Fig. 4 shows the visual results. Edges in the final model look overall sharper than from the reference model. Both models seem to have similar problems, for example the objects in white were not detected (Fig. 4, bottom row). Nonetheless, the final model has troubles in the detached housing area (Fig. 4, top row), unlike the reference model. An explanation for this are the extracted edges of this particular area, illustrated in Fig. 5. Thus, if the extracted edges are in a bad quality, the prediction accuracy will decline. This also shows the influence of given edges on the model's success.

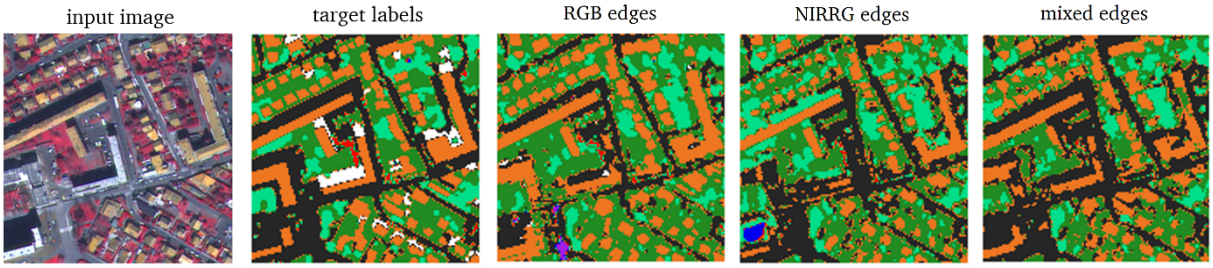


Fig. 3: Results of models with RGB edges, NIRRG edges or mixed edges. The colours represent buildings in orange, impervious surfaces in black, pervious surfaces in dark green, high vegetation in light green, cars in red, water in blue, sport venues in purple and void in white

Tab. 1: Numerical results of models with RGB, NIRRG and mixed edges.
Precision, recall and F1 are given as weighted average, accuracy as overall average

Input	Precision	Recall	F1	accuracy
RGB edges	0.60	0.56	0.55	0.56
NIRRG edges	0.60	0.54	0.54	0.54
Mixed edges	0.72	0.70	0.70	0.70

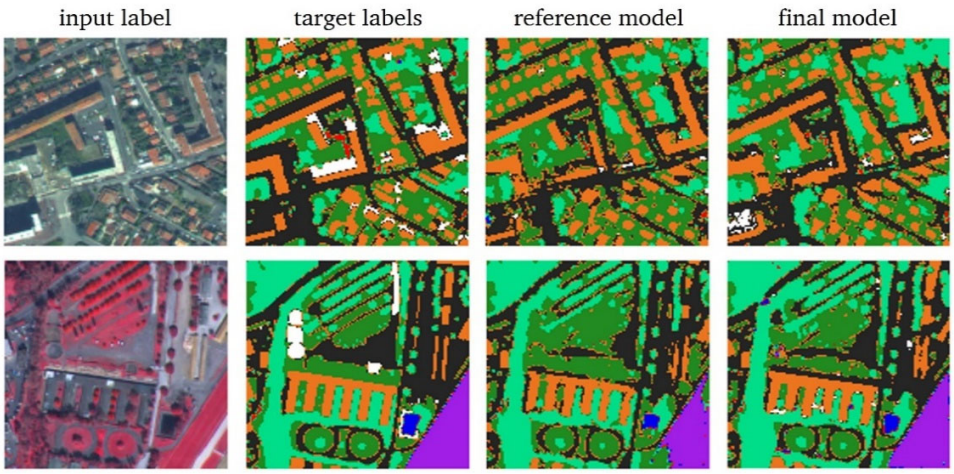


Fig. 4: Visual results of the reference and final model. The colours represent buildings in orange, impervious surfaces in black, pervious surfaces in dark green, high vegetation in light green, cars in red, water in blue, sport venues in purple and void in white

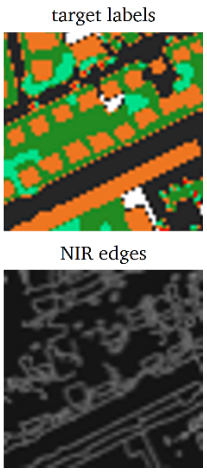


Fig. 5: Example of poorly extracted NIR edges

4.2.1 Numerical results

For the final model the highest F1 score was achieved by the building class (0.90), while the reference model reached 0.91. One class which was only rarely detected in all prior models, namely the sport venues, was now found 83 % of times (recall) by the final model, which is 12 % more than the reference model. Leading to a sport venues F1 value of 0.83 for the final model and 0.81 for the reference model. The least success was found in cars with an F1 score of 0.69, while the reference model reached 0.73. Lower results can derive from incorrect edges that guide the model into false directions. The overall F1 and accuracy is 0.80 for the final model, and 0.82 for the reference model.

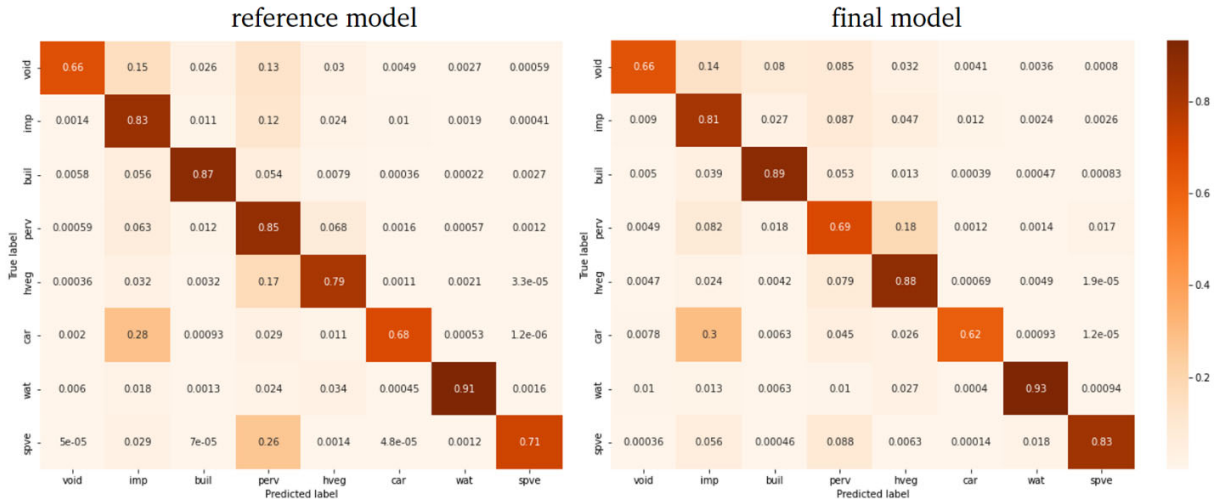


Fig. 6: Normalised confusion matrices of the reference and final model, with the latter including mixed edges. Given numbers represent consensus between predicted and true labels, with 1.00 as complete match. The abbreviations stand for: impervious surface (imp), building (buil), pervious surface (perv), high vegetation (hveg), water (wat) and sport venues (spve)

4.2.2 Model uncertainty

The class-wise recall for the reference and final model are portrayed in the confusion matrices in Fig. 6. Uncertainties of the reference and final model are shown in Fig. 7 for three classes, calculated as explained in the method (section 3). Selected exemplary classes were chosen because they have clean edges (building), unclear edges (high vegetation) or are a challenging class for the models (sport venues). The two confusion matrices show a good diagonal line which implicates that the models perform good and are rather confident.

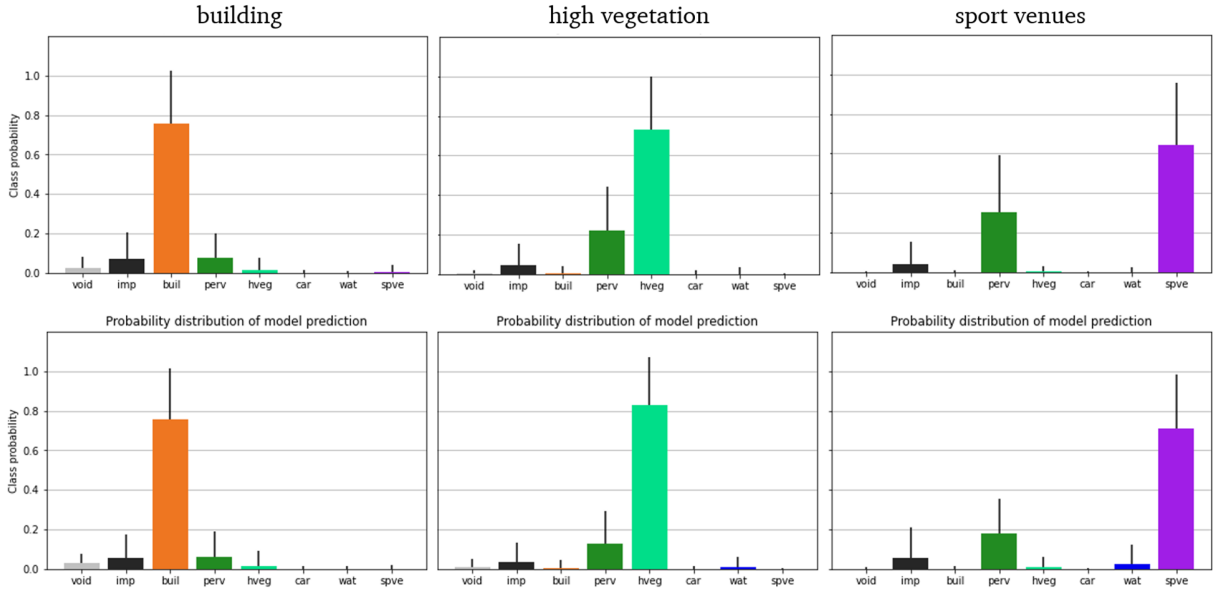


Fig. 7: Graphs show the probabilistic distribution per class of the reference model (top) and final model (bottom), with the latter including mixed edges. Median and standard deviation were calculated only on the pixels that belong to that specific class. For example, for building the pixel that are classified as buildings in the true labels were extracted. The black vertical line represents the standard deviation. See Fig. 6 for abbreviation explanations

The highest recall was reached in the water class for both models, with 0.91 in the reference model and 0.93 by the final model. The lowest percentage of consensus is the car class for both models. Cars were often classified as impervious surface, which is reasonable, as they are mostly surrounded by roads and parking lots. Other than that, they are small objects and only occupy 2 % of the total annotated pixels. Unfortunately, cars were often annotated in the benchmark in a poor quality, thus the true labels could in fact not correspond to the truth.

For the building class, the final model achieved an average confidence of approx. 0.76 ±0.26, very similar to the reference model (Fig. 7). Building pixels matched by 0.89 in the confusion matrix for the final model. Concerning high vegetation with a confidence of 0.83 ±0.24, the final model confused it at times with pervious surface. This is not surprising, as they can be very similar in the green and NIR channel. Including edges in the input dataset, improved the confidence in this class by 0.10 ±0.02, and recall from 0.79 to 0.88. As discussed before, the reference model had more difficulties with the sport venues class and misclassified 26 % of pixels as pervious surface. Naturally, these sites can include pervious surface (e.g. football fields), impervious surface (e.g. Tartan track), water (e.g. swimming pools) and buildings (e.g. stadium). For these reasons it is the

most diverse class and difficult to classify. Thus, the finale model reaches an average confidence of 0.71 ±0.28, significantly better than the reference with 0.64 ±0.32. Interestingly, the other way around, the models only rarely misclassify pervious surface as sport venues.

4.2.3 Comparison with other work

For comparison, MARMANIS et al. (2017) achieved with their most comparable network HED+FCN-H-V which is based on an FCN plus VGG weights, an overall accuracy of 0.85. It includes HED method and the ISPRS Vaihingen benchmark and Potsdam benchmark as input (include ground elevation data). Their overall accuracy is 0.05 better than the final model in this work. However, the model in this work only involves 23×10^6 trainable parameters, while their model has 300×10^6 (MARMANIS et al. 2017). On the Virtual KITTI dataset, semantic segmentation was performed with SSDNet-Sem by ZOU et al. (2020). These images include 13 classes plus void. Their complex model includes boundary-detection without depth maps. SSDNet-Sem needed 25×10^6 parameters and achieved an accuracy of 0.76 (ZOU et al. 2020). Therefore, it needed a few more parameters than the final model discussed in this work, as well as reaching 0.04 less in accuracy. In conclusion, the number of classes to predict is next to the architecture and input another crucial part for the success of a model, as the Vaihingen and Potsdam benchmark include 6 classes, this work 8 and the KITTY dataset 14 classes.

To summarise, the visual results seemed to be generally better in the final model. Contrarily, overall accuracy and (weighted) F1 value was 0.02 higher in the reference than the final model. However, recall was better in the final model for the classes building, high vegetation, water and sport venues. These classes involve clear and unclear edges as well as poorly distributed and challenging classes. Furthermore, the final model showed are clearer diagonal in the confusion matrix, thus delivers a higher confidence, which is also reflected in the class probability outputs (Fig 7). This indicates that the additional weights on the edge pixels guide the model's decisions. Unfortunately, if the edges are extracted in a poor quality, the model will be led into wrong directions during its training. In consequence, errors in the edge extracted images are an explanation for overall better accuracy of the reference model compared to the final model.

5 Conclusion and future work

In this work, experiments of edge extractions based on different image channels and combinations of image bands were analysed. Better results were achieved when constructing a closer connection between colour bands such as red, green, blue and NIR, and edge extracted bands within one image. The final model included these mixed bands, as well as dropout layers and depth-wise separable convolution.

To improve the model, more fine tuning and further variation of the architecture could be tested. In this work edges were extracted via the Canny operator. It is a simple and fast method, however with a more intricate approach the result could be improved. MARMANIS et al. (2017) mentioned to include information of class specific boundaries, so that the model not only classifies pixel by pixel but is also aware which classes are separated. This idea would immensely reduce misclassification of single pixels and reduce processing (MARMANIS et al. 2017).

SemCity Toulouse is a good benchmark for pixel-wise semantic segmentation, especially for remote sensing tasks. However, just as the ISPRS Vaihingen and Potsdam benchmark, it needs a higher annotation quality. Nonetheless the final model also showed good results in under-represented classes such as sport venues and water.
Final model results with edge extracted bands reach an accuracy of 0.80 and the reference model without edges 0.82. However, mean values cannot fully show the challenging semantic segmentation task for urban areas. The confusion matrices showed that including edge extraction improved classes with clear edges such as buildings as well as unclear edges such as high vegetation. Additionally, the confidence of the model and the visual results increased with edges. The final model with edges obtained predominantly clearer edge results than the reference. This work showed the impact of edge extracted bands on the model's performance. Finally, investing in high quality edge extracted images as well as combining edges and colour bands within single images will improve semantic segmentation results and is advised to include in future models.

6 References

CANNY, J., 1986: A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8.6, 679-698. https://doi.org/10.1109/TPAMI.1986.4767851.

CHEN, L., BARRON, J. T., PAPANDREOU, G., MURPHY, K. & YUILLE, A. L., 2016: Semantic Image Segmentation with Task-Specific Edge Detection Using CNNs and a Discriminatively Trained Domain Transform. IEEE Conference on Computer Vision and Pattern Recognition, 4545-4554.

CHEN, L., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A.L., 2017: DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, **40**(4), 834-848, https://doi.org/10.1109/TPAMI.2017.2699184.

CHINCHOR, N., 1992: MUC-4 evaluation metrics. 4th Conference on Message understanding - MUC4 '92, Association for Computational Linguistics, 22, https://doi.org/10.3115/1072064.1072067.8.

GAL, Y. & GHAHRAMANI, Z., 2016: Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. https://arxiv.org/abs/1506.02142.

HAFFNER, O. & RAVAS, R., 2014: Edge detection in image. https://www.researchgate.net/publication/269108799_EDGE_DETECTION_IN_IMAGE.

HE, K., ZHANG, X., REN, S. & SUN, J., 2015: Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition, 770-778, https://arxiv.org/abs/1512.03385.

KAMPFFMEYER, M., SALBERG, A. & JENSSEN, R., 2016: Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks. IEEE Conference on Computer Vision and Pattern Recognition, 1-9, https://doi.org/10.1109/CVPRW.2016.90.

KINGMA, D. P. & BA, J., 2014: Adam: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations, http://arxiv.org/pdf/1412.6980v9.

KRISHNAN, K. B., RANGA, S.P. & GUPTHA, N., 2017: A Survey on Different Edge Detection Techniques for Image Segmentation. Indian Journal of Science and Technology, **10**(4), 1-8, https://doi.org/10.17485/ijst/2017/v10i4/108963.

LIU, X., DENG, Z. & YANG, Y., 2019: Recent progress in semantic image segmentation. Artificial Intelligence Review, 52(2), 1089-1106, https://doi.org/10.1007/s10462-018-9641-3.

LONG, J., SHELHAMER, E. & DARRELL, T., 2015: Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3431-3440, https://doi.org/10.1109/CVPR.2015.7298965.

MA, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. & Johnson, B.A., 2019: Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, **152**, 166-177, https://doi.org/10.1016/j.isprsjprs.2019.04.015.

MARMANIS, D., SCHINDLER, K., WEGNER, J.D., GALLIANI, S., DATCU, M. & STILLA., U., 2017: Classification With an Edge: Improving Semantic Image Segmentation with Boundary Detection. ISPRS Journal of Photogrammetry and Remote Sensing, **184**, https://doi.org/10.1016/j.isprsjprs.2017.11.009.

NIKHILROXTOMAR, 2021: Semantic-Segmentation-Architecture: vgg19_unet.py. https://github.com/nikhilroxtomar/Semantic-Segmentation-Architecture/blob/main/TensorFlow/vgg19_unet.py, last access 21.08.2021.

RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical image computing and computer-assisted intervention, 234-241, https://doi.org/10.1007/978-3-319-24574-4_28.

ROSCHER, R., VOLPI, M., MALLET, C., DREES, L., & WEGNER, J. D., 2020: SemCity Toulouse: A benchmark for building instance segmentation in satellite images. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **V-5**-2020, 109-116, https://doi.org/10.5194/isprs-annals-V-5-2020-109-2020.

SCIKIT-IMAGE, n.d.: Module: feature — skimage v0.19.0.dev0 docs: canny. https://scikit-image.org/docs/dev/api/skimage.feature.html#skimage.feature.canny, last access 03.09.2021.

SIMONYAN, K. & ZISSERMAN, A., 2014: Very Deep Convolutional Networks for Large-Scale Image Recognition. https://arxiv.org/pdf/1409.1556.

ZOU, N., XIANG, Z., CHEN, Y., CHEN, S. & QIAO C., 2019: Boundary-Aware CNN for Semantic Segmentation. IEEE Access, 7, 114520-114528, https://doi.org/10.1109/ACCESS.2019.2935816.

ZOU, N., XIANG, Z., CHEN, Y., CHEN, S. & QIAO C., 2020: Simultaneous Semantic Segmentation and Depth Completion with Constraint of Boundary. Sensors, **20**(3), 635, https://doi.org/10.3390/s20030635.

It's also about timing! When do Pedestrians Want to Receive Route Instructions

ANTONIA GOLAB[1]

Abstract: The preferences of a pedestrian on the timing of route instructions are studied based on observations obtained during an outdoor wayfinding study conducted in an urban environment (N=52): Participants walked two different routes following auditory navigation instructions which they could request whenever and how often they wanted. By means of an exploratory analysis using regression-based modeling techniques, a set of user and environmental variables that influence the observed participant behavior are obtained. Results suggest the relevance of both user and environment characteristics for the prediction of preferred route instruction timing; among these are, e.g., landcover-related variables and psychological personality traits such as e.g., neuroticism (emotional stability).

1 Introduction

One of the core goals in the research of a pedestrian navigation system design is to reduce a person's cognitive load during wayfinding (e.g., GIANNOPOLOUS et al. 2014). Consequently, optimizing how route instructions are phrased and presented on mobile devices has seen much attention by research (e.g., KLIPPEL et al. 2009). However, the question of *when* to present a route instruction to a navigated pedestrian has remained a desideratum – even though there is an obvious potential to facilitate the wayfinding process by presenting an instruction at the optimal point in time: Receiving it too close to the approached decision point might cause uncertainty or presenting it too early might be disturbing to the wayfinder. The virtual reality study by GIANNOPOLOUS et al. (2017) presents the first one addressing this research gap explicitly. Using time-to-event modeling, the authors predicted the distance to a decision point at which a first presentation of a route instruction would be most preferable to the navigated person. The present work builds on this study and expands it considerably by means of an outdoor wayfinding study: On the one hand, similar time-to-event modeling is, again, used to predict the optimal timing of route instruction presentation (*RQ 1*). On the other hand, the problem of repeated requests for navigation instructions is approached, and hence, factors that impact why a person would want to listen to a route instruction for a second time are identified (*RQ 2*). The research question of when to present a route instruction for a first time was studied within an international, interdisciplinary cooperation, together with Prof. Ioannis Giannopolous (Geoinformation, TU Wien) and Dr. Markus Kattenbeck (Geoinformation, TU Wien), and Transportation researcher Dr. Georgios Sarlas (ETH Zürich) who applied the time-to-event modeling to the data which I collected and preprocessed. Our work was published as a journal paper for which I have taken first-author responsibilities, GOLAB et al. (2021). For the submission of the diploma thesis, an analysis on parameters impacting a person's preference to want to listen to a route instruction more than once (*RQ 2*) was added, within which

[1] TU Wien, Research Division Geoinformation, Gußhausstraße 27-29, A-1040 Vienna, E-Mail: antonia.golab@tuwien.ac.at

I applied the Generalized Estimating Equations (GEE) model. In the following sections of this paper, the cornerstones of my work such as the wayfinding study, applied modeling methodology and the core findings are presented.

