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

 DOI: 10.24407/KXP:1796026646

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.

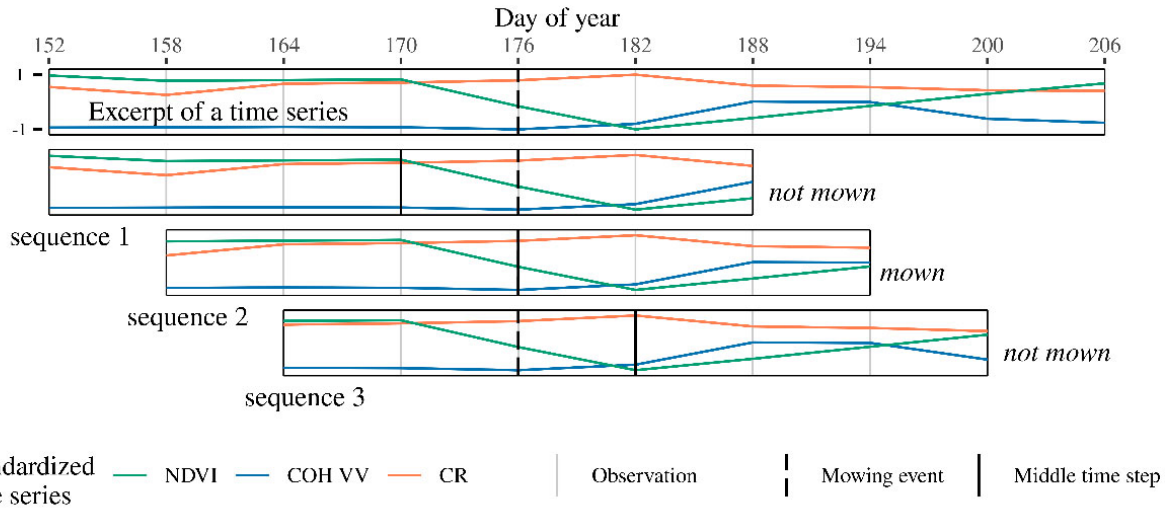

Fig. 1: Schema of the moving window approach applied to an exemplary excerpt of the time series of a grassland parcel in 2019. Depending on whether there was a mowing event on the middle time step of a resulting sequence, it is labeled as mown or not mown.

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.

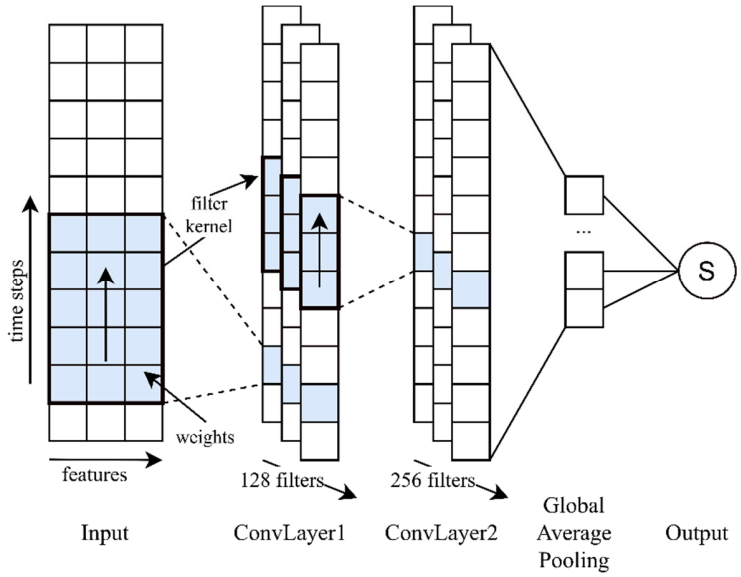

Fig. 2: Schematic architecture of the 1D-CNN by Lobert et al. (2021). The model consists of an input layer with three features of 11 timesteps. The input layer is followed by two convolutional Layers with kernel sizes five and three and filter numbers 128 and 256, respectively. Next is a global pooling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

