PFG 2013 / 4, 0333–0349
Stuttgart, August 2013

Article

Estimation and Mapping of Carbon Stocks in Riparian Forests by using a Machine Learning Approach with Multiple Geodata

LEONHARD SUCHENWIRTH, MICHAEL FÖRSTER, Berlin, FRIEDERIKE LANG, Freiburg & BIRGIT KLEINSCHMIT, Berlin

Keywords: organic carbon, floodplains, CART, OBIA, linear multiple regression

Summary: Floodplain ecosystems offer valuable carbon sequestration potential. In comparison to other terrestrial ecosystems, riparian forests have a considerably higher storage capacity for organic carbon (C_{org}). However, a scientific foundation for the creation of large-scale maps that show the spatial distribution of C_{org} is still lacking. In this paper we explore a machine learning approach using remote sensing and additional geographic data for an area-wide high-resolution estimation of C_{org} stock distribution and evaluate the relevance of individual geofactors. The research area is the Danube Floodplain National Park in Austria, one of the very few pristine riparian habitats left in Central Europe. Two satellite images (Ikonos and RapidEye), historical and current topographic maps, a digital elevation model (DEM), and mean groundwater level (MGW) were included. We compared classifications of C_{org} stocks in vegetation, soils, and total biomass based on two, three, four, and five classes. The results showed that a spatial model of C_{org} in riparian forests can be generated by using a combination of object-based image analysis (OBIA) and classification and regression trees (CART) algorithm. The complexity of floodplains, where patterns of C_{org} distribution are inherently difficult to define, clearly exacerbated the challenge of achieving high classification accuracy. In assessing the relevance of individual geofactors, we found that remote sensing parameters are more important for the classification of C_{org} in vegetation, whereas parameters from auxiliary geodata, e.g. elevation or historical riverbeds, have more influence for the classification of soil C_{org} stocks. This was also confirmed by a comparative linear multiple regression analysis.

Zusammenfassung: *Schätzung und Kartierung von Kohlenstoffvorräten in Auwäldern mithilfe eines Ansatzes des maschinellen Lernens und verschiedenartigen Geodaten*. Auenökosysteme haben ein hohes Speicherpotenzial für organischen Kohlenstoff (C_{org}), auch im Vergleich zu anderen terrestrischen Ökosystemen. Allerdings fehlt eine wissenschaftliche Grundlage für die Schaffung von großmaßstäbigen Karten, die die räumliche Verteilung des C_{org} zeigen. In diesem Beitrag untersuchen wir einen Ansatz des maschinellen Lernens mittels Fernerkundungs- und zusätzlichen geografischen Daten für eine flächendeckende hochauflösende Abschätzung der C_{org}-Verteilung und bewerten die Relevanz der einzelnen Geofaktoren. Das Untersuchungsgebiet ist der Nationalpark Donau-Auen in Österreich, eines der wenigen unberührten Auenhabitate in Mitteleuropa. Zwei Satellitenbilder (Ikonos und RapidEye), historische und aktuelle topografische Karten, das digitale Geländemodell und Grundwasserdaten wurden einbezogen. Wir verglichen die Klassifizierung des C_{org}-Gehalts in Vegetation, Boden und Gesamtbiomasse in zwei, drei, vier und fünf Klassen. Die Ergebnisse zeigen ein räumliches Modell der C_{org}-Verteilung in Auwäldern mit der Kombination einer objektbasierten Bildanalyse (OBIA) und einem CART (Klassifikations- und Regressionsbaum) -Algorithmus. Die Komplexität der Auen, in denen Muster von C_{org}-Verteilung von Natur aus schwer zu definieren sind, erschwerte es, eine hohe Klassifizierungsgenauigkeit zu erzielen. Bei der Beurteilung der Relevanz einzelner Geofaktoren zeigte sich, dass die Fernerkundungsparameter wichtig für die Klassifizierung von C_{org} in der Vegetation sind, während die Höhe oder die Lage des historischen Flussbetts mehr Einfluss auf die Klassifizierung des C_{org}-Gehalts im Boden haben. Dies wurde auch durch eine vergleichende lineare multiple Regression bestätigt.