2 Problem Statement

Tab. 1 illustrates the research questions addressed in this work by means of an example.

Tab. 1: Illustration of research questions (*RQ 1* and *RQ 2*).

The following scenario is considered: Alice walks from her home to the dentist and uses a mobile navigation system providing auditory route instructions. She starts walking and is approaching an intersection at which she does not know if she has to make a turn.	
RQ 1	When will she want to hear the navigation instruction?
RQ 2	Will she want to hear the navigation instruction for a second time before she makes a turn?
Alice might be distracted or feel uncertain by listening to a route instruction earlier or later than personally preferred. She might also want to hear a route instruction more than once in specific circumstances to, e.g., be reminded of it.	

These individual preferences are observed during an outdoor wayfinding study. By modeling the relationship between these observations and user and environment variables, the goal is to identify the influential parameters impacting the observed behavior. Following the paper by GIANNOPOLOUS et al. 2014 who suggest an impact by the user and environment on the complexity of decision situation, the following categories of features are considered in the analysis addressing *RQ 1* and *RQ 2*: **route**-related (e.g., route length), **environmental** (e.g., the density of points of interest), **behavioral** (timing of route instruction requests) and **personal** features (e.g., age).

3 Wayfinding Study

3.1 Materials

Prior to the outdoor experiments, participants answered questionnaires on demographic data (e.g., age, sex), the German language *FRS* questionnaire on spatial strategies (MÜNZER & HÖSCHEL 2011), and a questionnaire on personal characteristics based on the so-called Big Five Personality traits (RAMMSTEDT et al. 2012). A further essential part of this data collection was the outlining of familiar areas and routes using a self-programmed online GIS application. This data was used twofold: Firstly, this information allowed to introduce a binary variable describing whether the person is familiar or unfamiliar with the environment during the experiment, enabling the within-subject design of the experiment which is further explained in Section *3.2*. Secondly, the provided routes by the participants were processed to design routes for the outdoor experiments. For these, spoken navigation instructions were prepared using an algorithm described by ROUSSEL & ZIPF (2017). These were landmark-based, i.e. referencing salient objects in the environment, exclusively designed for turning points (following the concept of spatial chunking described in

KLIPPEL et al 2009) and formulated in German language, e.g., *Biegen Sie bei der Apotheke links ab.* (English translation: *Turn left at the pharmacy*.).

3.2 Outdoor Experiments

The experiment was designed as a within-subject experiment during which each participant walked a familiar route and one which was located in an unfamiliar environment, essentially resulting in two trials per participant. During each trial, locational data of participants was collected using a high-precision PPM 10-xx38 GNSS receiver, head movements were tracked using an Xsens MTi-300 IMU and gaze data was collected using a PupilLabs Invisible mobile eye-tracker. The participant's task during a trial was to walk to a specific destination by following auditory route instructions. In order to directly observe preferences of when people would want to hear a route instruction and whether people would want to hear instructions multiple times, participants were allowed to issue requests for route instructions whenever and how often they wanted by informing the experimenter using a custom-built clicking device (Fig. 1). The experiments took place during the months between June and October 2020. In total, 52 people participated which resulted in 104 conducted trials.

Fig. 1: **A:** A sample participant in full equipment. **B:** GNSS receiver (PPM 10-xx38). **C:** During the experiment, participants requested navigation instructions using a custom-built clicker-device (circled in red) which triggers a LED light located in the backpack informing the experimenter about the request.

4 Data Analysis

4.1 Preprocessing

The primary goal of the preprocessing procedures was to extract variables describing the observations and to retrieve associated features needed in the modelling part of this work. For this, the GPS data and logged times of requests were processed. Extracting all observations of first requests of route instructions, resulted in a number of 304 data points. These were further processed to obtain observational variables for the analysis of *RQ 1* and *RQ 2*. While the aim was to predict the optimal timing of a route instruction in addressing *RQ 1*, the spatial dimension of a request was chosen to be modeled. This decision was made in order to counterbalance potential biases resulting from walking speed fluctuations due to e.g. crows, red lights, etc. Therefore, the predicted output by the survival analysis was the distance between the position of a first instruction request and the last turning point, normalized by the overall distance between the approached intersection and the last turning point, yielding values in the range $[0,1]$. Further, the observations were labeled with a binary variable $\in \{0,1\}$, 1 indicating that a route instruction was requested more than once, which served as the dependent variable to analyze *RQ 2*. Associated features were extracted from the questionnaires which provided data on personal variables, and from information on walked routes from which route-specific features were extracted (e.g., type of approached intersection). Further, characteristics of the environment were calculated using spatial data drawn from OpenStreetMap and Urban Atlas databases (EUROPEAN COMMISSION 2016).

4.2 RQ 1: Rationale to use Survival Analysis

To address *RQ 1*, a survival model was applied. Such a model essentially predicts the probability of the occurrence of an event over time. WALDORF (2003) verified the conceptual equality of applying survival models to temporal and spatial processes which is exploited in this work as the distance (instead of actual time) at which a first presentation of a route instruction is preferred is predicted.

4.3 RQ 2: Rationale to use Generalized Estimating Equations (GEE)

The relation between the binary labels indicating whether participants requested a route instruction more than once and retrieved features was modeled using the Generalized Estimating Equations (GEE) model. This model was first described by LIANG & ZEGER (1986). It was chosen for this analysis because it allows within-subject correlation which is present in the dataset at hand, solves a marginal model describing population-average effects, and is adaptable to a logistic output.

5 Results and Discussion

Two models were obtained to answer the research questions at hand: Firstly, a survival model predicting the optimal distance since the last turning point at which a user of a navigation system would want to hear a route instruction for the upcoming turning point; secondly, a GEE model expressing the population mean of pedestrians' preference of wanting to hear an instruction more than once. Both models express the observed value as a function of user and environmental

characteristics. While initially including all retrieved features, each model was carefully reduced to the influential ones. This was done in a step-wise manner by increasing information criteria such as the AIC and the QIC. Features were finally classified as influential when their p-values were below 5%. During this model building process, attention was paid to avoiding multicollinearity and model overfitting by observing VIF and Cook's Distance values. The influence of a variable in the final models is interpreted based on the magnitude and sign of obtained covariates.
Overall, influential variables in both final models encompassed features from all initially considered categories which is in line with the research suggesting an impact of both user and environmental characteristics on the wayfinder's behavior (e.g., GIANNOPOLOUS et al. (2014) and the findings of GIANNOPOLOUS et al. (2017). Interestingly, results revealed a different effect by the variable age within the model by GIANNOPOLOUS et al. (2017) than within the survival model obtained in the present study: While the model by GIANNOPOLOUS et al. (2017) suggests that people of higher age would tend to request route instructions earlier, the opposite was found in the present study. Further research is needed to understand the reasons for this contradiction. Two potential explanations come to mind: This contradiction may stem from the different experiments' set-up (virtual vs. outdoor environment) which raises questions concerning the ecological validity of a study in virtual setting. Another reason may be the presence of a bias arising from differences in participant samples. However, there are also similarities among the findings by, e.g., the effect of route-related features such as the length of the currently walked route segment. Figure 2 displays the impact on the survival rate by the environmental and route-related variables in the obtained survival model. LC_1, LC_2 and LC_3 represent area shares of different landcover classes in the area the currently walked route segment lies in. $lgnSegm$ is a binary variable; equaled 1, it implies that the currently walked route segment is greater than $120m$ long. Considering an artificial observation ("base case"), the direct impact on the survival rate by an alternation of the values of selected variables is displayed (decrease/increase by one standard deviation or value change in a binary variable). For example, by the decrease of the variable LC_1 (representing areas with *Industrial, commercial, public, military and private units,* EUROPEAN COMMISSION 2020), the survival rate decreases at earlier distances which implies that when the variable's value increases, persons would be more prone to request a route instruction earlier. This quantification of an effect on the preference of route instruction timing holds the potential of being directly applied in the adaptation of a navigation system to a user's needs in different environments.
An interesting finding within the obtained GEE model was the effect of the personal characteristic *neuroticism* which was measured using the BFI-10 scale. The model indicates that people who have high scores on the *neuroticism* scale tend to request a route instruction at the beginning of the experiment more often than later on. As a person scoring high on this scale suffers from anxiety and stress (COSTA & MCCREA 2017), it is reasonable that this person also feels stressed during the experiment. The stress might decrease in time for anxious people as they might feel more comfortable over time. However, it remains unknown whether this observed effect can be reduced to the experiment situation or if people with a strong trait of neuroticism generally feel stressed when navigating. Having said this, spatial anxiety was reported to correlate positively with making more errors during wayfinding (HUND & MARINARIK 2016). Though, it is unknown how and if neuroticism is related to spatial anxiety. Nevertheless, it can be assumed that listening to a route instruction more often reduces uncertainty and therefore also the error rate during navigation.

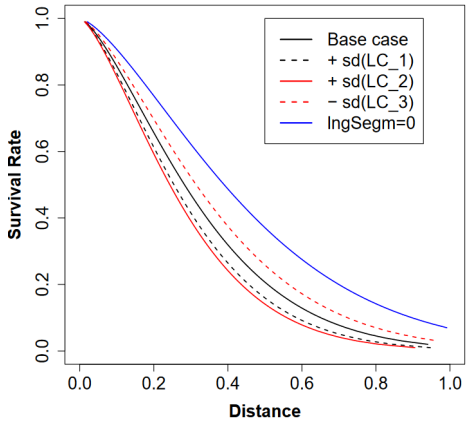


Fig. 2: Survival model: changing survival rate depending on effects given by the presence of different landcover classes (LC_1, LC_2, LC_3) and a distance between the last turning point and approached decision point which is below $120m$ ($lngSegm = 0$).

6 Conclusion and Future Work

A set of user and environmental variables influencing a pedestrian's preferences in respect to route instruction timing has been found, and the size and direction of their effect quantified. Given the observed differences to the findings by GIANNOPOLOUS et al. (2017) and the yet-to-be-conducted validation of the obtained results, further studies on this topic are needed. Besides, the contribution of this work has been evident as the conducted outdoor wayfinding study has been serving to answer different types of research questions, due to the careful experiment design, which incorporated various sensors and measurements of a multitude of behavioral variables: While the present work was dedicated to studying the timing of route instructions, ALINAGHI et al. (2021) analyzed the acquired eye-tracking data to predict a turning decision at intersections. Moreover, a further journal paper is currently being prepared which uses IMU and eye-tracking data collected during this study to predict a participant's familiarity.

7 List of References

ALINAGHI, N., KATTENBECK, M., GOLAB, A., & GIANNOPOULOS, I., 2021: Will You Take This Turn? Gaze-Based Turning Activity Recognition During Navigation. 11th International Conference on Geographic Information Science (GIScience2021)-Part II, https://doi:10.4230/LIPIcs.GIScience.2021.II.5.

COSTA, P. T. & MCCRAE, R. R., 2017: The NEO Inventories as instruments of psychological theory. The Oxford Handbook of the Five Factor Model, 11-37.

EUROPEAN COMMISSION, 2012: Mapping guide for a European Urban Atlas — copernicus. hhttps://land.copernicus.eu/user-corner/technical-library/urban-atlas-mapping-guide, last accessed 30.11.2020.

GIANNOPOULOS, I., JONIETZ, D., RAUBAL, M., SARLAS, G., & STÄHLI, L., 2017: Timing of pedestrian navigation instructions. 13th International Conference on Spatial Information Theory (COSIT), **86**, 16:1-16:13, https://doi:10.4230/LIPIcs.COSIT.2017.1.

GIANNOPOULOS, I., KIEFER, P., RAUBAL, M., RICHTER, K.-F., & THRASH, T., 2014: Wayfinding Decision Situations: A Conceptual Model and Evaluation. In Proceedings of the 8th International Conference on Geographic Information Science (GIScience2014), 221–234. https://doi:10.1007/978-3-319-11593-1_15.

GOLAB, A., KATTENBECK, M., SARLAS, G., & GIANNOPOULOS, I., 2021: It's also about timing! When do pedestrians want to receive navigation instructions. Spatial Cognition & Computation, 1-33. https://doi.org/10.1080/13875868.2021.1942474.

HUND, A. M. & MINARIK, J. L., 2006: Getting from here to there: Spatial anxiety, wayfinding strategies, direction type, and wayfinding efficiency. Spatial cognition and computation, **6**(3), 179-201, https://doi.org/10.1207/s15427633scc0603_1.

KLIPPEL, A., RICHTER, K. F., & HANSEN, S., 2009: Cognitively ergonomic route directions. Handbook of Research on Geoinformatics, 230-238. https://dio:10.4018/978-1-59140-995-3.ch029.

LIANG, K.-Y., & ZEGER, S. L. 1986: Longitudinal data analysis using generalized linear models. Biometrika, **73**(1), 13-22. https://doi:10.1093/biomet/73.1.13.

MÜNZER, S., & HÖLSCHER, C., 2011: Entwicklung und Validierung eines Fragebogens zu räumlichen Strategien. Diagnostica, **57**(3), 111-125. https://doi:10.1026/0012-1924/a000040.

RAMMSTEDT, B., KEMPER, C. J., KLEIN, M. C., BEIERLEIN, C., & KOVALEVA, A., 2012: Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit: Big-Five-Inventory-10 (BFI-10). GESIS-Working Papers, **22**(2), 1-32.

ROUSELL, A., & ZIPF, A., 2017: Towards a landmark-based pedestrian navigation service using OSM data. ISPRS International Journal of Geo-Information, **6**(3). https://doi:10.3390/ijgi6030064.

WALDORF, B., 2003: Spatial patterns and processes in a longitudinal framework. International Regional Science Review, **26**(3), 269-288. https://doi.org/10.1177/0160017603253788.

Multi-Target Multi-Camera Drone Tracking

YUCHANG JIANG[1]

Abstract: The use of drones has increased dramatically in recent years. To monitor different use cases of drones, a visual tracking system for multi-target multi-camera (MTMC) drone tracking tasks is required. Yet intensive studies have been conducted in the field of MTMC, most of them focus on pedestrian tracking or vehicle tracking, leaving drone tracking underexplored. To tackle this problem, we have proposed an approach to track multiple drones in a roughly synchronized static camera network with unknown camera poses. This setting simulates practical applications, such as monitoring drones in a public webcam network. First, the existing method developed for pedestrian tracking is evaluated on our drone data to generate single-camera tracking results. Then cross-camera matching is formulated as an optimization problem with the single-camera tracking result as input. The results of single-camera tracking show that we can apply methods developed for pedestrian tracking to drone tracking. Besides, the experiments of cross-camera association demonstrate the limits of visual features in the drone tracking scenario and prove the utility of geometry features extracted from drone trajectories.

1 Introduction

Due to the increasing popularity of drones in recent years, drone monitoring is eagerly in demand. UAV-related accidents have been reported more often recently because it is hard to observe a small drone and a missing alarm can lead to catastrophic results. Therefore, it is an essential and urgent task to develop a reliable drone tracking system. Furthermore, thanks to the widely-installed public camera networks, achieving drone surveillance with a multi-camera network is cost-effective, scalable, precise, and desired in real applications or commercial projects.

To perform drone surveillance in practical applications (e.g. drone monitoring in the airport), a high-level task, multi-drone multi-camera tracking (MDMCT), is required but needs further investigation. The common application domain of multi-target multi-camera tracking is pedestrian tracking. Its classic solution is a two-step approach: firstly tracking objects in each camera view independently and then associating objects across cameras to find their global identities. For the first step, we can reuse the single-camera tracker designed for pedestrian tracking in our drone tracking task. However, it is hard to directly utilize the commonly used features like appearance and geometry in the multi-camera association step in our drone tracking task. The primary challenges come from the small, fast-moving drones and the public camera network with unknown camera poses. The target objects, drones, are usually blurry and small in camera views, creating obstacle to gain enough information from appearance. Therefore, appearance features are of limited usefulness for drones. The second challenge is caused by unknown camera poses in camera networks. Although it is common to leverage geometric relationships to associate tracks across camera views, the use of geometric features requires camera poses, which are unknown in most practical cases (e.g. outdoor webcam data).

[1] ETH Zürich, Switzerland, E-Mail: yujiang@ethz.ch

This presented work focuses on building a tracking system for multiple drones in a camera network with overlapping views. To simulate real applications like drone surveillance in an airport, we consider the camera network consists of synchronized stationary cameras without spatial calibration. This work builds a two-step approach to track multiple drones in a synchronized camera network without knowing camera poses. As shown in Fig. 1, single-camera tracking (SCT) is performed independently on each camera view and then in the multi-camera tracking (MCT) stage, resulted single-camera tracks are fed into multi-camera matchers to figure out their global identities. In the single-camera stage, an established multi-object tracking (MOT) method from pedestrian tracking is applied to generate tracking results and a post-processing procedure is utilized to produce more robust long tracks. Then, both geometric and visual features are employed to match tracks across camera views in the multi-camera stage. In the end, this work generates tracking results for drones in a synchronized camera network.

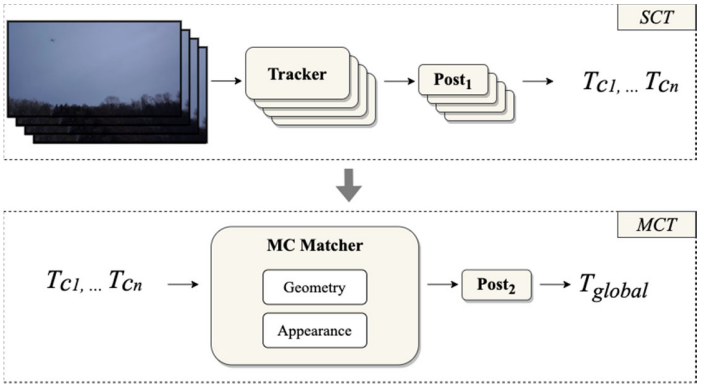


Fig. 1: Pipeline of the proposed method

2 Methodology

2.1 Data

Our dataset consists of video sequences from four synchronized static cameras. The four cameras are Sony a5100, Sony G, Samsung S10, and Sony nex5n. This challenging dataset contains three drones in a snow-covered background. Three drones are Pixhawk drone, DJI Phantom, and DJI Mavic. Annotations (ground truth tracks of drones) of acquired video sequences are labelled by human operators. For each camera, the whole video sequence is around 10 minutes. After synchronization, 16414 image frames for each camera are split into 1642 train images and 14772 test images. To observe different cases such as single-drone cases and multi-drone cases efficiently, the test set is further divided into six subsets and each subset has around 2461 frames. All train images are used to train the deep learning-based single-camera tracker, while all test images are utilized for evaluation. Several examples images are presented in Fig. 2 and Fig. 3, which shows the main difficulties of this dataset: homogeneous snow-covered background and small fast-moving objects. The homogeneous background adds the difficulty to extract distinct keypoints for camera calibration while small fast-moving objects complicates object detection and tracking.