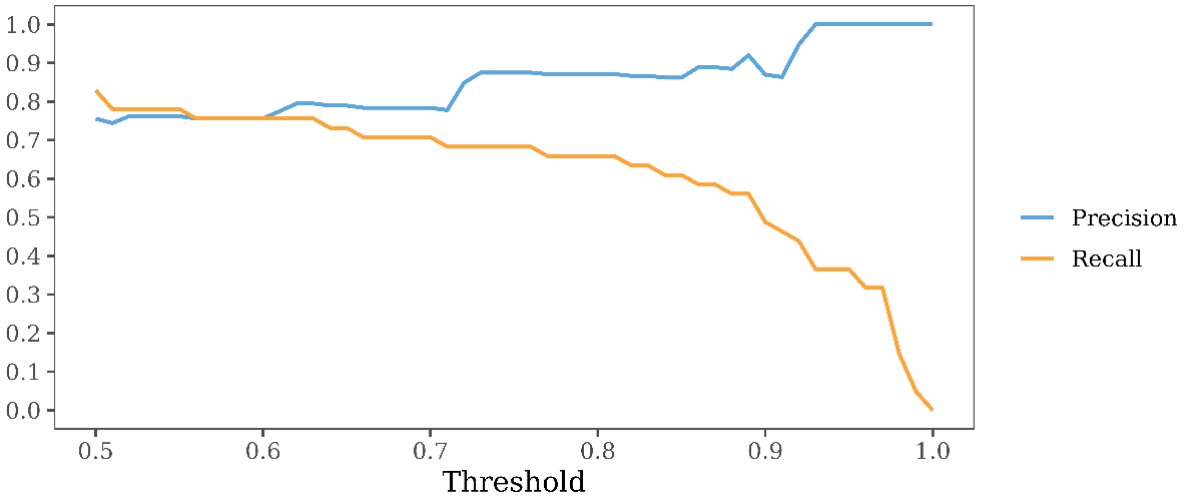

Fig. 3: Precision and recall measured on the grassland test data depending on different threshold values of the sigmoid activated model output to be classified as mown.

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.

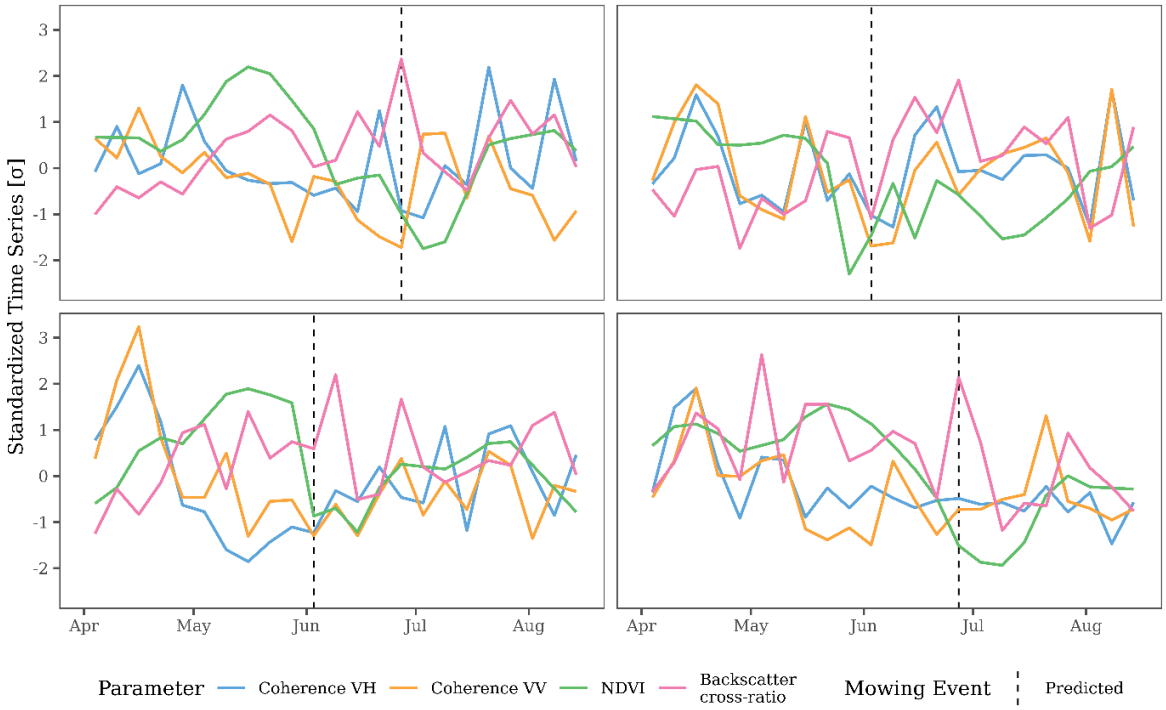

Fig. 4: Time series and predictions made by the model for four exemplary fallows in Brandenburg in 2019 with predicted mowing before the 15th of July with a threshold of 0.93.

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.