DOI: 10.1127/1432-8364/2013/0181 1432-8364/13/0181 $ 4.25

1 Introduction

Floodplain ecosystems offer valuable carbon sequestration potential. Riparian forests have a considerably higher storage capacity for organic carbon (C_{org}) than other terrestrial ecosystems (Cierjacks et al. 2010, Hoffmann et al. 2009, Mitra et al. 2005). Among the different floodplain compartments, it is essential to pay special attention to riparian forest vegetation, but also to soils, which often dominate C_{org} pools (Baritz et al. 2010, Harrison et al. 1995, Hofmann & Anders 1996, Kooch et al. 2012, Lal 2005).

Despite the importance of floodplains for carbon sequestration, a scientific foundation for creating large-scale maps showing the spatial distribution of C_{org} is still lacking. Carbon distribution can be mapped at a global or national level, but regional validation is usually not available (Gibbs et al. 2007, Groombridge & Jenkins 2002, UNEP-WCMC 2008). In particular, there are no maps showing the actual allocation of the C_{org} storage within riparian soils and vegetation at the local or regional level. Various studies have focussed on C_{org} stocks in ecosystems, such as in alder fens (Busse & Gunkel 2002), coastal plain floodplains (Giese et al. 2000), boreal lakes in Ontario (Hazlett et al. 2005) or timber plantations in Scandinavia (Backéus et al. 2005, Cao et al. 2010). In tropical and subtropical wetlands there has been research on mangroves and shrimp farms in Thailand (Matsui et al. 2009), seasonal sequestration in the Okavango delta (Mitsch et al. 2010) and Panama (Grimm et al. 2008). Cierjacks et al. (2011) provided statistical models on the spatial distribution of C_{org} stocks in Danubian floodplain vegetation and soils. Rheinhardt et al. (2012) used indicators based on the distance to river for biomass estimations in a river system in North Carolina. However, these studies rely on data collected by cost-intensive field surveys. For improving the estimation of C_{org}, including larger or less accessible wetland and riparian areas, combined methods of remote sensing, geographic information systems (GIS) and machine learning are promising techniques.

A wide range of remote sensing methods (Farid et al. 2008, Munyati 2000, Ozesmi & Bauer 2002) and in particular object-based image analysis (OBIA) (Kollár et al. 2011, Rokitnicki-Wojcik et al. 2011, Wagner 2009) have been used for mapping of wetland habitats. However, these studies related to the differentiation of vegetation classes and did not focus on the assessment of biomass or C_{org}.

In addition, various remote sensing analyses of C_{org} stocks have been done for non-floodplain habitats, but most of these studies have focused either on C_{org} stocks in soil (Behrens & Scholten 2006, McBratney et al. 2003) or in vegetation (Awaya et al. 2004, Hilker et al. 2008, Olofsson et al. 2008). So far, no studies on the estimation of total C_{org} stocks in riparian forests have been done. And despite advances in remote sensing and geodata analysis, these techniques have not yet been applied to the analysis and estimation of area-wide C_{org} stocks in floodplains.