Fig. 2: Examples of images containing large objects from four cameras

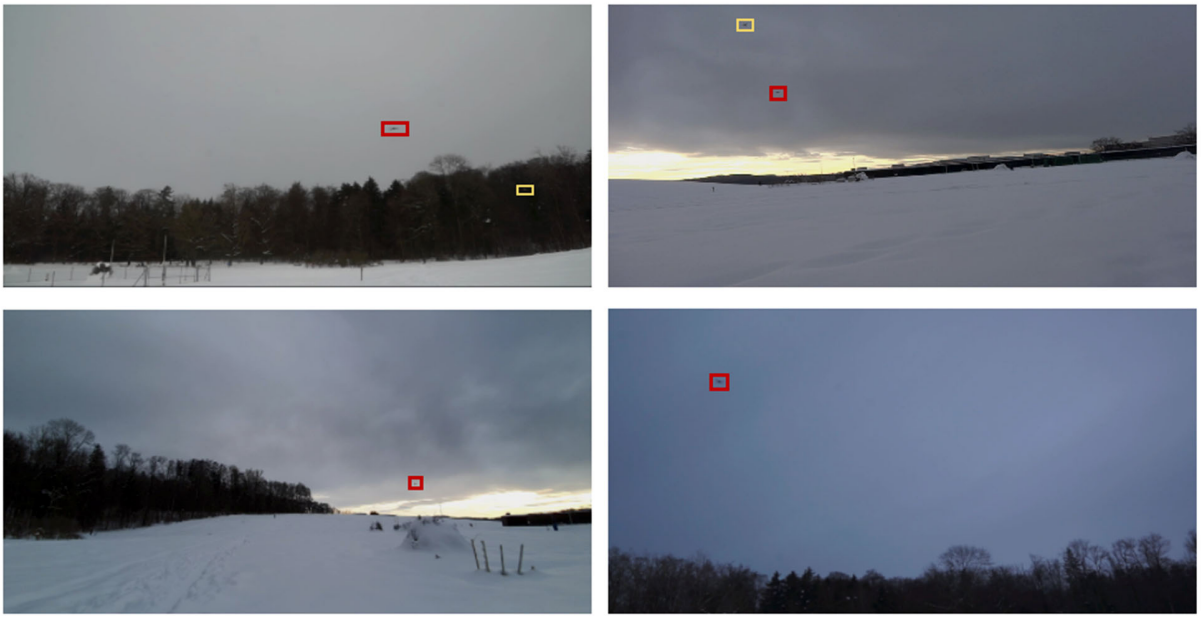


Fig. 3: Examples of images containing small objects from four cameras (different colors stand for different drones)

2.2 Generating Multi-Object Single-Camera Drone Tracks

The established tracker, Chained tracker (CTracker) (PENG et al. 2020) is utilized to track drones in each camera view. Since it raises the concept of node chaining to leverage temporal information of tracks, it is suitable to track fast-moving objects, drones. As indicated in Fig. 4, node chaining means treating a pair of consecutive frames as a node and defining a pair of successive nodes as a

node chaining. If the objects detected in the common frame (F_t) are highly overlapped based on Intersection-over-Union(IoU) score, the related detection in this node chaining (F_{t-1}, F_t, F_{t+1}) should belong to the same track.

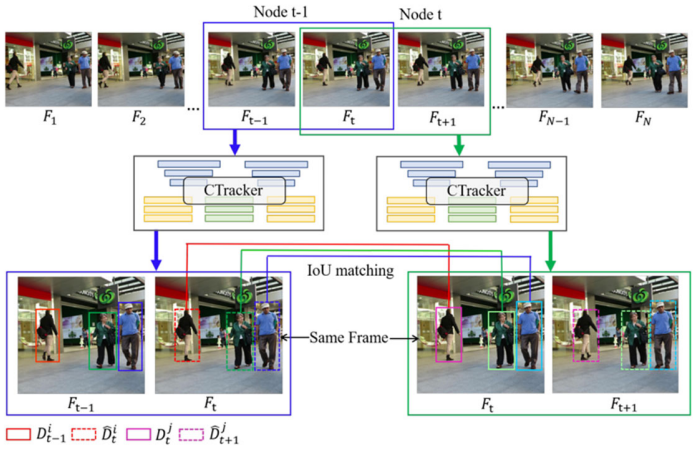


Fig. 4: Concept of node chaining (PENG et al. 2020)

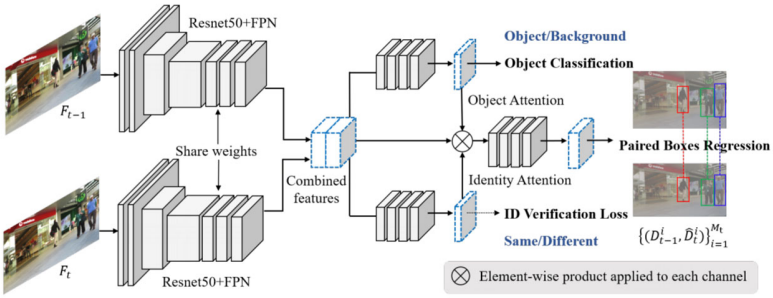


Fig. 5: Network architecture of CTracker (PENG et al. 2020)

More specifically, the detailed structure of CTracker is sketched in Fig. 5. It takes a pair of consecutive image frames as input and uses a pretrained neural network to extract features for each image. Then features from both images are concatenated and fed into two sub-models: a classification model to classify whether the detection contains the foreground and a re-identification (ReID) model to check if the detection from two frames belongs to the same object. After that, a joint attention mechanism is designed to extract useful information from two sub-models. Then a regression layer is employed to output a pair of bounding boxes (position of a bounding box in F_{t-1} and its corresponding position in F_t). During inference, the paired bounding boxes from a node chaining are used to compute the IoU score and decide whether they should be connected. The structure of CTracker and the concept of node chaining seem applicable to drone tracking tasks because the temporal information between consecutive frames is considered, which handles the fast movements of drones.

After applying the single-camera trackers, tracking results for each camera are obtained. However, these tracking results are usually a large number of short tracks, which will increase the computational burden for camera association in the MCT stage. Since connecting short tracks into

long trajectories before cross-camera association helps to reduce the computational burden, post-processing is performed to connect short tracks and remove noisy detection.

2.3 Generating Multi-Object Multi-Camera Drone Tracks

With single-camera tracking results, the multi-camera association aims to assign global identities to tracks. The MCT stage has two main components: a pairwise camera association method and a post-processing method. The pairwise camera association extracts features from a pair of cameras to match tracks from these two cameras. Then post-processing method is performed to merge pairwise matching results and resolve conflict to assign global identities to all tracks. Here we adopt a centralized way in the multi-camera tracking: one camera is defined as the reference camera and then the pairwise association is performed between the reference camera and each of the remaining cameras. Based on the matching results, global identities are assigned.

In the single-camera multi-object problem, optimization methods like Hungarian algorithm (KUHN 1955) are usually chosen to find the one object-to-one object (1V1) assignment and match objects in consecutive frames. As one object from the previous frame can match to at most one object in the current frame, this is a standard assignment problem. However, in the multi-camera scenario, associating tracks from two cameras is more complicated since a long track in camera A can be matched to multiple short tracks in camera B, making pairwise camera association an NP-hard problem. To fulfil this multiple track-to-multiple tracks (NVN) requirements approximately and simplify the optimization problem, we split the whole time domain (e.g. 2000 frames) into equal-sized sliding windows (e.g. 300 frames) and then solve the 1V1 assignment problem for each sliding window independently. The assumption is that the 1V1 relationship is approximately preserved in a short sliding window. With this assumption, we can extract features to compute the cost or affinity between tracks and then perform the classic Hungarian algorithm to solve the assignment problem in each sliding window. The pseudo-code of this algorithm is listed in Alg. 1. This algorithm takes two sets of tracks from two cameras as input and outputs the matches of tracks.

Input tracks from two cameras, T_{c1}, T_{c2}
Output matches of tracks from two cameras $M = \{(t_{c1}, t_{c2}, start, end) | t_{c1} \in T_{c1}, t_{c2} \in T_{c2}\}$

procedure MATCH PAIRWISE CAMERAS
  $M \leftarrow \emptyset$
  for each sliding window in time **do**
    $start \leftarrow start_{window}$ ▷ starting time of the current window
    $end \leftarrow end_{window}$ ▷ ending time of the current window
    $cost \leftarrow \infty$
    for each pair of t_i from $c1$, t_j from $c2$ inside current window **do**
      if t_i, t_j overlap in time **then**
        $cost(t_i, t_j) = \Psi_{affinity}(t_i, t_j)$
    Add dustbin node to $cost$
    $M' = Hungarian\ algorithm(cost)$
    $M = M \cup M'$

Alg. 1: Algorithm of pairwise camera association

In each sliding window, the track-to-track cost matrix $C \in \mathbb{R}^{n \times n}$ is formed (n is the total number of tracks in this sliding window) and each element $c_{i,j}$ contains the distance or affinity measure of two tracks, i and j. During the computation of the cost matrix, dustbin nodes are concatenated to

deal with objects that only appear in one camera view, motivated by the dustbin concept in (SARLIN et al. 2020). Then Hungarian algorithm is applied to the cost matrix to find tracks belonging to the same object.

To measure track affinity ($\Psi_{affinity}$), two types of features, appearance (Ψ_{app}) and geometry (Ψ_{geo}), are considered. The pretrained ReID component from CTacker extracts appearance features of detections in each track. The average appearance feature of each track is defined as the representative feature and the Euclidean distance between any two tracks are computed as appearance cost. Compared to visual features, geometric features are more complicated. To leverage geometric relationships, camera poses are usually required, which is unknown in our case. A natural solution is to extract keypoints and then compute fundamental matrix to obtain geometric relationships. However, in the practical scenario of drone tracking, the outdoor background of video sequences is usually sky or ground. Hence, it is hard to find robust and distinctive keypoints from the static background. Inspired by the previous work (LI et al. 2020), we consider the dynamic part, the detected drones, as keypoints. For a pair of tracks, center points of detection from temporal-overlapping frames are treated as corresponding points. Geometric relationships such as fundamental matrix and homography are obtained from these detections. There are two affinity measurements for geometric features: homography error (H) and epipolar line error (F). Homography of two tracks can be computed with the RANSAC method (FISCHLER & BOLLES 1981). Then points in image 2 can be transformed to image 1 to calculate Euclidean distance as affinity score. Similarly, the fundamental matrix of two tracks can be computed and then the epipolar line error can reflect the affinity between two tracks. To combine both features, the affinity score is defined as Eq. 1, which means the affinity score prefers appearance feature more if the object's bounding box is close to the defined size of 'large' drones, λ_{box}. This equation is based on the assumption that larger objects usually provide richer visual features while small objects can be considered as points to utilize geometry features.

$$\Psi_{affinity} = w_{app}\Psi_{app} + (1 - w_{app})\Psi_{geo}$$
$$w_{app} = min(\frac{box\ size}{\lambda_{box}}, 1) \quad (1)$$

After accomplishing pairwise camera association, a post-processing procedure is designed to combine these matching results across all cameras and resolve possible conflicts to build trajectories with global identities. In the end, each track is assigned with a new global identity and across-camera tracking is achieved.

3 Results and Discussion

For evaluation, IDF1 (identity F1 score) is adopted as the primary metric. It penalizes false negative and false positive predictions based on a bijective mapping between trajectories and ground truth tracks. In experiments, all test sets are evaluated and their average IDF1 values are summarized for both single-camera tracking and multi-camera tracking stage.

Fig. 6: Visual results of single-camera tracking (left: before post-processing, right: after post-processing)

In the single-camera tracking stage, it is noticed that CTracker tends to produce short tracks but these tracks can be connected into longer ones with a post-processing method. For example, from Fig. 6, several short tracks in the left image (before post-processing) are combined into a longer one in the right image (after post-processing), which shows the effect of post-processing and the result single-camera tracking. At the end of single-camera tracking, the IDF1 can achieve 65.27%.

With single-camera tracking results, different features are adopted to realize multi-camera association. In Tab. 1, results of using appearance feature only, geometry feature only, and the combination of both features are summarized. It proves the effectiveness of geometry features in the drone tracking application.

Tab. 1: Result of multi-camera trackers with different affinity measures

	Appearance	**Geometry (H)**	**Geometry (F)**	**App + Geo (H)**	**App + Geo (F)**
IDF1	40.55 %	43.29 %	43.50 %	50.04 %	43.45 %

4 Conclusion

The proposed pipeline can partially solve the multi-drone multi-camera tracking problem. It can generate single-camera tracks with the existing algorithm and then associate tracks with appearance and geometric features. The contributions of this work include: evaluate the state-of-the-art tracker on a drone dataset, explore possible features in multi-camera drone tracking scenarios, and leverage geometric features in a camera network without camera calibration. However, the proposed solution in the multi-camera association step approximates the 1-to-N association problem with temporal windows. Still, the information in all frames from all cameras remained underexploited, suggesting the possibilities to reformulate the problem exactly in order to fully use the information. Besides, more training data can improve the single-camera tracker's performance and the results of the multi-camera association.

5 Bibliography

FISCHLER, M.A. & BOLLES, R.C., 1981: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, **24**(6), 381-395.

LI, J., MURRAY, J., ISMAILI, D., SCHINDLER, K.& ALBL, C., 2020: Reconstruction of 3D flight trajectories from ad-hoc camera networks. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1621-1628.

KUHN, H.W., 1955: The Hungarian method for the assignment problem. Naval research logistics quarterly, **2**(1-2), 83-97.

PENG, J., WANG, C., WAN, F., WU, Y., WANG, Y., TAI, Y., WANG, C., LI, J., HUANG, F. & FU, Y., 2020: Chained-tracker: Chaining paired attentive regression results for end-to-end joint multiple-object detection and tracking. European Conference on Computer Vision, Springer, Cham, 145-161.

SARLIN, P.E., DETONE, D., MALISIEWICZ, T. & RABINOVICH, A., 2020: Superglue: Learning feature matching with graph neural networks. IEEE/CVF conference on computer vision and pattern recognition, 4938-4947.

Combining Multisensor Satellite Remote Sensing Time Series to Assess Disturbances in Central European Temperate Forests

SIMON KÖNIG[1,2,3], FRANK THONFELD[4,5], OLENA DUBOVYK[1,6] & MARCO HEURICH[2,3,7]

Abstract: Bark Beetle outbreaks are an important disturbance agent in Central European forests, with strong impacts in recent years. Here, dense time series of Landsat, Sentinel-2, and Sentinel-1, plus their combinations, were used to detect bark beetle infestations in the Bavarian Forest National Park, Germany. Various processing steps were applied to all data, and a consistent data cube was built, of which a variety of parameters were computed. The parameters with the best ability to separate infested from healthy plots were selected for the detection. From these, infestation probability time series were derived and classified into infested/healthy according to increasing probability thresholds, and the respective detections were evaluated. The results show that Sentinel-2 works best, both in terms of spatial accuracy and an early detection. Combining Sentinel-2 and Landsat didn't provide benefits, and Sentinel-2 SAR data or the fusion of SAR and multispectral data performed worse. The results of this study have important implications for forest monitoring and management.

1 Introduction

Temperate forests are a major part of the biosphere, hosts of biodiversity, and an important component of land-atmosphere interactions. If healthy and intact, they serve as net carbon sinks, and thus are a key asset in the mitigation of climate change. Besides, they provide a variety of ecosystem services to humankind, including the provision of goods such as timber, regulating effects like the protection against erosion, or cultural services such as the potential for recreation (BROCKERHOFF et al. 2017). Natural disturbances are crucial for the dynamics of temperate forest ecosystems. This includes both abiotic (e.g. windthrow, drought) and biotic disturbances (e.g insect outbreaks, pathogens). By changing the structure, composition, and function of forests, they facilitate succession, renewal, and reorganization. Thus, they contribute to forest heterogeneity, biodiversity, and resilience (SEIDL et al. 2017). In recent decades, disturbance regimes have changed around the globe, whereby many of these changes can be attributed to climate change.

[1] Geographisches Institut, Universität Bonn, Meckenheimer Allee 156, 53115 Bonn
[2] Nationalpark Bayerischer Wald, Freyunger Straße 2, 94481 Grafenau, E-Mail: [simon.koenig, marco.heurich]@npv-bw.bayern.de
[3] Fakultät für Umwelt und Natürliche Ressourcen, Universität Freiburg, Tennenbacher Straße 4, 79106 Freiburg
[4] Deutsches Fernerkundungsdatenzentrum, Deutsches Zentrum für Luft- und Raumfahrt, Münchener Straße 20, 82234 Weßling, E-Mail: frank.thonfeld@dlr.de
[5] Institut für Geographie und Geologie, Universität Würzburg, Am Hubland, 97074 Würzburg
[6] Zentrum für Fernerkundung der Landoberfläche, Universität Bonn, Genscherallee 3, 53113 Bonn, E-Mail: odubovyk@uni-bonn.de
[7] Department of Forestry and Wildlife Management, Inland Norway University of Applied Sciences, Anne Evenstads Veg 80, 2480 Koppang, Norwegen

Likewise, further shifts in forest disturbance regimes are to be expected under further warming climate, potentially resulting in so-called "megadisturbances" which may alter temperate forest ecosystems beyond their resilience (MILLAR & STEPHENSON 2015; SEIDL et al. 2017).
Outbreaks of bark beetles are one of the most important disturbance agents in temperate forests. These insects reproduce under the bark of their host trees. By feeding on the tree's phloem, i.e. the living tissue under the bark, many bark beetle species disrupt water and nutrient transport and consequently kill their host tree (HLÁSNY et al. 2021). In Central Europe, the bark beetle species whose outbreaks have historically caused the most widespread consequences is the European Spruce Bark Beetle (*Ips Typographus* L.). In recent years, facilitated by a series of hot and dry summers and past forest practice resulting in widespread old-growth Norway Spruce (*Picea abies* L.), i.e. the main host tree of the European Spruce Bark Beetle, Central Europe has experienced bark beetle-induced tree mortality levels beyond any historical level. For example, 16 % of all spruce growing stock have been subject to die-off between 2018 in 2020 in Germany alone (BUNDESMINISTERIUM FÜR ERNÄHRUNG UND LANDWIRTSCHAFT 2021).
Satellite remote sensing has been proven to be a highly valuable tool in the monitoring of forest disturbances, very much including bark beetle infestations (SENF et al. 2017). Bark beetle disturbances happen in three distinct stages: during the *green attack* phase, the tree foliage does not visibly change. This changes in the second, i.e. *red attack* phase: as the die-off of the tree proceeds, its needles first turn yellow and then proceed to a red color. As time progresses, the tree starts to lose its needles and the tree bark slowly starts to detach. This constitutes the *grey attack* phase. The detection of dead trees in the grey attack phase is relatively easy using satellite imagery (WULDER et al. 2006). However, the early detection of infested trees at the beginning of the red attack or even in the green attack phase is more challenging while at the same time very important, because in managed forests, infested trees should be logged and removed quickly to prevent further spread. Bark beetle infestations induce physiological changes in a tree, including loss of water and chlorophyll content in the needles, but these changes induce only subtle shifts in the reflectance, at least in the beginning of an infestation (ABDULLAH et al. 2018). A variety of approaches exist for the detection of bark beetle infestations. Many of these relied on multispectral data, e.g. from Landsat and Sentinel-2 (ABDULLAH et al. 2019; BÁRTA et al. 2021; HAIS & KUČERA 2008; MEIGS et al. 2011). SAR imagery from sensors such as Sentinel-1 or TerraSAR-X has only seldomly been applied (HOLLAUS & VREUGDENHIL 2019). In addition, approaches using high time series density have not frequently been applied (BÁRTA et al. 2021; HUO et al. 2021), even though a high observation density is beneficial to ensure an early detection of bark beetle infestations (SENF et al. 2017). The same holds true for the combination of different kinds of data to increase the density of applied time series, which is especially important for forestry applications in cloud-poor regions (REICHE et al. 2018). This study tries to close these gaps by using full time series of multiple satellites. Its goal is to assess how well data of different sensors, as well as their combinations, can be used to detect bark beetle infestations. This refers to both the spatial accuracy of detection as well as the date on which a detection happens. Assessing the capabilities of dense time series as well as such derived from multiple satellites is crucial for a potential future development of large-scale near-realtime bark beetle infestation monitoring systems.

2 Data and Methods

2.1 Study Area

The Bavarian Forest National Park in Southeastern Germany served as a study area for the approach. Founded in 1970, Germany's oldest National Park comprises an area of ca. 250 km^2. The park is located at altitudes between 600 and 1450 m and largely dominated by coniferous forests (HEURICH et al. 2010). Mostly characterized by old-groth Norway, the park has strongly been affected by disturbances of the European Spruce Bark Beetle. In the park's core zone, infested trees are not cleared, which enables the longterm monitoring of bark beetle outbreaks (LATIFI et al. 2021). Yearly infested areas in the park have been documented in a spatially explicit manner based on aerial imagery since 1988 (HEURICH et al. 2010). This is a unique dataset that served as a reference dataset for this study.