Goetz et al. (2009) distinguished three approaches for using remote sensing data to map carbon stocks. In the simplest method, the stratify and multiply (SM) approach, e.g. as used by Mayaux et al. (2004) or Suchenwirth et al. (2012), a single value or a range of values is assigned to each class of land cover, vegetation type, or other site characteristic. This approach is limited due to the range of biomass within any given thematic class and the ambiguities concerning the identification of given types. The second approach, combine and assign (CA), extends the SM approach to a wider range of spatial data to improve classifications (Gibbs et al. 2007). It has the advantage of using finer spatial units of aggregation and weighted data layers, but is limited due to the moot representativeness of class values and difficulties in acquiring consistent information as the study area size increases. The third approach, direct remote sensing (DR), uses machine learning techniques and extends satellite measurements directly to maps, i.e., a classification algorithm is trained to develop an optimized set of rules through iterative repeated data analysis (Breiman 2001) for the estimation of biomass and carbon (Baccini et al. 2012). This approach results in continuous values for biomass based on easily understandable rules, such as those described for the Amazon basin (Saatchi et al. 2007), Russian forests (Houghton et al. 2007), or the African continent (Williams et al. 2007).

Suchenwirth et al. (2012) used remote sensing data and a digital elevation model to map carbon densities in a floodplain. They used an OBIA approach to classify vegetation types. The total carbon storage of soils and vegetation was quantified using a Monte-Carlo simulation for all classified vegetation types, and spatial distribution was mapped.

We want to improve this method by including additional data and using a machine learning technique. Due to the complexity of the spatial distribution of C_{org} in the Danube floodplains (Cierjacks et al. 2010, 2011, Suchenwirth et al. 2012), and the amount, variety, and variable consistency of available data, our goal is to establish a machine learning approach for an area-wide modeling of C_{org} stocks. To include remote sensing data and several additional geodata, we chose a classification and regression tree (CART) approach (Breiman et al. 1984, Loh 2011).

The specific aims of this paper are as follows:

(1) to evaluate a machine learning algorithm (CART) for estimating and mapping C_{org} stocks in vegetation (C_{org_veg}), soil (C_{org_soil}) and total biomass (vegetation, soil and deadwood; C_{org_tot}) in riparian forests based on classification accuracies, and (2) to rank the parameters in terms of their ability to predict C_{org}.

2 Materials and Methods

2.1 *Research Area*

The research area has a size of 11.3 km^2 and is situated within the Danube Floodplain National Park (*Nationalpark Donau-Auen*) in Austria (16.66° E, 48.14° N). The national park is located between the Austrian capital Vienna and the Slovak capital Bratislava and stretches along the river Danube for about 36 km (Fig. 1). The river has an average width of about 350 m, and the banks are generally fixed by riprap. Only a few human impacts on the area happened apart from the construction of the *Hubertusdamm* dike in the 19th century to protect areas on the northern riverbank from inundation. In the 1960s, natural forest structures were altered by widespread planting of hybrid poplars (*Populus* x *canadensis*), especially on the southern riverbank. In 1996, the area was declared a national park, and thus commercial enterprises were banned within its precincts. Despite of the mentioned human interventions, the area remains one of the last large pristine riparian habitats in Central Europe and has been recognized by the International Union for Conservation of Nature (IUCN) as a Riverine Wetlands National Park, Category II.

The national park's environmental features include the secondary streams (the Danube river itself is an international waterway), side channels and oxbow lakes, gravel banks, riparian forests and meadows, reed beds and xeric habitats. Within the forests, we can dif-

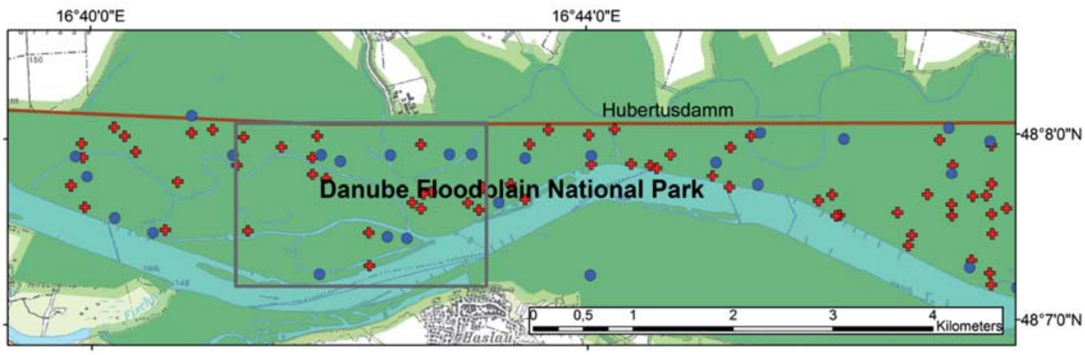

Fig. 1: Research Area, green: Danube Floodplain National Park, red cross: locations of the terrestrial sample points training data, blue dot: test data. The red line represents the Hubertusdamm dike. The grey box represents the outline of the subsets in Fig. 2.

ferentiate between hardwood forest (dominated by *quercus robur*, *fraxinus excelsior* and *acer campestre*), softwood forest (dominated by *salix alba* and *acer negundo*) and cottonwood forest (consisting of hybrid poplar plantations of the 1960ies) (Cierjacks et al. 2010). The main soil type is haplic fluvisol (calcaric). Calcaric gleysols are less important. The climate is continental with a mean annual temperature of 9.8 °C and a mean annual precipitation of 533 mm [Schwechat climate station, 48°07' N, 16°34' E, 184 m above sea level (ZAMG 2002)].

The mean carbon storage in the area was estimated as 359.1 Mg C ha^{-1} (472,186 Mg in an area of 13.1 km^2) by Cierjacks et al. (2010).

2.2 *Data*

The following available comprehensive data from the research area were included in the analysis: two very high spatial resolution (VHSR) satellite images from Ikonos and RapidEye sensor, historical and current topographic maps, a digital elevation model (DEM), and data on the mean groundwater level (MGW).

We purchased a preprocessed cloudfree Ikonos 2 image, recorded on April 22, 2009 with a spatial resolution of 1.0 m (panchromatic) and 4 m (multispectral), as well as a satellite image from RapidEye recorded on August 1, 2009 and processed at L3A with a spatial resolution of 5.0 m (multispectral), provided by the German Aerospace Centre. Both images were provided in the UTM WGS 1984 projected coordinate system and were reprojected into the Austrian MGI M34 projected coordinate system. We used this local system as the majority of local data was also projected in this way.

In addition to the spectral values, several ratios and texture parameters (Haralick et al. 1973) were calculated (Tab. 1). A digital elevation model derived from lidar data was used to compute height and slope. Increased slope values can suggest former riverbeds of the main stream or overgrown side channels, which can serve as an indicator of softwood (Suchenwirth et al. 2012), which cannot be detected directly through spectral values. Also the height above ground has been included in the knowledge-base. The following vegetation types were determined by OBIA from the Ikonos image and the DEM: meadow, reed bed, cottonwood, softwood and hardwood forests (Suchenwirth et al. 2012).

Historical and current topographic maps were provided by the Austrian Federal Office for Metrology and Survey (Österreichisches Bundesamt für Eich- und Vermessungswesen, BEV). The historical maps are derived from three topographic land surveys, the First (1764–1806), the Second (1806–1869) and the Third Military Mapping Survey (1868–1880). We digitized the riverbeds and channels as well as oxbows and coded them, either if there was a historic water body or not. A groundwater model indicating median ground water depth was provided by the Vienna University of Technology.

During two terrestrial surveys in 2008 and 2010, a total of 104 samples from vegetation and soil were taken [69 samples in 2008 (Cierjacks et al. 2010) and 35 samples in 2010 (Rieger et al. 2013), Fig. 1]. All data were collected in a stratified randomized sampling design throughout the research area in 10 x 10 m plots. In each sample plot, forest stand structure was measured and soil samples were taken. A detailed description of the C_{org} calculation is given by Cierjacks et al. (2010) and Rieger et al. (2013). These data were randomly separated in training data (70 %) and test data (30 %) for the classification.