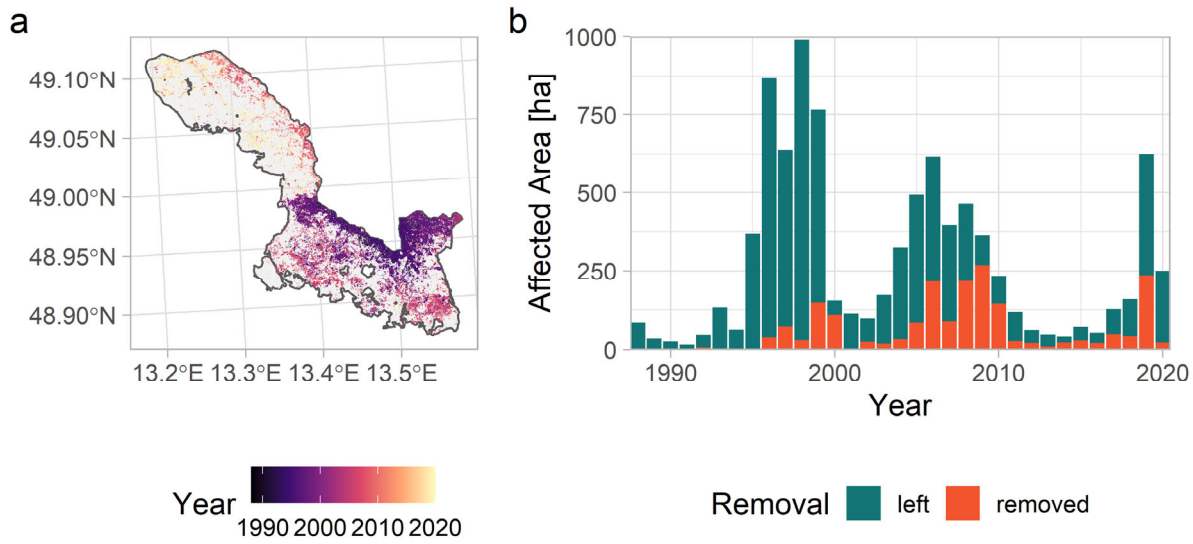


Fig. 1: Overview of bark beetle disturbances in the Bavarian Forest National Park. a) shows a map of all disturbances in the National Park since the 1980s. b) is a summary of the total area infested per year

2.2 Satellite Data

The following satellite data was used: the Landsat satellite family, providing data since 1982 with a spatial resolution of 30 m, as well as Sentinel-2, providing data since 2015 with a spatial resolution of 10 to 20 m, were used as sources of multispectral imagery, whereby Landsat has a repeat interval of 16 days, and Sentinel-2 offers a three-day revisit. Sentinel-1 was included as well, delivering SAR data with a spatial resolution of approximately 10 m and a revisit of six days. In total, 3177 multispectral and 1081 SAR images were used.

2.3 Methods

2.3.1 Data Processing

Since one major goal of this study was to compare and combine multiple kinds of satellite data, a first important step was the generation of analysis-ready data (ARD) to ensure that all data can be used in conjunction. Therefore, extensive processing steps were applied to both the multispectral and the SAR data. Landsat and Sentinel-2, besides their differences regarding spatial and temporal resolution, are spectrally similar and can hence be regarded as a "virtual constellation" (WULDER et al. 2015). If appropriate processing steps are applied, data from both satellites can be used in conjunction. Here, the *Framework for Operational Radiometric Correction for Environmental monitoring* (FORCE) was used to ensure this appropriate processing (FRANTZ 2019). FORCE applies a consistent processing scheme to both Landsat and Sentinel-2 (FRANTZ et al. 2016). This includes the radiometric correction and harmonization of the data as well as the generation of quality flags to remove pixels covered by clouds, cloud shadows, snow, etc. Thereby, the software uses a modified version of the popular Fmask algorithm (ZHU & WOODCOCK 2012) that considers parallax effects in the Sentinel-2 imagery to overcome the lack of a thermal band of this satellite (FRANTZ et al. 2018). Using the same software, the spatial characteristics of both types of data were aligned, i.e. all data were projected into one common coordinate system, resampled to the same resolution of 10 m, and subdivided into tiles of 30 × 30 km. As the study area fits in one of these tiles, all other tiles were disregarded to reduce data volume.

Likewise, Sentinel-1 was processed extensively as well data as well, whereby the majority of processing was performed on the Google Earth Engine cloud processing platform (GORELICK et al. 2017). The Sentinel-1 GRD imagery available on this platform is already preprocessed, including thermal noise removal, radiometric calibration, and terrain correction. This data was further processed using one consistent workflow (MULLISSA et al. 2021). Specifically, a border noise correction was applied, as well as a multitemporal Quegan speckle filter. To account for terrain and use both ascending and descending Sentinel-1 orbits in conjunction, radiometric terrain flattening was applied as well (VOLLRATH et al. 2020). The processed data was exported from Google Earth Engine, and using the FORCE software again, its spatial characteristics were aligned with Landsat and Sentinel-2. Thereby one consistent data cube of all available Landsat, Sentinel-2, and Sentinel-1 images, with a spatial resolution of 10 m, was created. A total of 25 parameters were computed from this data, including 10 "traditional" vegetation indices, 11 indices based on Sentinel 2's red edge channels, and 4 SAR parameters from Sentinel-1. An overview of the parameters is listed in Table 1. To remove seasonal fluctuations and noise from the data, every time series of every index in every pixel were smoothed using locally estimated Scatterplot Smoothing (LOESS) models. LOESS is an implementation of local regression (HASTIE et al. 2009).

2.3.2 Sampling and Class Separability Assessment

The fitting of the detection algorithm implemented in this study requires data representing both healthy and disturbed forests. This data was gathered with the reference dataset mentioned above. Yearly median composites for the years 2014 to 2018 were computed for all 25 parameters, and these were sampled according to the following strategy: to gain data representing disturbed forests,

these composites were sampled with the reference data of the previous year. Vice-versa, composites were sampled with the reference data of two years afterwards to gain data representing healthy forests. E.g. the median composite of 2017 was sampled with areas that were subject to bark beetle infestations in the year before to ensure that the sampled values definitely represent infested forests. The same composite was also sampled with areas that were infested in 2019 based on the assumption that forests infested in this year had still been healthy in 2017. To assess which of the 25 parameters separate healthy and infested forests best, the Jeffries-Matusita (JM) distance was computed for any of them. This measure has a finite range of 0 to 2, with 0 indicating complete inseparability and 2 indicating full separability (LALIBERTE et al. 2012). The multispectral parameter $v1$ and SAR parameter $v2$ with the highest separability, respectively, were selected for further analysis.

2.3.3 Disturbance Detection

Disturbances were detected for all pixels that were subject to bark beetle infestations in 2019 as well as an equal number of pixels that were known to be still healthy during this year. To detect disturbances, a Bayesian approach was applied (REICHE et al. 2018, 2015). From the distributions of the sampled values for the two best-performing variables, Gaussian probability density functions were computed, and the probability densities $p(v1|H)$ and $p(v2|H)$ that a pixel belongs to a healthy forest as well as densities $p(v1|I)$ and $p(v2|I)$ that it belongs to an infested forest were computed. From these, the conditional probability of an infestation could be derived using Bayes' theorem (MCELREATH 2020), assuming equal priors for healthy and infested:

$$P(I|v1_t) = \frac{p(v1_t|I)}{p(v1_t|I) + p(v1_t|H)} for\ t \in \mathrm{T}_{v1} \quad (1)$$

where t is the date in the time series of a given satellite observation (REICHE et al. 2015). Analogously, the parameter time series of both selected variables for every pixel were transformed into infestation probability time series. This procedure is illustrated for one example pixel in Fig. 2. The detection itself was performed by binarily classifying the time series into healthy and infested using 100 increasing probability thresholds between 0.01 and 1. Thereby, the detection capabilities of five different sensor configurations were examined:

- Multispectral, i.e. all observations of $P(I|v1_t)$.
- Landsat, i.e. $P(I|v1_t)$ with a subset of t corresponding to the Landsat observation dates.
- Sentinel-2, i.e. $P(I|v1_t)$ with a subset of t representing the Sentinel-2 observation dates.
- Sentinel-1, i.e. all observations of $P(I|v2_t)$.
- Fused, i.e. $P(I|v1_t)$ and $P(I|v2_t)$ combined, arranged by t.

By comparison to areas in the national park that were still healthy and pixels that were infested in 2019, the spatial accuracy of detections was assessed. If a pixel was detected as infested under a given threshold, the date of the satellite observation during which the detection occurred was recorded. Infestations that were recorded before April 1st, 2019 (the approximate beginning of the bark beetle season) were regarded as false detections and consequently not counted. The same holds true for observations detected later than September 30th, 2020 (one year after the approximate end of the bark beetle season in 2019). Besides, to increase confidence in the results, only

detections were counted that were confirmed by two additional observations that were higher than the applied threshold. The overall spatial accuracies, omission and commission error of the infested class, and summary statistics on the detection date were computed.

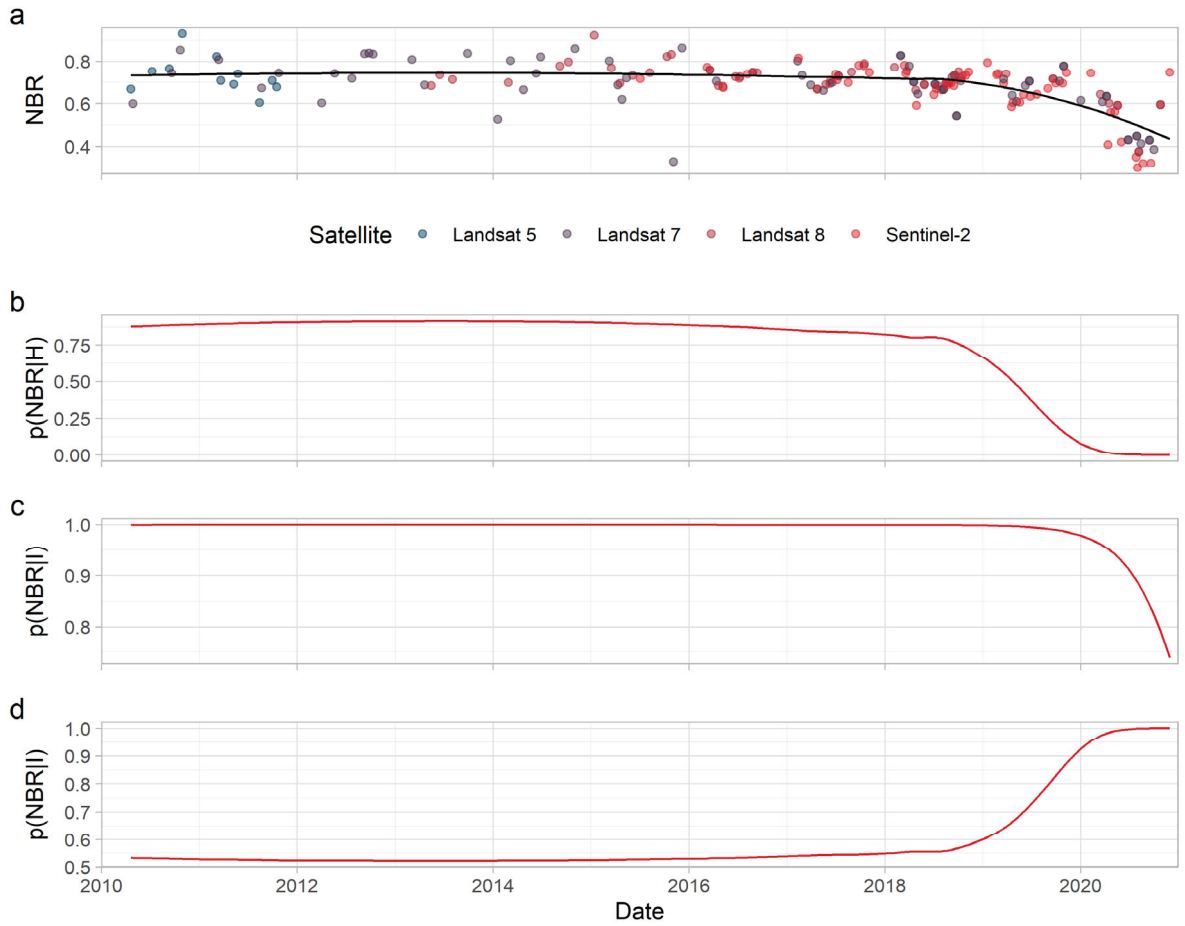


Fig. 2: Illustration of the derivation of infestation probabilities for one random example pixel infested in 2019, for the Multispectral sensor configuration. a) the original NBR time series. Colors indicate the satellite of the respective observation, the black line is the LOESS-smoothed time series. b) the derived healthy probabilities, c) the derived infested probabilities, d) conditional probabilities of an infestation

3 Results and Discussion

Fig. 3 depicts the results of the class separability according to the JM distance. The Normalized Difference Burn Ratio (NBR) was the parameter with the highest separability, scoring a Jeffries-Matusita Distance of 1.6 on a scale of 0 to 2. In general, the NBR and other vegetation indices that incorporate the Landsat/Sentinel-2 Shortwave Infrared (SWIR) bands, i.e. the Normalized Difference Moisture Index (NDM) and Normalized Difference Tillage Index, achieved the best results, as did some Sentinel-2 red-edge indices. This is in line with existing literature documenting the sensitivity of SWIR bands of Landsat and Sentinel-2, as well as the red edge region of the electromagnetic spectrum, to physiological changes in trees induced by bark beetle disturbances

(ABDULLAH et al. 2018; GOODWIN et al. 2008; MEIGS et al. 2011). Other parameters performed considerably worse, especially the SAR-based metrics, all of which achieved JM distance values below 0.5. Thereby, the VV band offered the relative highest separability (Jeffries-Matusita Distance of 0.3) of all four SAR parameters. According to the methodology described above, the NBR and the VV backscatter were selected as $v1$ and $v2$ for the detection.

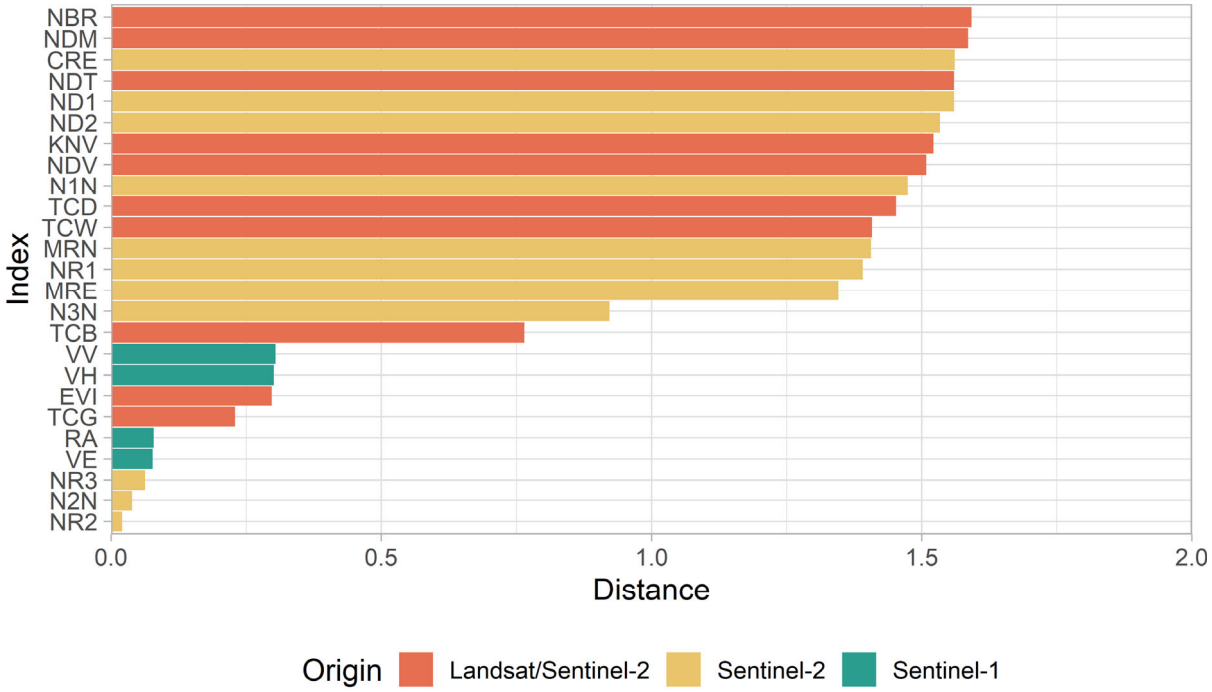


Fig. 3: Jeffries-Matusita Distance for all 25 computed parameters. Colors represent the origin of the parameter

Fig. 4 shows the overall spatial accuracies of the five different sensor configurations, depending on the applied thresholds, together with the corresponding omission and commission errors. Here, Sentinel-2 achieved the best results, with a maximum overall accuracy of 0.86 at a probability threshold of 0.9. Landsat and the combination of Landsat and Sentinel-2 achieved somewhat inferior results, whereby both sensor configurations are highly similar. The commission errors for Sentinel-2 are slightly higher, but this is more than compensated by the lower omission errors, making Sentinel-2 the best-performing sensor configuration. Sentinel-1 and the combination of all sensors by comparison perform much worse. While the maximum overall accuracies of the all-sensor configuration are still relatively good for some of the thresholds, Sentinel-1 only barely reaches a maximum overall accuracy of 0.55, and for lower probability thresholds, the overall accuracy is 0.5 (i.e. as good as a random classifier). This is the case because – while commission errors are comparably low, the omission errors are very high, i.e. 100 % for thresholds < 0.25. This means that in general, very few disturbances are detected at all by both sensor configurations. Since the better-performing multispectral data is included as well, the fusion configuration based on all datasets exhibits somewhat lower omission errors at least for thresholds > 0.5

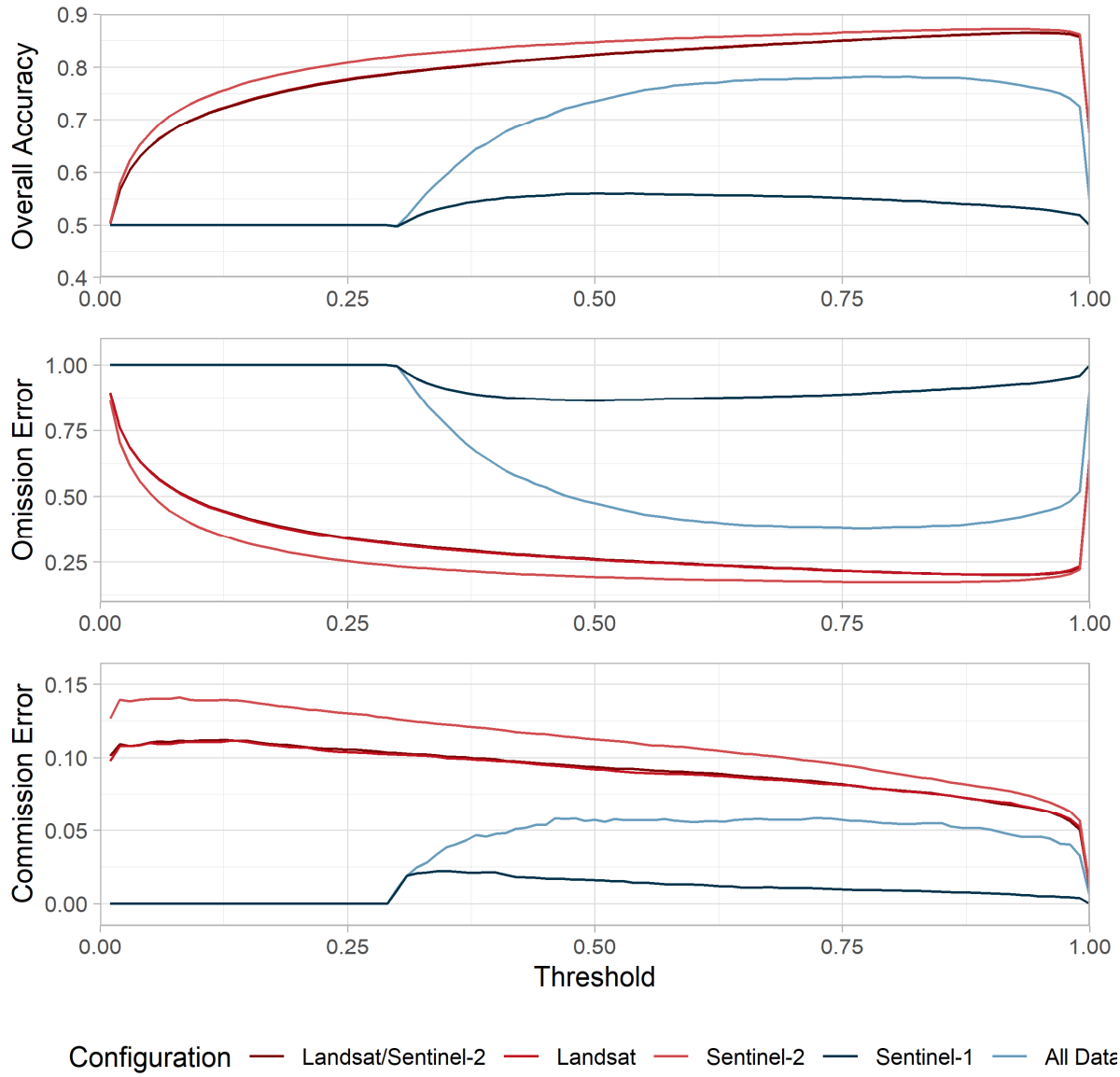


Fig. 4: Overall accuracies (top) as well as omission (center) and commission errors (bottom) for the infested class, according to the applied probability threshold. Colors represent the five sensor configurations

If the date of detection is compared between Landsat, Sentinel-2, and their combination for the respective threshold of highest spatial accuracy (Fig. 5), Sentinel-2 is at an advantage here as well. The median detection date of Sentinel-2 is nearly 1 month earlier compared to Landsat and the combination of both, whereby the two latter sensor configurations achieve very similar results once again. A detailed assessment of the time it takes from an infestation in the field to be detected by the applied satellite configurations was not possible, because the reference dataset only documents the year of an infestation, but not the infestation date at a sub-annual level. Yet, the fact that Sentinel-2 on average detects the same infestations one month earlier than Landsat or the combination of both leads to the conclusion that Sentinel-2 is indeed better equipped for an early

detection. The approach developed here allows for the choice of a threshold that suits user needs: if a lower threshold is applied, earlier detections are possible, at the price of higher errors, whereby the commission errors are generally very low and with Sentinel-2, omission errors are low as well even with relatively low thresholds (e.g. 0.25). If very high spatial accuracies (> 0.8) are the goal, e.g. for detailed damage inventories, higher probability thresholds can be applied, even though this leads to disturbances being detected later.