2.3 *Methods*

We developed a spatial model for the estimation and mapping of C_{org} stocks in soils and vegetation based on a machine learning algorithm. For this, we chose a classification and regression tree (CART) approach. CART creates classification rules in the shape of a decision tree. Decision trees show hierarchical rules according to which a dataset is classified. At the beginning of a decision tree is the basic population of the data. During the classification process, the dataset is divided according to binary rules (Breiman et al. 1984, Loh 2011, Quinlan 1986). The advantages of CART include the flexibility to handle a broad

Tab. 1: Available geodata, derived parameters and used abbrevations.

Available geodata	Derived parameters	Abbreviation
Ikonos image (April 22, 2009)	Blue channel	Ikonblu
	Green channel	Ikongrn
	Red channel	Ikonred
	Near infrared channel	Ikonnir
	NDVI (normalized difference vegetation index) (Tucker 1979, Rouse et al. 1973)	Ikonndvi
	Vegetation classification derived by OBIA (Suchenwirth et al. 2012)	Classification
RapidEye image (August 1, 2009)	Blue channel	b1-REblue
	Green channel	b2-REgreen
	Red channel	b3-REred
	RedEdge channel	b4-REredEdge
	Near infrared channel	b5-REnir
	NDVI	RE_NDVI
	Transformed NDVI [$((b5+b3)+0.5)^{0.5}$] (Deering et al. 1975)	tNDVI
	modNDVI [(b5-b4)/(b5+b4-2*b1)] (Datt 1999)	modNDVI
	b4NDVI [(b5-b4)/(b5+b4)] (Gitelson & Merzlyak 1994)	b4NDVI
	Solar Reflectance Index [b5/b3] (Rouse et al. 1973)	b4sri
	[b2-b1]	b2mb1
	[b3-b1]	b3mb1
	[b3-b2]	b3mb2
	[b5-b4]	b5mb4
	[b3/b1]	b3db1
	[b4/b2]	b4db2
	[b5/b2]	b5db2
	Texture parameters (Haralick et al. 1973)	
	Gray-level co-occurrence matrix (GLCM) homogeneity	GLCM_Homogeneity
	GLCM mean	GLCM_Mean
	GLCM correlation	GLCM_Correlation
	GLCM contrast	GLCM_Contrast
	Gray-level difference vector (GLDV) entropy	GLDV_Entropy
Digital elevation model	Elevation	DEM
	Slope	slope
Historical and current topographic maps	Existence of historic riverbed during:	
	First Military Mapping Survey (1773 – 1781)	hist1
	Second Military Mapping Survey (1806 – 1869)	hist2
	Third Military Mapping Survey (1868 – 1880)	hist3
	Current distance to river based on current topographic map ÖK50	dist
Ground water model	Ground water level	MGW
C_{org} ground survey data from 2008 and 2010	Above ground carbon stocks	C_{org_veg}
	Below ground carbon stocks	C_{org_soil}
	Total carbon stocks	C_{org_tot}

range of response types, such as numeric and categorical data, the ease and robustness of construction, and the ease of interpretation (DE'ATH & FABRICIUS 2000).

For our work, we used the software package eCognition 8.7.1. It allowed us to combine CART and OBIA and thus make use of the vast amount of data including remote sensing and other spatially continuous geodata. OBIA has been successfully applied to classifications of diverse habitats from wetlands (KOLLÁR et al. 2011, ROKITNICKI-WOJCIK et al. 2011) and floodplains (WAGNER 2009) to forests (CHUBEY et al. 2006) and drylands (LALIBERTE et al. 2007). The CART approach in eCognition is based on the original algorithms described by BREIMAN et al. (1984) and has been implemented by the OPENCV-WIKI (2010) and eCognition (ECOGNITION 2012).