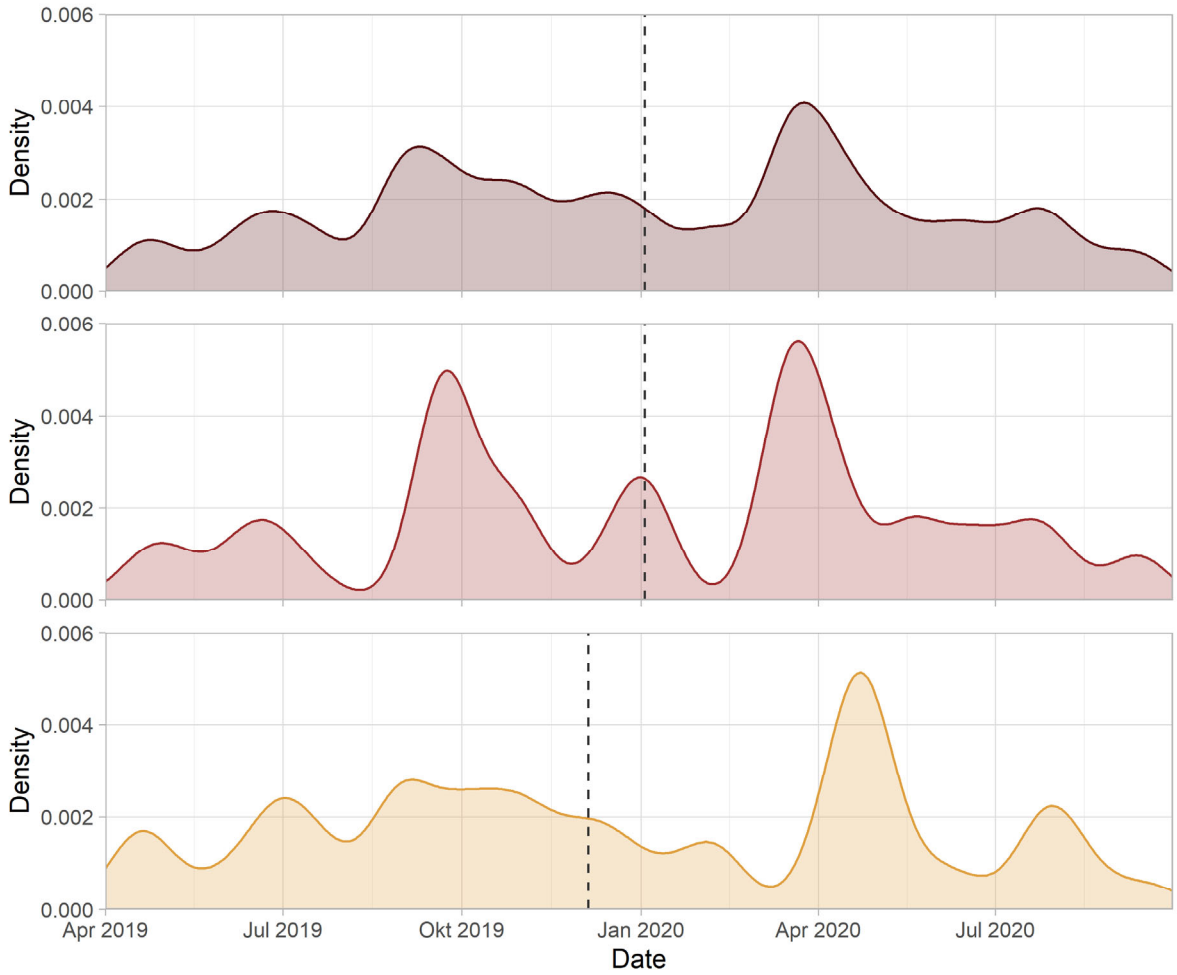


Fig. 5: Detection Dates of all pixels infested in 2017 according to the reference data. Top: Landsat and Sentinel-2 combined, center: Landsat only, bottom: Sentinel-2 only. Dashed lines represent mean detection dates

The overall good performance of Sentinel-2 can be attributed to its relatively high spatial, temporal, and spectral resolution, confirming the high potential that this satellite has for forestry applications. With similar spectral characteristics, Landsat achieves somewhat inferior results probably due to its lower spatial and temporal resolution, which also impacted the combination of both. The worse performance of the two configurations that included SAR data are likely a result of the suboptimal suitability of Sentinel-1 due to its C-band radiation not being able to penetrate forest canopy, as well as topographic and speckle effects that were not fully removed. This

however is a result of the approach applied here, relying on the automated processing of a large number of images. Possibly, more elaborate processing of fewer Sentinel-1 images and the tryout of different speckle filters may improve its results in the future.

4 Conclusion

This study utilized full time series of Landsat, Sentinel-2, and Sentinel-1 as well as a combination of multispectral and SAR data to detect forest disturbances induced by bark beetle infestations in the Bavarian Forest National Park, Germany. As elaborated in the results, all multispectral sensor configurations produced sensible results, with Sentinel-2 as the best-equipped satellite for the detection of bark beetle disturbances in terms of both time and space. The combination of Landsat and Sentinel-2 provided good results but is not necessary as it yields no advantage over applying Sentinel-2 alone. By comparison, using SAR data did not provide meaningful results, and neither did the combination of SAR and multispectral data. With the approach developed here, different probability thresholds can be applied for the detection, allowing potential users of an operational monitoring based on this approach to decide whether the goal is to detect infested trees very early (i.e. in Near-Realtime) or if a highly accurate inventory of disturbances is the goal. This is a key advantage of a potential monitoring tool based on this study. To conclude, the results of this study gave important insights about the usage of satellite data in the detection of bark beetle infestations. This is an important step in a better monitoring of these disturbances in Central Europe.

5 References

ABDULLAH, H., DARVISHZADEH, R., SKIDMORE, A.K., GROEN, T.A. & HEURICH, M., 2018: European spruce bark beetle (Ips typographus, L.) green attack affects foliar reflectance and biochemical properties. International Journal of Applied Earth Observation and Geoinformation, **64**, 199-209. https://doi.org/10.1016/j.jag.2017.09.009.

ABDULLAH, H., SKIDMORE, A.K., DARVISHZADEH, R. & HEURICH, M., 2019: Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (Ips typographus, L.) compared with Landsat-8. Remote Sensing in Ecology and Conservation, **5**, 87-106. https://doi.org/10.1002/rse2.93.

BÁRTA, V., LUKEŠ, P. & HOMOLOVÁ, L., 2021: Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. International Journal of Applied Earth Observation and Geoinformation, **100**, 102335. https://doi.org/10.1016/j.jag.2021.102335.

BROCKERHOFF, E.G., BARBARO, L., CASTAGNEYROL, B., FORRESTER, D.I., GARDINER, B., GONZÁLEZ-OLABARRIA, J.R., LYVER, P.O., MEURISSE, N., OXBROUGH, A., TAKI, H., THOMPSON, I.D., VAN DER PLAS, F. & JACTEL, H., 2017: Forest biodiversity, ecosystem functioning and the provision of ecosystem services. Biodiversity and Conservation, **26**, 3005-3035. https://doi.org/10.1007/s10531-017-1453-2.

BUNDESMINISTERIUM FÜR ERNÄHRUNG UND LANDWIRTSCHAFT, 2021: Waldbericht der Bundesregierung 2021. Bonn.

FRANTZ, D., 2019: FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. Remote Sensing, **11**, 1124. https://doi.org/10.3390/rs11091124.

FRANTZ, D., HAß, E., UHL, A., STOFFELS, J. & HILL, J., 2018: Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sensing of Environment, **215**, 471-481. https://doi.org/10.1016/j.rse.2018.04.046.

FRANTZ, D., RÖDER, A., STELLMES, M. & HILL, J., 2016: An Operational Radiometric Landsat Preprocessing Framework for Large-Area Time Series Applications. IEEE Transactions on Geoscience and Remote Sensing, **54**, 3928-3943. https://doi.org/10.1109/TGRS.2016.2530856.

GOODWIN, N.R., COOPS, N.C., WULDER, M.A., GILLANDERS, S., SCHROEDER, T.A. & NELSON, T., 2008: Estimation of insect infestation dynamics using a temporal sequence of Landsat data. Remote Sensing of Environment, **112**, 3680-3689. https://doi.org/10.1016/j.rse.2008.05.005.

GORELICK, N., HANCHER, M., DIXON, M., ILYUSHCHENKO, S., THAU, D. & MOORE, R., 2017: Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, **202**, 18-27. https://doi.org/10.1016/j.rse.2017.06.031.

HAIS, M. & KUČERA, T., 2008: Surface temperature change of spruce forest as a result of bark beetle attack: remote sensing and GIS approach. European Journal of Forest Research, **127**, 327-336. https://doi.org/10.1007/s10342-008-0208-8.

HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J.H., 2009: The elements of statistical learning: data mining, inference, and prediction, 2nd ed. Springer, New York.

HEURICH, M., BEUDERT, B., RALL, H. & KŘENOVÁ, Z., 2010: National Parks as Model Regions for Interdisciplinary Long-Term Ecological Research: The Bavarian Forest and Šumavá National Parks Underway to Transboundary Ecosystem Research, Long-Term Ecological Research: Between Theory and Application. Müller, F., Baessler, C., Schubert, H., Klotz, S. (Eds.), Springer Netherlands, Dordrecht, 327-344. https://doi.org/10.1007/978-90-481-8782-9_23.

HLÁSNY, T., KÖNIG, L., KROKENE, P., LINDNER, M., MONTAGNÉ-HUCK, C., MÜLLER, J., QIN, H., RAFFA, K.F., SCHELHAAS, M.-J., SVOBODA, M., VIIRI, H. & SEIDL, R., 2021: Bark Beetle Outbreaks in Europe: State of Knowledge and Ways Forward for Management. Current Forestry Reports, **7**, 138-165. https://doi.org/10.1007/s40725-021-00142-x.

HOLLAUS, M. & VREUGDENHIL, M., 2019: Radar Satellite Imagery for Detecting Bark Beetle Outbreaks in Forests. Current Forestry Reports, **5**, 240-250. https://doi.org/10.1007/s40725-019-00098-z.

HUO, L., PERSSON, H.J. & LINDBERG, E., 2021: Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). Remote Sensing of Environment, **255**, 112240. https://doi.org/10.1016/j.rse.2020.112240

LALIBERTE, A.S., BROWNING, D.M. & RANGO, A., 2012: A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. International Journal of Applied Earth Observation and Geoinformation, **15**, 70-78. https://doi.org/10.1016/j.jag.2011.05.011.

LATIFI, H., HOLZWARTH, S., SKIDMORE, A., BRŮNA, J., ČERVENKA, J., DARVISHZADEH, R., HAIS, M., HEIDEN, U., HOMOLOVÁ, L., KRZYSTEK, P., SCHNEIDER, T., STARÝ, M., WANG, T., MÜLLER, J. & HEURICH, M., 2021: A laboratory for conceiving Essential Biodiversity

Variables (EBVs)—The 'Data pool initiative for the Bohemian Forest Ecosystem.' Methods in Ecology and Evolution, **12**, 2073-2083. https://doi.org/10.1111/2041-210X.13695.

MCELREATH, R., 2020: Statistical rethinking: A Bayesian course with examples in R and Stan, 2nd ed. Taylor and Francis, CRC Press, Boca Raton.

MEIGS, G.W., KENNEDY, R.E. & COHEN, W.B., 2011: A Landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests. Remote Sensing of Environment, **115**, 3707–3718. https://doi.org/10.1016/j.rse.2011.09.009.

MILLAR, C.I. & STEPHENSON, N.L., 2015: Temperate forest health in an era of emerging megadisturbance. Science, **349**, 823-826. https://doi.org/10.1126/science.aaa9933.

MULLISSA, A., VOLLRATH, A., ODONGO-BRAUN, C., SLAGTER, B., BALLING, J., GOU, Y., GORELICK, N. & REICHE, J., 2021: Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. Remote Sensing, **13**, 1954. https://doi.org/10.3390/rs13101954.

REICHE, J., DE BRUIN, S., HOEKMAN, D., VERBESSELT, J. & HEROLD, M., 2015: A Bayesian Approach to Combine Landsat and ALOS PALSAR Time Series for Near Real-Time Deforestation Detection. Remote Sensing, **7**, 4973-4996. https://doi.org/10.3390/rs70504973.

REICHE, J., HAMUNYELA, E., VERBESSELT, J., HOEKMAN, D. & HEROLD, M., 2018: Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. Remote Sensing of Environment, **204**, 147-161. https://doi.org/10.1016/j.rse.2017.10.034.

SEIDL, R., THOM, D., KAUTZ, M., MARTIN-BENITO, D., PELTONIEMI, M., VACCHIANO, G., WILD, J., ASCOLI, D., PETR, M., HONKANIEMI, J., LEXER, M.J., TROTSIUK, V., MAIROTA, P., SVOBODA, M., FABRIKA, M., NAGEL, T.A. & REYER, C.P.O., 2017: Forest disturbances under climate change. Nature Climate Change, **7**, 395-402. https://doi.org/10.1038/nclimate3303.

SENF, C., SEIDL, R. & HOSTERT, P., 2017: Remote sensing of forest insect disturbances: Current state and future directions. International Journal of Applied Earth Observation and Geoinformation **60**, 49-60. https://doi.org/10.1016/j.jag.2017.04.004.

VOLLRATH, A., MULLISSA, A. & REICHE, J., 2020: Angular-Based Radiometric Slope Correction for Sentinel-1 on Google Earth Engine. Remote Sensing, **12**, 1867. https://doi.org/10.3390/rs12111867.

WULDER, M.A., DYMOND, C.C., WHITE, J.C., LECKIE, D.G. & CARROLL, A.L., 2006: Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. Forest Ecology and Management, **221**, 27-41. https://doi.org/10.1016/j.foreco.2005.09.021.

WULDER, M.A., HILKER, T., WHITE, J.C., COOPS, N.C., MASEK, J.G., PFLUGMACHER, D. & CREVIER, Y., 2015: Virtual constellations for global terrestrial monitoring. Remote Sensing of Environment, **170**, 62-76. https://doi.org/10.1016/j.rse.2015.09.001.

ZHU, Z. & WOODCOCK, C.E., 2012: Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, **118**, 83-94. https://doi.org/10.1016/j.rse.2011.10.028.

Early Detection of Bark Beetle Induced Forest Stress using Sentinel-2 Data

LISA MANDL[1,2] & STEFAN LANG[2]

Abstract: Abiotic as well as biotic disturbances are central shapers of forest ecosystems. In contrast to sudden disturbance agents such as wind, avalanches and fire, bark beetle infestation is characterized by successive progression. When infestation is observable by the human eye, trees are already in the final stages of infestation - the so-called red- and grey-attack. In the relevant phase - the green-attack - biochemical and biophysical processes take place, which, however, are not or hardly visible. In this study, we applied a time series analysis based on semantically enriched Sentinel-2 data and spectral vegetation indices (SVI's) to detect early traces of bark beetle infestation in the Berchtesgaden National Park. Results show that the best separability of healthy and early-infested pixels is in the vegetation red-edge and SWIR parts of the electromagnetic spectrum. Regarding the indices, water stress-related ones have proven to be most sensitive. This method detects initial bark beetle infestation up to three weeks before sightings in the field. From a site-specific perspective, bark beetle areas are reflected with an accuracy of 76% compared to reference data.

1 Introduction

Whilst access to environmental data grew rapidly in the last decade, constant monitoring for a full understanding of forest disturbance mechanisms is still a big challenge. The launch of the Sentinel-2 satellites within the European Copernicus programme has made a major contribution to the monitoring of ecological and forestry issues and marked a paradigm shift in Earth observation (EO) monitoring capabilities (FER et al. 2021). The ultimate goal of space assets such as EO is to convert these "big Earth data" to valuable information in order to contribute to the understanding of - inter alia - natural processes (SUDMANNS et al. 2020). When it comes to the monitoring of early stress factors for vegetation, the temporal availability of satellite data is of particular importance. In this study, we developed a method to detect early infestation stages of the European Spruce Bark Beetle (*Ips typographus L.*) using semantically enriched Sentinel-2 time series and spectral vegetation indices (SVI's). Bark beetle infestation impacts forest ecosystem dynamics and consequently affect ecosystem services. Although insect disturbances lead to disrupted structures and alter the composition of the ecosystem, they also foster diversity on the landscape and thus contribute to a various number of ecosystem services (THOM & SEIDL 2016). Regardless whether considering such disturbances an opportunity or a risk, installing a monitoring system is key to gain knowledge about the ecological processes or to limit the ecological damage caused by bark beetle infestations. Both scenarios depend on the management strategy of the respective site, which can be either more conservation or usage-oriented. The present study was conducted in the Berchtesgaden National Park (southern Germany), as a highly protected area. The major challenge in detecting infested trees timely is

[1] National Park Administration Berchtesgaden, Department of Research and Monitoring, Doktorberg 6, D-83471 Berchtesgaden, E-Mail: lisa.mandl@npv-bgd.bayern.de
[2] University of Salzburg, Department of Geoinformatics, Schillerstrasse 30, A-5020 Salzburg, E-Mail: lisa-maria.mandl@stud.sbg.ac.at, stefan.lang@plus.ac.at

the fact that they appear physiologically green and do not show symptoms clearly observable by the human eye (ABDULLAH et al. 2019a). Nevertheless, biochemical changes during the initial phase of bark beetle infestation become visible due to the interaction between electromagnetic (EM) radiation of the sun and the leaves of the trees. Thus, certain portions in the EM spectrum can be used as stress indicators (HENDRY et al. 1987).

2 Material and Methods

2.1 Data

Input satellite data are Copernicus Sentinel-2 level 2A and level 3 data. We used best-available pixel composites between March and October only where illumination conditions for mono-temporal imagery were unfavourable. The 13 channels of Sentinel-2 are optimized for land surface observations and the high resolution of up to 10 m and the swath width of 290 km are ideal for detecting changes in vegetation. The revisiting time of these satellites is about five days (ESA 2017). In total, ten Sentinel-2 scenes from March to October 2020 were used (for concrete dates of data acquisition, Fig. 3). When atmospheric conditions permitted, two images per month were acquired. As reference data, we used aerial imagery from the management zone with a spatial resolution of 0.2 m from 2020. Grey-attack stages were derived by a supervised machine learning approach using a Random Forest ensemble model in R's machine learning package *SuperML*. The resulting dataset – the grey-attack detection – shows the state at the end of 2020. This means that sometime between April and August, the early stage of bark beetle infestation must occur on these plots, as this is the period when climatically optimal conditions for an outbreak are prevalent. Exactly this knowledge is essential if certain indicators are to be derived later. The plots classified as grey-attack are referred to as "test plots" from here on. Besides that, ground truth data was included. Search teams walk the same spots in the management zone every four weeks and enter bark beetle detections into this app. Field workers identify early-infested trees by opposite slope observation, paying particular attention to minimal changes in the colour of the spruce trees. Once such a tree is identified, an in-situ sighting is carried out, looking for possible dry dust and holes in the bark. These are ground-based signs of early infestation. Once early-infested plots are recorded with the app, data can be exported as a table and visualized in a GIS. Here, we used this information to assess the temporal accuracy of the developed method.

2.2 Software

The indices-based computations and the feature engineering were both conducted using the statistical software R (version 3.6.2) and the packages *sen2r* and *SuperML*, which provides a fit to Python's "scikit learn" package (R CORE TEAM 2021; RANGHETTI et al. 2020). Besides that, we applied a knowledge-based approach using the software SIAM™. This software – as opposed to the indices – considers the entire feature space of an image and divides it into so-called spectral-based semi-concepts (pre-classes). SIAM™ (release 88v7) stands for **S**atellite **I**mage **A**utomatic **M**apper and executes a preliminary, spectral rule-based classification, which is a discrete finite set of mutually exclusive and totally exhaustive spectral-based semi-concepts also referred to as spectral categories. It considers the bands Blue, Green, Red, NIR, SWIR-1 and SWIR-2. The benefit of SIAM™ for the purpose of this study is the capability to categorize satellite data into 96 spectral-based semi-concepts (BARALDI 2011). In this way, even minor changes in the "greenness" of vegetation can be detected and computed together with the

changes in biophysical properties, represented by indices. These semi-concepts convert the continuous sub-symbolic variable "reflectance" into a given set of discrete categorical variables, whereas this process has to be reversible (BARALDI & BOSCHETTI 2012). However, these classes do not claim absolute uniqueness; several semantic categories may be valid options derived from one and the same pixel colour, like water or shadow.