The ground survey dataset containing total carbon stocks was grouped into classes (Tab. 2) as were the separate stocks for vegetation and soil. We compared classifications of above ground biomass (C_{org_veg}), below ground biomass for soil depth up to 1 m (C_{org_soil}) and total carbon stocks (C_{org_tot}) using classifications based on two, three, four and five quantile classes. We used quantiles in order to have equal numbers of samples for each class. We applied this approach for different numbers of classes to define an optimum number of classes with acceptable classification accuracy.

The OBIA was performed on a multiresolution segmentation with a scale parameter of 200 and the homogeneity criterion including a shape of 0.1 and a compactness of 0.5. Each spectral band of the RapidEye and Ikonos satellite imagery, as well as each additional geodata layer was weighted equally. However, calculated indices or ratios were not further weighted. Equal segmentation settings were used for all classifications in order to facilitate the comparability of area units among the classifications.

The internal CART algorithm was trained with the respective quantile classes and applied onto the parameters using the "classifier" tool in the software package eCognition 8.7.1, with a classifier depth of 10, a minimum sample count of 6 and 9 cross validation folds.

To evaluate the accuracy of the individual classifications, we calculated the overall accuracy (OA). We additionally decided to follow the suggestions of PONTIUS & MILLONES (2011) who recommend the use of allocation and quantity disagreement for accuracy assessment rather than the use of kappa. The two measures are described as follows:

a) Allocation disagreement (AD) is the number of pixels that have a less than optimal spatial allocation in the comparison map with respect to the reference map. Allocation disagreement is the distance above the quantity disagreement line.
b) Quantity disagreement (QD) is the absolute difference between the number of pixels of a certain class in the reference map and the number of pixels of the same class in the comparison map.

The lower the values of allocation and quantity disagreement, the better is the accu-

Tab. 2: C_{org} ranges (Mg C_{org} ha^{-1}) for C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks for different numbers of classes.

class	Five quantile classes			Four quantile classes			Three quantile classes			Two quantile classes		
	C_{org_veg}	C_{org_soil}	C_{org_tot}	C_{org_veg}	C_{org_soil}	C_{org_tot}	C_{org_veg}	C_{org_soil}	C_{org_tot}	C_{org_veg}	C_{org_soil}	C_{org_tot}
1	< 55.0	<132.8	<231.0	< 75.0	<140.0	<255.5	< 86.5	<161.0	<281.0	<134.9	<186.4	<325.9
2	55.0 - 99.9	132.8 - 173.9	231.0 - 300.0	75.0 -135.0	140.0 - 186.5	255.5 - 326.9	86.5 - 180.0	161.0 - 203.2	281.0 - 373.0	>135.0	>186.5	>326.0
3	100.0 -134.0	174.0 - 197.3	300.1 - 360.9	135.1 - 200.0	186.5 - 227.0	327.0 - 407.0	>180.0	>203.2	>373.0			
4	134.1 - 193.0	197.4 - 240.0	361.0 - 445.0	>200.0	>227.0	>407.0						
5	>193.0	>240.0	>445.0									

racy. Both disagreement values are calculated as percentages.

Furthermore, we calculated for each classification the root-mean-square error (RMSE), frequently used to check the internal model quality with the advantage of being independent of the number of used classes (Kanevski et al. 2009, Richter et al. 2012). For our application, we used the arithmetic mean of each class (of the training plots) as the estimated value, and used the terrestrial value of each test plot as the measured value.

To calculate the relevance of the individual datasets, we summarized the use frequency of the individual parameters, normalized by the overall sum of all use frequencies. Additionally, we considered how many parameters derived from a specific dataset were applied, normalized by the total number of the available parameters of a certain dataset. Erasmi et al. (2013) described the concept as "normalized importance".

3 Results

3.1 *Modelled C_{org} Distribution and Accuracies*

Fig. 2 shows the classification results in the form of maps for a part of the research area. The subset comprises all classes and all environmental features inside the research area. We can see that C_{org_veg} stocks are equally scattered across the area, while C_{org_soil} stocks increase as the distance to the river increases.