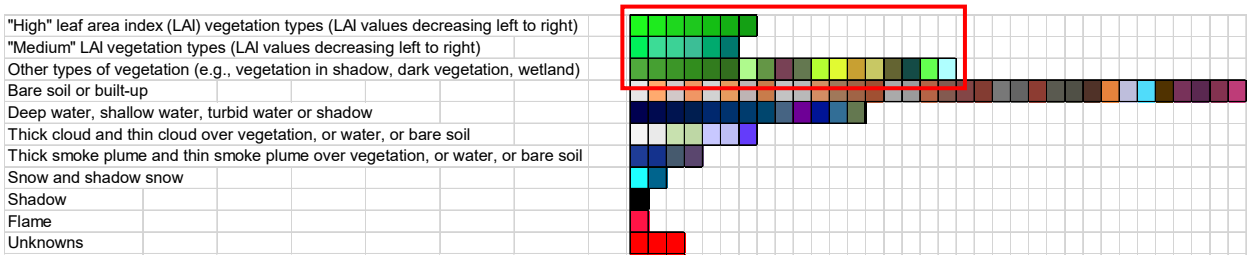


Fig. 1: Spectral-based semi concepts of the SIAM output (most important ones for this study in the red box), source: BARALDI et al. (2006)

As input, SIAM™ expects at least top-of-atmosphere (ToA) corrected data, meaning satellite data that is radiometrically calibrated into ToA (BARALDI 2011). For Sentinel-2 Level 2A data exhibiting bottom-of-atmosphere (BoA, also referred to as "SURF") correction, this requirement is given. In order to execute the SIAM™ software, Sentinel-2 Level 2A data were pre-processed including spatial resampling, image stacking and conversion to 8-bit ENVI format, the creation of an artificial thermal band as well as a noData mask. Finally, a SIAM batch file was created to calculate the spectral categories of the input data in batch mode.

2.3 Study Design

The core of this knowledge-based approach is a feature space partitioning, temporal response patterns and SVI's in combination with semantically enriched Sentinel-2 data (SIAM spectral categories). As a-priori knowledge is needed for first parametrization, the analyses were performed retrospectively for the year 2020. Grey attack test plots show where at the end of 2020 trees are attacked by bark beetles but there is no information available when infestation took place. Nevertheless, the grey-attack dataset is essential for further analysis. The two main processes during initial bark beetle infestation – subtle changes in the green tone of spruces and changes in the biochemical properties of trees – are represented by changes in the SIAM spectral categories and changes in pre-selected SVI's respectively. Both changes were solely observed within the grey-attack test plots because there is assured knowledge, that these areas must be attacked within the course of the year (Fig. 2). Within the test plots, inter-annual trajectories of pixels were observed. Besides the SIAM-based change detection, we developed a comprehensive set of six SVI's by applying a Principal Component Analysis (PCA) on 23 potentially suitable water stress and chlorophyll-related indices found in a literature research. Selected SVI's are shown in Fig. 5. As opposed to other methods, we defined index-specific thresholds based on statistical measurements. For this, we started three model runs using a set of descriptive statistics (min, max, quartiles and IQR) as potential thresholds. For each model run, the SVI change detection output was reclassified according to the thresholds. We found the minimum value serves best as lower threshold, while the 3rd quartile led to the best results for the upper threshold. The two basic assumptions are that (1) changes occuring in the SIAM spectral categories indicate bark beetle infestation at an early stage. This means that SIAM change detection values act as predictors and all changes in SIAM spectral categories within

the test plots are considered as potentially early infestation pixels. (2) the interplay with SVI's concretise these first hints from the SIAM change detection by determining whether changes occur due to bark beetle infestation or due to other factors such as phenology, poor illumination, pixel shift or non-stationary surface type properties. In order to address only the host trees of bark beetles, the norway spruce tree (picea abies) is masked out by grouping SIAM spectral categories. This compensates for a missing data set for tree identification at the species level. At the end, conditional statements classify a raster cell as *early-infested*, applying following decision rule:

$$\textbf{Early infested} = \begin{cases} 1, \text{ if SIAM predictor } \| \text{ SIAM coniferous } \| \text{ index A } \| \text{ index B } \| \\ \text{index C } \| \ \ldots = 1, \text{ 0 otherwise} \end{cases}$$

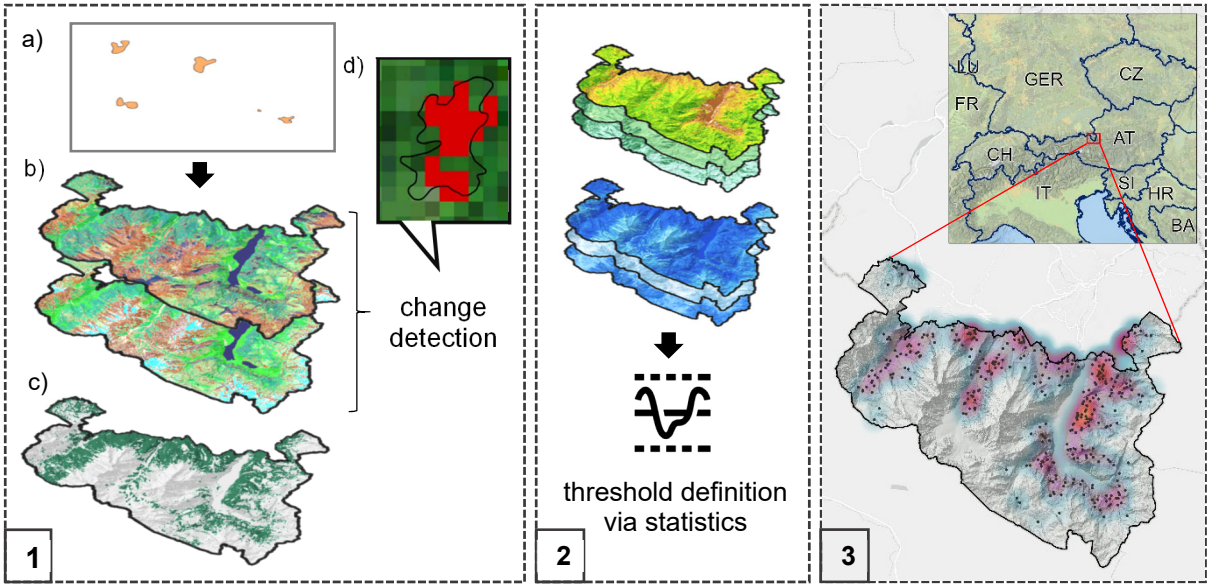


Fig. 2: Panel 1: SIAM-based analysis; a) Test plots within change detection is carried out, b) SIAM change detection, here for two time stamps, c) coniferous mask d) reclassified SIAM predictor values within test plot (black outline). Panel 2: Indices-based analysis: Change detection of chlorophyll-related indices (greenish) and water stress-related indices (blueish), based on which the thresholds are derived. Panel 3: Location of the study site and final infestation map, here for the entire year

All cells classified as *early-infested* are double-checked in order to minimize the potential error that may result from pixel shift or adverse illumination conditions. Then, in the second iteration, it is not only checked for which cells conditional statements are true, but also whether a particular cell has been classified as *early-infested* in the previous time step. If this hold true, the cell is considered infested; otherwise, it is reset to *healthy*. Finally, the impact of the six selected SVI's was analysed by computing feature importance based on random forest (RF).

3 Results

3.1 Spectral separability and temporal variation of canopy reflectance under bark beetle infestation

In total, 350 randomly selected pixels for each of the classes *healthy, early-infested* and *grey-attack* were included into the analysis; their spectral reflectance (median) was plotted together with the error bands.

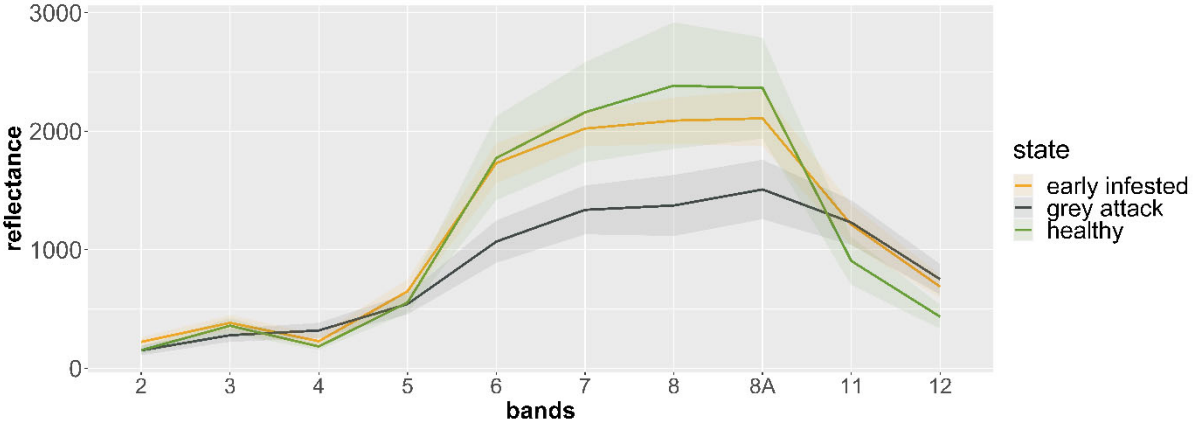


Fig. 3: Spectral behaviour of green-attack-, grey-attack and healthy pixels in the electromagnetic spectrum

Fig. 3 shows that healthy pixels differ significantly from grey-attack pixels in terms of their spectral behaviour. Especially in the red-edge bands and in the NIR and SWIR part of the EM spectrum (bands 5 to 12) the differences are clearly visible. The signal is much weaker in the VIS range. When comparing healthy pixels with the pixels classified as early-infested, there are significantly smaller differences, both in the VIS as well as in the red-edge, NIR and SWIR range. The greatest distinguishability between the two stages is seen at band 8 and 8A, where early-infested pixels have lower reflectance values and at bands 11 and 12, i.e. in the NIR-1, NIR-2 and SWIR range, where reflectance is higher compared to healthy pixels. By analysing single time stamps, inter-annual differences can be observed. Particularly big differences of reflectance per time stamp emerge for bands 8, 8A, 11 and 12 (Fig. 4).

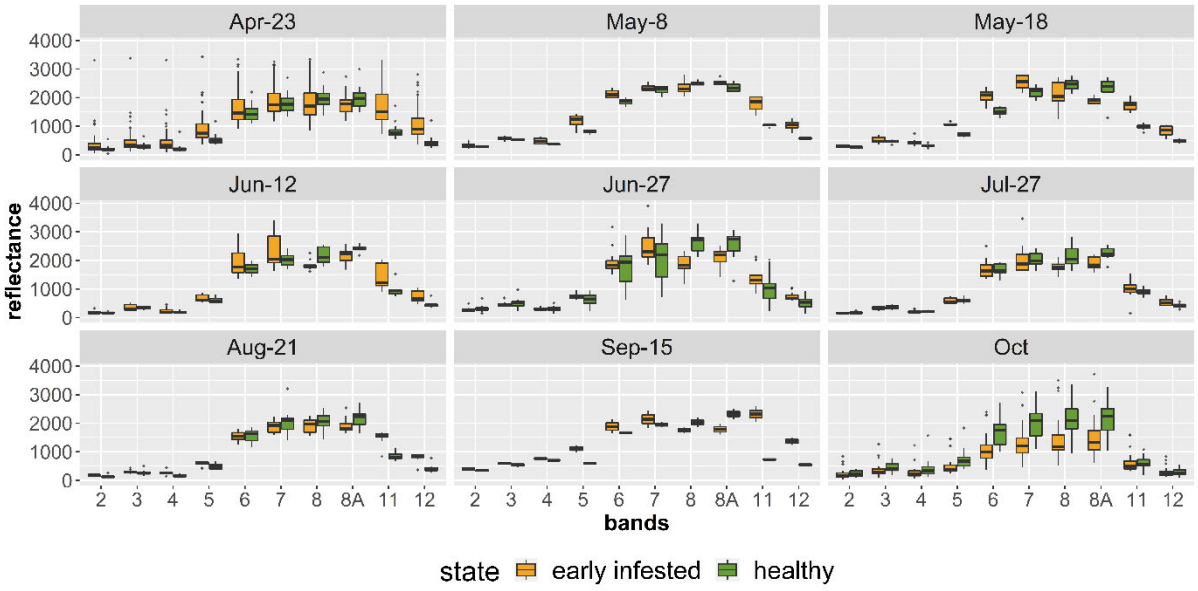


Fig. 4: Reflectance for healthy and early-infested pixels per band and time step. There is stronger scatter for healthy pixels, whereas the signal appears clearer for early-infested pixels

In these wavelength ranges, the reflectance for pixels classified as *early-infested* differs significantly from the class *healthy* as also shown by a t-test (see Tab. 1).

Tab. 1: Significance levels for the bands 8 (NIR-1), 8A (NIR-2), 11 (SWIR-1) and 12 (SWIR-2). p-values are computed by applying a student-t-test

Date	*p*-value B8	*p*-value B8A	*p*-value B11	*p*-value B12
2020-05-08	0.211	0.102	**< 0.001**	**< 0.001**
2020-05-18	0.101	0.034	**< 0.001**	**0.002**
2020-06-12	**0.020**	**0.026**	**0.013**	**0.004**
2020-06-27	**< 0.001**	**0.003**	0.065	**0.002**
2020-07-27	**0.005**	**0.001**	0.205	**0.004**
2020-08-21	0.195	**0.020**	**< 0.001**	**< 0.001**
2020-09-15	0.306	0.172	**< 0.001**	**< 0.001**
2020-10	**0.003**	**0.004**	0.667	0.938

* in bold values with $p < 0.05$

Tab. 1 shows significance levels for bands 8, 8A, 11 and 12. Especially for bands 8A, 11 and 12 – NIR-2, SWIR 1 and SWIR 2 – the *p*-value mostly is below the significance threshold of 0.05. That means the null hypothesis – there are no significant differences between healthy and early-infested pixels – should be rejected. For NIR-1 (band 8), differences seem to be less clear as it is the case for the NIR-2 and SWIR region. Only four time steps fall below the significance threshold of 0.05 and can be considered *significantly* different. The remaining *p*-values > 0.05 give evidence that differences between healthy and early-infested pixels are not significant. This is particularly true for band 8 and reflects the findings in ADAMCZYK AND OSBERGER (2015) who claim that the NIR range is particularly useful for detecting advanced stages of bark beetle infestation but not that much for early stage detection.

3.2 Impact of indices on classification result based on impurity-based feature importance

The results of the impurity-based feature importance (Gini-importance) for the whole year shows that DSWI is the most relevant index for the model output. It is followed by NDI45 and NDWI. The impact of TCW and especially NDRE3 and NGRDI is relatively low (Fig. 5a).

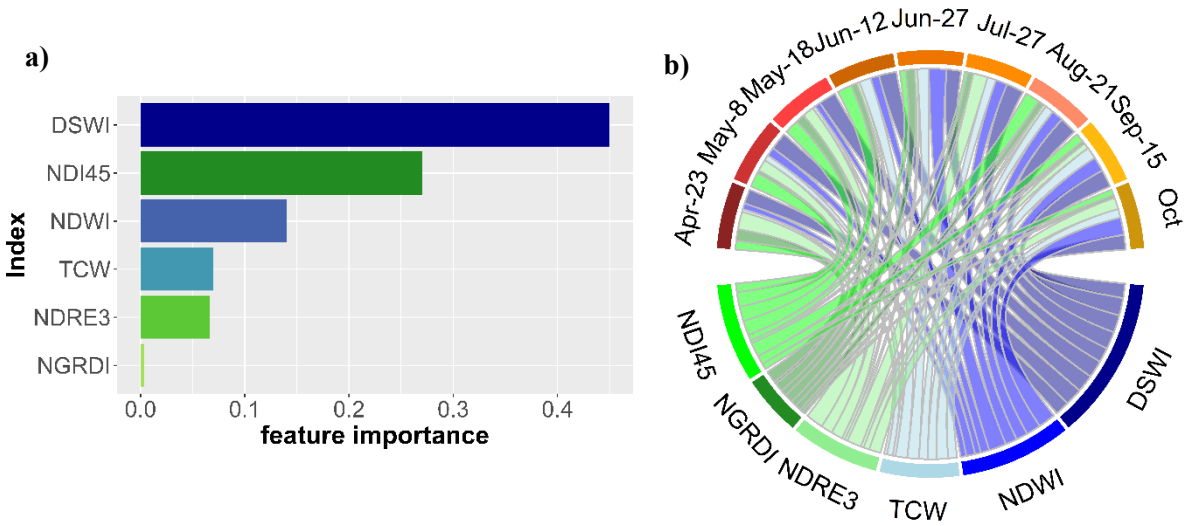


Fig. 5: a) Feature importance computed for the whole year. b) Relationship between the single time steps satellite data was acquired (top), the according indices (bottom) and the feature ranking (indicated by the thickness of the connecting lines)

When comparing predictor importance for the single time steps (Fig. 5b and 6), it can be observed that the order of indices is swapping. The chord chart (Fig. 5b) shows the interrelation between the time stamps of the satellite data, the SVI's and feature importance of SVI's. Thicker lines indicate a higher ranking of the corresponding index depending on the date. It is striking that for the months April to May the pigment- and chlorophyll-related indices (in green) dominate the top ranking positions, while from June onwards the water stress-related indices (in blue) get more important. This is also shown in Fig. 6.

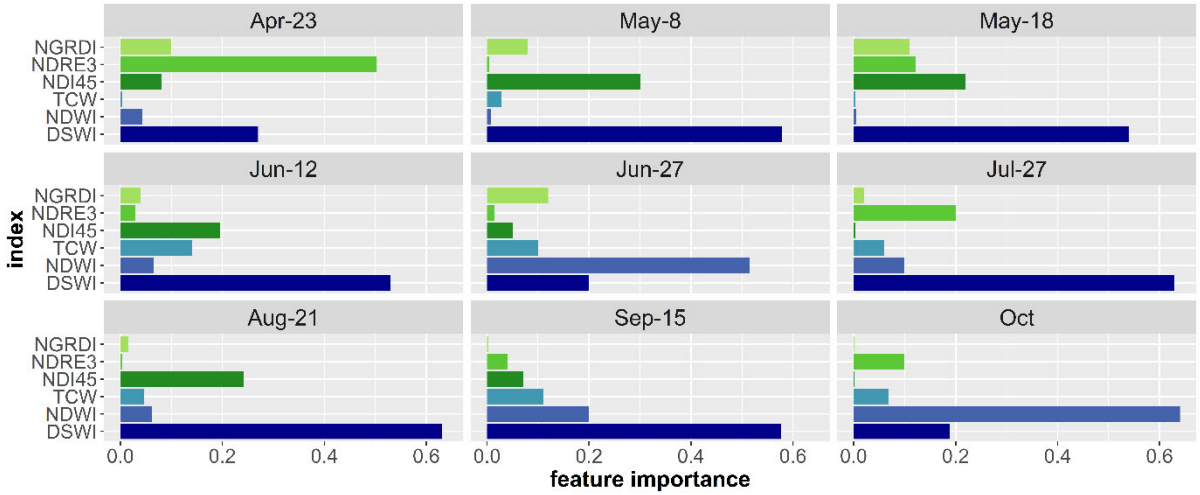


Fig. 6: Feature importance computed for single time steps. Inter-annual differences in the feature ranking is observable

Tab. 2: Accuracy of the RF model for determining feature importance

Time step	Accuracy	OOB error
April 23rd	36.22%	0.673
May 8th	54.667%	0.231
May 18th	72.00%	0.158
June 12th	62.50%	0.273
June 27th	81.667	0.108
July 27th	66.667%	0.237
August 21st	70.588%	0.143
September 15th	78.286%	0.114
October	37.99%	0.683

The seasonal dependence of SVI's on predictor importance is a key finding and should be considered in follow-up studies. The predictive performance of the RF feature importance model for the whole year is 61.8%. This value seems to be relatively low, but can be explained when evaluating feature importance and accuracy for each single time stamp, which is just like out-of-bag (OOB) error, quite variable as shown in Tab. 2. In summary, intra-annual differences in predictor importance and accuracy as well as out-of-bag (OOB) error are noticeable. The rather low accuracy of the model for April and October can be attributed to the poorer illumination conditions of the satellite data in these months.

3.3 Validation

Both spatial and temporal dimensions were considered in the validation process using the grey-attack dataset and ground truth data as reference data (cf. chapter 2.1). Notably, the data situation is not good enough for an all-encompassing and fully representative accuracy assessment. Reference data is only available for the so-called management zone of the Berchtesgaden National Park, which accounts for less than one third of the entire area (~ 6,300 ha). Nevertheless, based on available data, we will evaluate the spatial and temporal performance of the method.

From a spatial perspective, the method we developed in this study detects early bark beetle patches with an accuracy of 76% compared to the grey-attack reference dataset. For better comparability, we rasterized the polygons of the grey-attack dataset and assessed for both, on pixel and on plot level, where plots refer to a group of more than four individual trees. That means that all single trees classified as early-infested were excluded for plot level accuracy assessment. We found out two main error sources: (1) the method generally includes too many pixels that should not be considered early-infested according to the reference data (error of commission); (2) *single* infested trees cannot be detected but only patches – meaning an accumulation of trees. Reasons leading to these sources of error are discussed in the following chapter.

From a temporal perspective, earlier respectively later detection scatters from three weeks earlier until three weeks later as compared to ground truth data. The results of the temporal accuracy assessment show that there are no trends in temporal accuracy between months. Thus, bark beetle infestation is not recorded particularly early or late in any of the time steps. Accordingly, the temporal differences are most likely due to data availability, but not due to data quality. Overall, the temporal agreement between the reference data and the modelled results is good. A high proportion of the predicted values indicate bark beetle infestations prior to the findings in the field. The maximum deviation in negative direction, i.e. temporally after the detection in the field, is three weeks. Undoubtedly, there is a high dependency between

accuracy and data availability such that the timing of data acquisition is a major driver when it comes to an early detection.

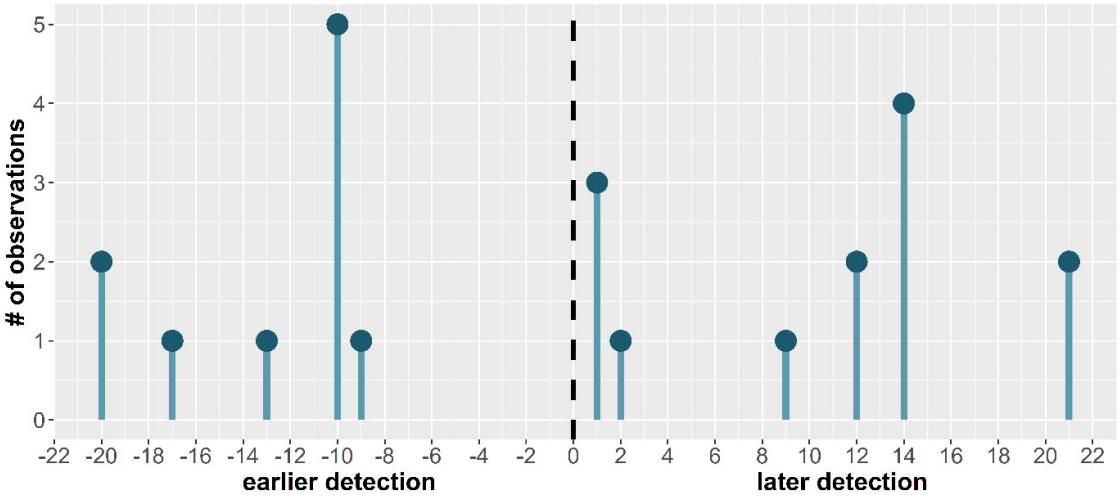


Fig. 7: Graphical representation of the temporal accuracy. The dashed line indicates the date of ground truth recording. All entries left from '0' mean earlier detection compared to ground truth data, all entries right mean later detection

4 Discussion

4.1 Methodological perspective

The method presented in this study is able to detect early stages of bark beetle infestation with the help of the lightweight software SIAM™ and the use of spectral vegetation indices. Both calculations can be realised within a very short timeframe, as an important feature when it comes to the operational use of this approach. Compared to field sights, the presented method detects bark beetle attacks up to three weeks earlier or later than the field workers do. The fairly wide range may be attributed to data availability. It is challenging to determine whether the method is really capable to detect green-attack stages. The whole workflow is based on alterations in SIAM spectral categories. Undoubtedly, the software can only indicate changes when changes already occur in the spectral range. Therefore, the term green-attack may be misleading here, which is why we use the term early detection in this study. On the other hand, the field sightings are also based on minimal spectral changes, which is why green-attack is not quite correct there either, but commonly used in fact. In summary, the term green-attack is a matter of definition and even in literature not always consistent. Moreover, bark beetle attacks are a continuous process, meaning that various environmental factors drive this particular forest ecosystem dynamic in time and space (BÁRTA et al. 2021). Furthermore, indices used as proxies, always need to be viewed critically, regardless of whether they are multi-dimensional indices such as the Tasseled Cap Components or unidimensional indices. This "lack of confidence" is attributable to the difficult understanding and interpretation of them. Of course, they give hints to facts like a decrease in the water availability. However, they also reflect atmospheric and landscape variations that are extremely difficult to impossible to filter out or to map them to fixed points on the index scale (MOFFIET et al. 2006). The information content of spectral vegetation indices is therefore limited. On the other hand, the use of indices has already proven its great potential in a variety of applications. Thus, they definitely have their

right to exist in remote sensing, but at the same time, their limitations must be known and taken into account.

4.2 Resolution, data quality and data availability

Concerning the effect of data resolution, there will always be mixed pixels at a spatial resolution of 10 m. Detection methods – no matter if early detection or grey-attack detection may come to its limit when there are pixels that do not purely represent spruce stands (MEDDENS et al. 2013). An evaluation of spectral mixture effects would require very high resolution (VHR) data, i.e. UAV data or the application of a radiative transfer model like the DART model, which can be used for sensitivity analysis of forest reflectance (BÁRTA et al. 2021). Generally speaking, high spatial resolution is necessary when bark beetle infestations occur at endemic levels, as the infested areas are usually isolated patches. In contrast, the relevance of spatial resolution decreases when the infestation is epidemic, since the infested areas are then mostly large, contiguous patches (FERNANDEZ-CARRILLO et al. 2020; KAUTZ et al. 2011). In addition, radiometric resolution is a potential source of error, as it affects the actual information content, and the ability to differentiate between slight differences in EM reflectance. This is especially important for this application purpose. It could be argued in this context that a better result could have been achieved if the input for SIAM™ did not have to be scaled down to 8 bit. However, since 8-bit data is a practical constraint for using SIAM™, such an assumption cannot be verified. Finally, it must also be remembered that high radiometric resolution alone is not sufficient to detect early-infested bark beetle trees. ABDULLAH et al. (2019b) proved that in a comparison of Sentinel-2 and Landsat data.

4.3 Accuracy

The high error of commission (39%) in the spatial accuracy assessment can primarily explained by the relatively early image flight campaign (early September), to which the grey-attack machine learning approach is applied and from which the grey-attack reference dataset results. Especially from July onwards, the accuracy assessment is subject to uncertainties, as it is highly likely that trees will not have discoloured when reference data was acquired and thus will not be detected by the grey-attack algorithm, even if these trees are infested. All pixels classified as *early-infested* from July and later, but not appearing in the grey-attack dataset, would thus require re-analysis with data from subsequent years. When evaluating the temporal accuracy, it should be mentioned that field recording of early infestation is not "obvious", meaning that even in the field the error is relatively high. Thus, the absence of reference data on a spot classified as *early-infested* does not necessarily indicate an error of omission, but may equally be due to difficulty in locating infested spruce trees, especially in a challenging terrain such as in the Berchtesgaden National Park. Finally, accuracy may be improved by the availability of data that classifies forest stands at the species level. At the moment, our approach is applied to all coniferous stands, no matter if there is Norway spruce or any other coniferous tree present.

5 Conclusion and Outlook

Our proposed method provides a novel prototypical tool to detect forest stress from bark beetles by considering space, time and spectral characteristics of remotely sensed data. The method can be considered a form of "hybrid AI", i.e. combining a deductive, learning-from-rules approach with an inductive, learning-from-samples method. The distinctive feature of the

method is the use of semantically enriched Sentinel-2 data via the highly advanced, knowledge-based software SIAM™ and the derivation of predictor values, which contributed significantly to the success of this study. Unlike the use of SVI's alone, spectral categories of SIAM™ include the complete feature space, so that no – possibly relevant – information is lost. This unique research design outperforms conventional methods of early bark beetle detection, both in terms of the time aspect of fieldwork and the accuracy of other remote sensing approaches. Further advantages of using SIAM™ is the complete automation of the software, which does not require any parameters from users or samples. The high level of automation also contributes to the robustness to changes in input data. In addition to Sentinel-2 data, all other at least top-of-atmosphere corrected optical satellite can be used as input, including VHR data exhibiting infrared bands. Thus, the workflow can also be adapted for other scale levels. A shortcoming of this method is that single trees cannot be detected due to the resolution of Sentinel-2 data. In addition, at least two satellite scenes per month are required to monitor the early stages of bark beetle infestation reliably. Otherwise, temporal accuracy decreases significantly. Results from this approach proved that especially water-stress related indices as well as the vegetation red-edge and the SWIR ranges of the EM spectrum show high sensitivity for early bark beetle detection. Future development shall focus on building up a samples database in order to apply more sophisticated and self-adaptive deep learning algorithms, which are currently state-of-the-art in recognition performance. This allows also to run the tool in real-time, i.e. without a-priori data (here the grey-attack dataset). Alternatively, this approach can be repeated for several years to derive more robust index-specific thresholds, thus staying with a purely index-based method. Regardless of the choice of the methodology, a monitoring system that detects bark beetle infestations at an early stage makes an important contribution both to the prevention of large-scale tree mortality and to ecological research. The interaction of modern, satellite-based technologies and the on-site knowledge of forest experts proves to be very beneficial and shows that ecological field research and remote sensing-based monitoring complement each other very well and profitably.

6 References

ABDULLAH, H., SKIDMORE, A., DARVISHZADEH, R. & HEURICH, M., 2019a: Timing of red-edge and shortwave infrared reflectance critical for early stress detection induced by bark beetle (Ips typographus, L.) attack. International Journal of Applied Earth Observation and Geoinformation, **82**, 101900.

ABDULLAH, H., SKIDMORE, A. K., DARVISHZADEH, R. & HEURICH, M., 2019b: Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (Ips typographus, L.) compared with Landsat-8. Remote Sensing in Ecology and Conservation, **5**(1), 87-106. https://doi.org/10.1002/rse2.93. https://doi.org/10.1002/rse2.93.

ADAMCZYK, J. & OSBERGER, A., 2015: Red-edge vegetation indices for detecting and assessing disturbances in Norway spruce dominated mountain forests. International Journal of Applied Earth Observation and Geoinformation, **37**.

BARALDI, A., 2011: Satellite Image Automatic Mapper™ (SIAM™) - A Turnkey Software Executable for Automatic Near Real-Time Multi-Sensor Multi-Resolution Spectral Rule-Based Preliminary Classification of Spaceborne Multi- Spectral Images. Recent Patents on Space Technology, **1**, 81-106.

BARALDI, A. & BOSCHETTI, L., 2012: Operational Automatic Remote Sensing Image Understanding Systems: Beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA). Part 1: Introduction. Remote Sensing, **4**(9), 2694-2735. https://www.mdpi.com/2072-4292/4/9/2694.

BARALDI, A., PUZZOLO, V., BLONDA, P., BRUZZONE, L. & TARANTINO, C., 2006: Automatic Spectral Rule-Based Preliminary Mapping of Calibrated Landsat TM and ETM+ Images. Geoscience and Remote Sensing, IEEE Transactions on, **44**, 2563-2586.

BÁRTA, V., LUKEŠ, P. & HOMOLOVÁ, L., 2021: Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. International Journal of Applied Earth Observation and Geoinformation, **100**, 102335. https://www.sciencedirect.com/science/article/pii/S0303243421000428.

DELEGIDO, J., VERRELST, J., ALONSO, L. & MORENO, J., 2011: Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. Sensors, **11**(7).

ESA. 2017: Sentinel-2: Color vision for Copernicus.

FER, I., GARDELLA, A. K., SHIKLOMANOV, A. N., CAMPBELL, E. E., COWDERY, E. M., DE KAUWE, M. G., DESAI, A., DUVENECK, M. J., FISHER, J. B., HAYNES, K. D., HOFFMAN, F. M., JOHNSTON, M. R., KOOPER, R., LEBAUER, D. S., MANTOOTH, J., PARTON, W. J., POULTER, B., QUAIFE, T., RAIHO, A., SCHAEFER, K., SERBIN, S. P., SIMKINS, J., WILCOX, K. R., VISKARI, T. & DIETZE, M. C., 2021: Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration. Global Change Biology, **27**(1), 13-26.

FERNANDEZ-CARRILLO, A., PATOČKA, Z., DOBROVOLNÝ, L., FRANCO-NIETO, A. & REVILLA-ROMERO, B., 2020: Monitoring Bark Beetle Forest Damage in Central Europe. A Remote Sensing Approach Validated with Field Data. Remote Sensing, **12**(21).

GALVÃO, L. S., FORMAGGIO, A. R. & TISOT, D. A., 2005: Discrimination of sugarcane varieties in Southeastern Brazil with EO-1 Hyperion data. Remote Sensing of Environment, **94**(4), 523-534. http://www.sciencedirect.com/science/article/pii/S0034425704003669.

HARDISKY, M., KLEMAS, V. & SMART, A., 1983: The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina Alterniflora canopies. Photogrammetric Engineering and Remote Sensing, **48**, 77-84.

HENDRY, G. A. F., HOUGHTON, J. D. & BROWN, S. B., 1987: The degradation of chlorophyll - A biological enigma. New Phytologist, **107**(2), 255-302.

HUNT, E., DAUGHTRY, C., EITEL, J. & LONG, D., 2011: Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index. Agronomy Journal, **103**, 1090.

KAUTZ, M., DWORSCHAK, K., GRUPPE, A. & SCHOPF, R., 2011: Quantifying spatio-temporal dispersion of bark beetle infestations in epidemic and non-epidemic conditions. Forest Ecology and Management, **262**, 598-608.

MEDDENS, A. J. H., HICKE, J. A., VIERLING, L. A. & HUDAK, A. T., 2013: Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. Remote Sensing of Environment, **132**, 49-58. https://www.sciencedirect.com/science/article/pii/S0034425713000060.

MOFFIET, T., MENGERSEN, K., KING, R., ARMSTON, J. & WITTE, C., 2006: Modelling of foliage projected cover using Landsat-7 spectral imagery: Spectral indices for greenness and brightness. 13th Australian Remote Sensing and Photogrammetry Conference, Canberra, Australia, 20-24.

R CORE TEAM, 2021. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

RANGHETTI, L., BOSCHETTI, M., NUTINI, F. & BUSETTO, L., 2020: sen2r: An R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data. Computers & Geosciences, **139**. https://sen2r.ranghetti.info.

SENF, C., PFLUGMACHER, D., HOSTERT, P. & SEIDL, R., 2017: Using Landsat time series for characterizing forest disturbance dynamics in the coupled human and natural systems of Central Europe. ISPRS journal of photogrammetry and remote sensing : official publication of the International Society for Photogrammetry and Remote Sensing (ISPRS), **130**, 453-463. https://doi.org/10.1016/j.isprsjprs.2017.07.004.

SHI, Y., WANG, T., SKIDMORE, A. K. & HEURICH, M., 2020: Improving LiDAR-based tree species mapping in Central European mixed forests using multi-temporal digital aerial colour-infrared photographs. International Journal of Applied Earth Observation and Geoinformation, **84**. http://www.sciencedirect.com/science/article/pii/S0303243419300340.

SUDMANNS, M., TIEDE, D., LANG, S., BERGSTEDT, H., TROST, G., AUGUSTIN, H., BARALDI, A. & BLASCHKE, T., 2020: Big Earth data: disruptive changes in Earth observation data management and analysis? International Journal of Digital Earth, **13**(7), 832-850. https://doi.org/10.1080/17538947.2019.1585976.

THOM, D. & SEIDL, R., 2016: Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests: Disturbance impacts on biodiversity and services. Biological Reviews, **91**.

Forecasting Urban Development from Satellite Images

NANDO METZGER[1]

Abstract: Forecasting where and when new buildings will emerge is a rather unexplored niche topic, but relevant in disciplines such as urban planning, agriculture, resource management, and even autonomous flight. In this work, we present a method that accomplishes this task using satellite images and a custom neural network training procedure. In stage A, a DeepLapv3+ backbone is pretrained through a Siamese network architecture aimed at solving a building change detection task. In stage B, we transfer the backbone into a change forecasting model that relies solely on the initial input image. We also transfer the backbone into a forecasting model predicting the correct time range of the future change. For our experiments, we use the SpaceNet7 dataset with 960 km² spatial extension and 24 monthly frames. We found that our training strategy consistently outperforms the traditional pretraining on the ImageNet dataset. Especially with longer forecasting ranges of 24 months, we observe F1 scores of 24% instead of 16%. Furthermore, we found that our method performed well in forecasting the times of future building constructions. Hereby, the strengths of our custom pretraining become especially apparent when we increase the difficulty of the task by predicting finer time windows.

1 Introduction

Understanding the evolution of land use (e.g., constructions and change of land type) is crucial in fields like urban planning, agriculture, natural resource management, anticipating housing market prices, and even autonomous driving and flying. One of the critical components in purely vision-based positioning systems is having an up-to-date map of the environment. Especially in emergencies, it is crucial to have a map with minimal uncertainty to find safe landing sites. These uncertainties are reducible by regularly conducting resource-intensive survey flights. However, a system that can anticipate the changes could assist in indicating the locations of future survey flights and potentially save unnecessary flights in regions with no change.

While existing approaches aim to forecast Land use/Land cover class changes, none explicitly targets building footprint as the output variable. Moreover, there appears to be a lack of research using advanced deep learning methods that aim to forecast the time of when the changes will occur. In this work, we aim to forecast where and when changes in building footprints are going to happen. We consider the problem of change forecasting as a binary segmentation problem with labels "change" and "no change", whereas the forecasting range of the model is already implicitly defined by the selection of the training data (i.e., using a consistent forecasting range in the training samples). We use satellite data series as the primary input data source because they provide global, uniform coverage.

We develop a 2-stage training strategy that is centered around the state-of-the-art segmentation network DeepLabv3+ (CHEN et al. 2018). The stages are as follows:

[1] ETH Zürich, Professorship for Photogrammetry and Remote Sensing
E-Mail: nando.metzger@geod.baug.ethz.ch

- **Stage A**: We train a Siamese change detection network on the task of change detection (See Fig. 1 for the task overview).
- **Stage B**: We use the pretrained backbone of stage A, to build two variants of forecasting models. The first model variant deals with the task of change forecasting, the other with the task of time range forecasting task (see Fig. 1 for the task overview).

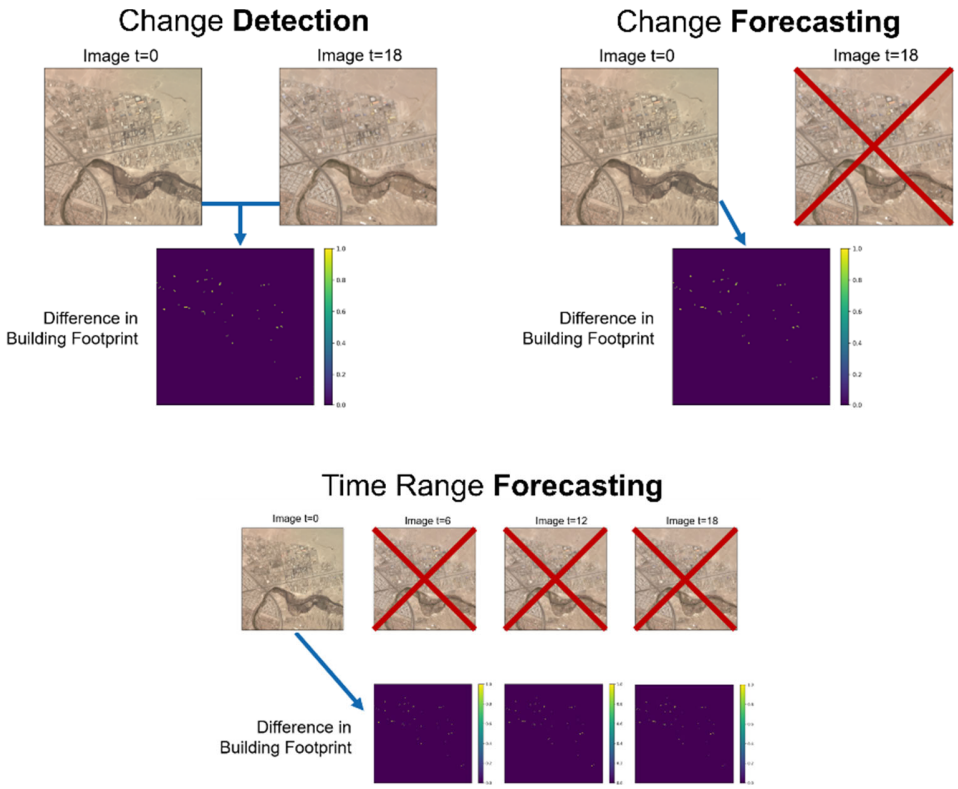


Fig. 1: Visual overview of tasks. The red crosses stress the unavailability of a variable in each task.

2 Methodology

2.1 Data

To validate our method, we make use of the SpaceNet7 dataset. It is published by VAN ETTEN 2021 and created for the building tracking competition featured at the NeurIPS 2020. The dataset consists of 60 locations, each containing a series of 24 Planet Labs satellite image mosaics of 4×4 km at timestamps that are one month apart. The ground sampling distance of the images is 4 m, and the total covered area of this dataset is 960 km^2. The dataset also contains a set of human-labeled building footprints, where each image of the time series is labeled individually.
For this work, we derive a dataset that considers image pairs from the same locations but at different time steps and calculate the changes in the corresponding building footprints. Note, that for simplicity reasons, we neglect the cases where buildings are removed. Using this method, we obtain about 11,000 unique image pairs. If we divide them into training patches of 224×224 pixels,

we get 237,000 mutually exclusive pairs. However, for the task of change forecasting, one implicitly defines the forecasting range by the choice of the training sub-datasets with a consistent forecasting range. For example, when we want to forecast for the smallest possible range (i.e., one month), the subset of pairs that are one month apart contains 5,000 samples, whereas, with the largest forecasting range of 24 months, we can only obtain 200 samples. The dataset also exhibits a severe label imbalance. The average percentage of positive pixels in the whole dataset amounts to 0.3%, and only every seventh patch has more than 0.5% positive labels.

2.2 Experimental Setup

We allocate the locations in the dataset into a training set with 42 locations (70%), a validation set for parameter tuning with 6 locations (10%), and a test set to calculate the final metrics with 12 locations (20%). As our evaluation metric, we use the optimal F1 score. The optimal F1 score has appealing theoretical properties, but it uses knowledge about the true labels to classify the predictions. Therefore, we also implement our adaptive threshold by a moving average over the last 500 optimal training thresholds. We found that the discrepancies between the optimal F1 score and the F1 score from our adaptive threshold are diminishingly small, such that we only report the optimal F1 score.

2.3 General Idea

The task of change forecasting is naturally ill-posed. One satellite image alone does not define for which future time range the binary classification of "change"/ "no change" should be predicted. It is therefore important that the model is trained on a subset of pairs with a fixed temporal difference, which implicitly defines the forecasting range. However, using these small subsets, one cannot exploit the full potential of the data. To still use all samples, we propose to first train the components of the model to perform the change detection task, where the time range is explicitly defined by the two input images of each pair. We call this pretraining procedure stage A, whereas our main task of change forecasting is called stage B (Fig. 2).

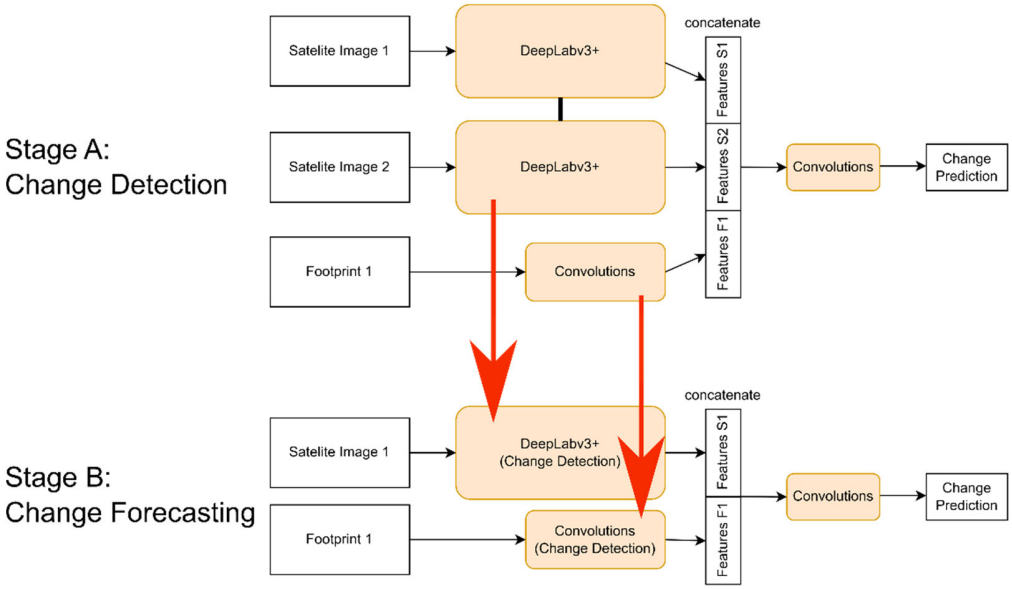


Fig. 2: Modular model architectures and general workflow.

2.4 Stage A: Change Detection

In stage A, we train a Siamese change detection network with the DeepLabv3+ backbone on the task of change detection using ordered pairs of satellite images. On the one hand, we can use all the available training pairs in this stage. On the other hand, this type of pretraining already allows the adaption to the satellite images domain, in contrast to the traditional pretraining that is usually performed on the ImageNet domain. We use the building footprints belonging to the first image as an additional feature and process it using a convolutional neural network. Please refer to Fig. 2 for further detail about the architecture. As a loss function, we use the binary cross-entropy loss.

2.5 Stage B: Forecasting

We extract the components of the change detection model of stage A and use them as initializations for two versions of forecasting models:

- The **change forecasting model** predicts a binary change label and is trained using the binary cross-entropy loss. We train one classifier for each of the sub-datasets with forecasting ranges {1,3,6,9,12,15,18,21,24}, which helps us better understand the strengths and weaknesses of our approach.
- The **time range forecasting model** is set up in a multitasking learning approach, where the model needs to predict the time of the change on an ordinal scale (*n* time windows) and the binary change label. We fix the maximal range to *24* months, and in a first set up, we split it into two prediction windows – e-g, *n=2*, such that the two ordinal labels are defined as *1-12*, and *13-24* months respectively. Furthermore, we also test a setup of three labels (*n=3*) where the labels are defined as *1-8*, *9-16*, and *17-24* months. The output dimension of the model is *n+1*, where the last output represents the "no change"-class. Finally, we train the model using a multi-class- and a binary cross-entropy loss for time range forecasting and change forecasting, respectively.

In both cases, we first freeze the backbone and footprint convolutions and train the remaining model for 5,000 batch iterations. We then reduce the learning rate by a factor of 10 and train the entire model.

3 Results and Discussion

3.1 Change Forecasting

In the comparison of Fig. 3 (left), we notice a consistent improvement by using our pretraining strategy compared to using the traditional ImageNet-pretraining strategy. Most notably, the model experiences a much more significant absolute increase in performance of +8% F1 score for the forecasting range of 24 months. The baseline models are only trained on the sub-datasets that correspond to their forecasting range, while our training strategy implicitly allows the models to access the information contained in the whole dataset through the pretraining of stage A. Therefore, it makes sense that the performance gap gets more significant as the forecasting range progresses because that is also when the availability of change forecasting samples is scarce.

Furthermore, we provide the precision-recall curve in Fig. 3 (right) for the classifier that is assigned to the 24 months prediction range. It shows what tradeoffs may be considered when implementing a real-world system.
Fig. 4 helps to get an understanding of which visual cues the model is using to make predictions. It is apparent that the models mainly go for construction sites and cluttered or empty spaces in the vicinity of already existing buildings.

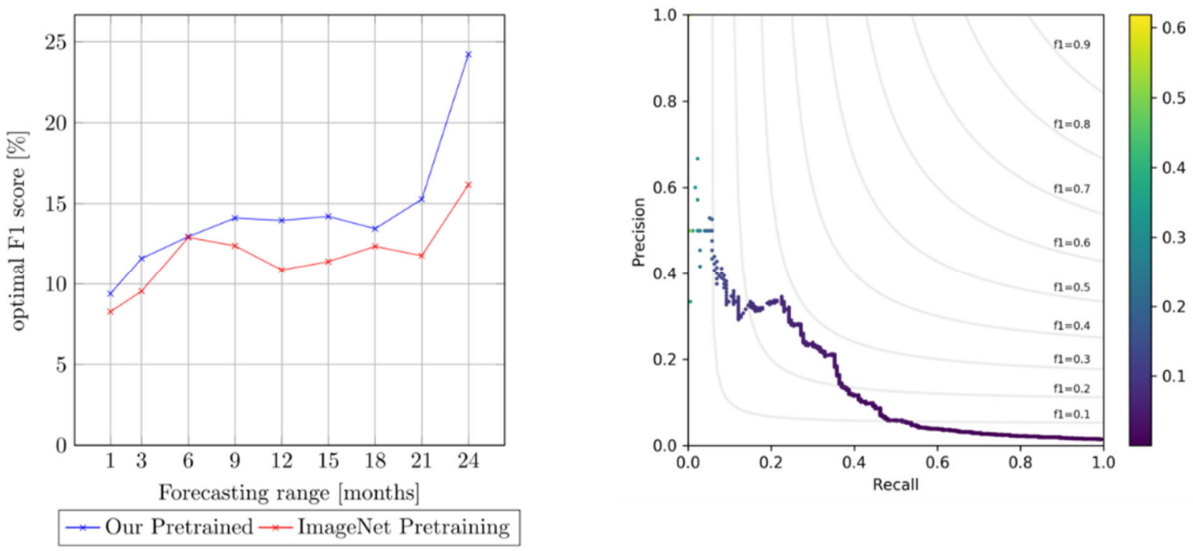


Fig. 3: Left: Performance of our pretraining strategy compared to ImageNet. Each dot denotes a different model that is specialized in the corresponding time range; Right: The precision-recall curve of the model that predicts the 24 months window. The colors of the dots resemble the corresponding thresholds

3.2 Time Range Forecasting

The resulting confusion matrices and F1 scores for the time range forecasting tasks with two and three target labels are presented in Tab. 1 and

Tab. 2, respectively. The results attest that all classifiers exhibit a noticeable signal on the diagonals of their confusion matrices. Our pretraining method outperforms the baseline method in terms of F1 score in both tasks, whereas the lead grows significantly for the three-label setting, which is the more challenging task. Moreover, we note that the class 9-16 months in the three-label setting creates the most considerable confusion. After all, it seems that the pretraining is effective enough to learn an operational classification.

Tab. 1: Setup with two labels (1-12 months, and 13-24 months).
The rows are the ground truth classes, while the columns are the predicted classes.

	Our Pretraining (F1= 68.4%)			**ImageNet Pretraining** (F1= 67.6%)	
	(1-12)m	***(13-24)m***		***(1-12)m***	***(13-24)m***
(1-12)m	44k	28k		39k	33k
(12-24)m	19k	60k		14k	65k

Tab. 2: Setup with three labels (1-8 months, 9-16 months, and 17-24 months).
The rows are the ground truth classes, while the columns are the predicted classes.

	Our Pretraining (F1= 48.2%)				**ImageNet Pretraining** (F1= 42.2%)		
	(1-8)m	***(9-16)m***	***(17-24)m***		***(1-8)m***	***(9-16)m***	***(17-24)m***
(1-8)m	25k	10k	18k		26k	16k	13k
(9-16)m	4k	13k	22k		11k	9k	21k
(17-24)m	2k	16k	32k		6k	16k	29k

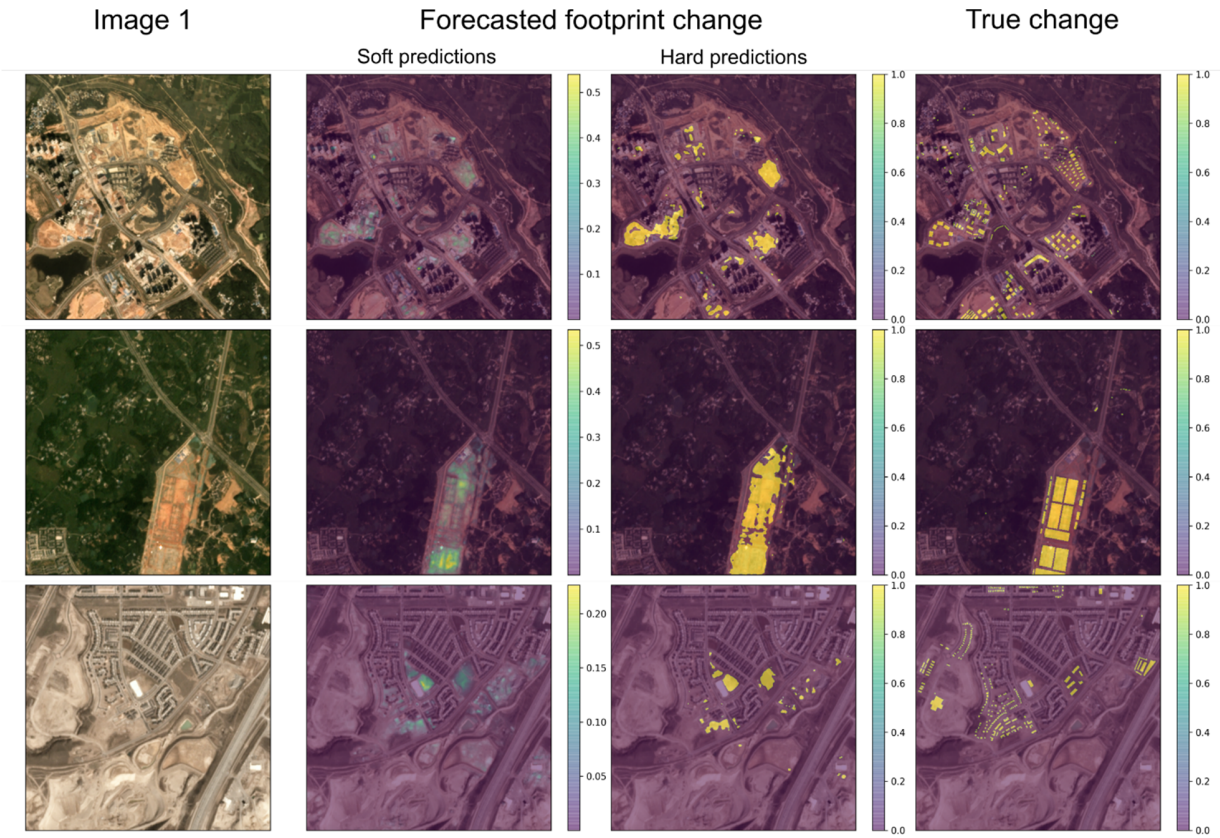


Fig. 4: Visual samples of the 18 months change forecasting task. We overlay the predictions and ground truth labels with their corresponding input image for better interpretability.

4 Conclusion and Outlook

The main contribution of this work is a model to forecast where and when changes in building footprints will happen. In this niche, there seems to be a scarcity of research. Furthermore, we develop a 2-stage transfer learning procedure that is superior to the traditional applied ImageNet pretraining and validate it on the task of change forecasting and time range forecasting. In stage A of our methodology, we pretrain a Siamese DeepLabv3+ on the task of change detection. In stage B, we use the pretrained components to tackle the task of change forecasting and time range forecasting. In all tasks, our method performs consistently better than traditional pretraining on ImageNet. In particular, for change forecasting, the method proved to be the most efficient for the

longer forecasting ranges - i.e., for the 24-month range, it achieves an F1 score of 24%, while the baseline achieves only 16%. In the range forecasting task, we found the largest performance increases for the more challenging case where three time ranges need to be predicted.
We argue that the tasks at hand with these particular settings are very challenging. On the one hand, very few visual cues give away a potential future change in land use. On the other hand, building change detection and forecasting tasks inherently lead to a significant imbalance in the label distribution, as new buildings are built very slowly and infrequently.
In our future work, we want to exploit the temporal dimension of the data more efficiently. We assume that time series analysis, such as convolutional recurrent neural networks (RNNs), can yield a model that recognizes spatial and temporal patterns indicating future changes. Moreover, the usage of higher resolution imagery would reveal more valuable cues to solve the task.

5 Bibliography

CHEN, L. C., ZHU, Y., PAPANDREOU, G., SCHROFF, F. & ADAM, H., 2018: Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the European conference on computer vision (ECCV), 801-818, https://arxiv.org/abs/1802.02611v3.

VAN ETTEN, A., HOGAN, D., MANSO, J. M., SHERMEYER, J., WEIR, N. & LEWIS, R., 2021: The Multi-Temporal Urban Development SpaceNet Dataset. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6398-6407.

AUTORENVERZEICHNIS

A

ACKERMANN, SYLVIA 297
ADLER, PETRA 40
ALAMOURI, AHMED 1
ARMATEY, LILLIAN 208

B

BANNEHR, LUTZ 220
BORRMANN, PETER 65
BRANDMEIER, MELANIE 27
BRAUNECK, JENS 9
BRUNN, ANSGAR 52, 173, 208
BUDDE, LINA E. 256, 304

C

CAMENZIND, MORITZ 19
CHEEMA, JASMIN 220
COORS, VOLKER 237
CRAMER, MICHAEL 193

D

DONAUBAUER, ANDREAS 27
DREES, LUKAS 140
DUBOVYK, OLENA 330

E

EHRHARDT, INA 84
ELTNER, ANETTE 243
ERASMI, STEFAN 117
ERBE, KARIN 27

F

FEUERSTEIN, ULF 19

G

GANZ, SELINA 40
GERKE, MARKUS 1, 149
GOCHT, ALEXANDER 117
GOEBEL, MONA 304
GOLAB, ANTONIA 315

H

HAALA, NORBERT 280
HÄNDLER, VERENA 52
HERZIG, PAUL 65
HEURICH, MARCO 330
HOLSTEIN, KATHARINA 84
HOLWEG, DANIEL 186

I
IWASZCZUK, DOROTA 94, 256, 304

J
JIANG, YUCHANG 322
JOACHIM, LENA 297

K
KILIAS, DAVID 65
KIRMER, ANITA 220
KLÜCK, HANS-CHRISTIAN 65
KNAUER, UWE 65, 84
KOLBE, THOMAS H. 27
KÖLLE, MICHAEL 290
KÖNIG, SIMON 330
KOPPENHAGEN, INES 186
KRÜGER, NINA 186
KRZEPEK, KATRIN 94

L
LAMPERT, ASTRID 1
LANG, STEFAN 342
LAUSCH, ANGELA 220
LIEBOLD, FRANK 105
LIEBSCHER, JAN-ANDREAS 27
LOBERT, FELIX 117
LUBIG, PATRICK 193

M
MAAS, HANS-GERD 105, 127
MABOUDI, MEHDI 149
MADER, DAVID 127, 243
MANDL, LISA 342
MARASHDEH, QUSAI 140
MAURER, ANDREAS 65
MAWAS, KARAM 149
MECHERNICH, SILKE 160, 267
MENSAH, EDMUND 208
METZGER, NANDO 355
MEYER, THERESA 173
MINNICH, THEA 208
MUERTH, MARKUS 186

N
NIE, RUWEI 193

O
OKYERE, FRANZ 208
OTOO-KWOFIE, CARLOS 208

P
PADSALA, RUSHI 237
PILLEN, KLAUS 65
PROKOPH, SOPHIE 220

R
REUß, OLIVER 52
RICHTER, KATJA 127
RÖDER, NORBERT 117
ROSCHER, RIBANA 140
ROTH, JONAS 52

S
SANTHANAVANICH, THUNYATHEP 237
SCHMELZ, DENNY 84
SCHMIDT, JAKOB 94
SCHMITT, MICHAEL 27
SCHNEIDER, PHILIPP JACOB 280
SCHULZE, ROBIN 243
SCHWIEDER, MARCEL 117
SEIFFERT, UDO 65
SELENSKI, JÜRGEN 52
SKUDDIS, DAVID 280
SÖRGEL, UWE 290
SPROLL, MORITZ 208
STEIER, JANIK 256
STILLA, UWE 173

T
THALER, MICHAEL 19
THIELERT, BONITO 84
THONFELD, FRANK 330

W
WEIß, ROBERT 160, 267
WESTFELD, PATRICK 127
WIRTH, RAINER 9
WÜRSTLE, PATRICK 237

X
XUE, RUIHANG 290

Y
YU, KANG 19

Z
ZHANG, WEI 280
ZHANG, XINLONG 290
ZULKOWSKI, MARCO 186

PROGRAMMKOMITEE

Abmayr, Thomas

Bareth, Georg

Bongartz, Jens

Brunn, Ansgar

Coors, Volker

Cramer, Michael

Gerke, Markus

Gülch, Eberhard

Haala, Norbert

Hastedt, Heidi

Hochschild, Volker

Hoegner, Ludwig

Jacobsen, Karsten

Jung, András

Kaufmann, Viktor

Kersten, Thomas P.

Kleinschmit, Birgit

Krzystek, Peter

Luhmann, Thomas

Mandlburger, Gottfried

Mocnik, Franz-Benjamin

Motagh, Mahdi

Nebiker, Stephan

Reigber, Andreas

Reiterer, Alexander

Schmitt, Michael

Schneider, Danilo

Sörgel, Uwe

Spreckels, Volker

Vohland, Michael

Wagenknecht, Stefan

Weinmann, Martin

Westerholt, René