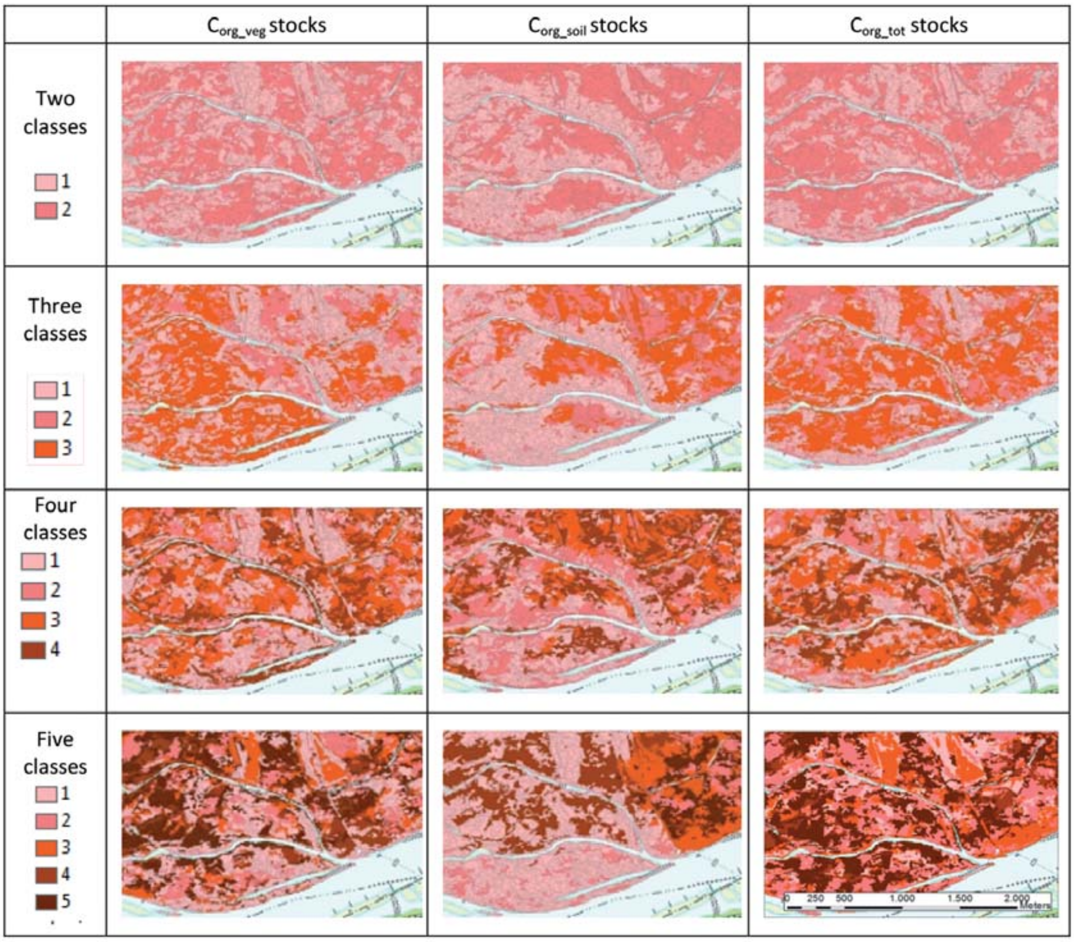

Fig. 2: Modelled distribution of C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks for different numbers of classes. The increasing amount of stored C_{org} is represented by colour graduations increases from pink to red to brown.

The influence is less visible for C_{org_tot} but can still be seen for a classification with four classes.

We compared the derived accuracies (OA, AD, QD) for C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks for all numbers of classes (Fig. 3), as well as RMSE. The comparison of classification accuracies for C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks revealed that the accuracy is highest for two classes and lowest for five classes (Fig. 3). Models with three or four classes range in between and represent a good compromise between complexity and acceptable accuracy.

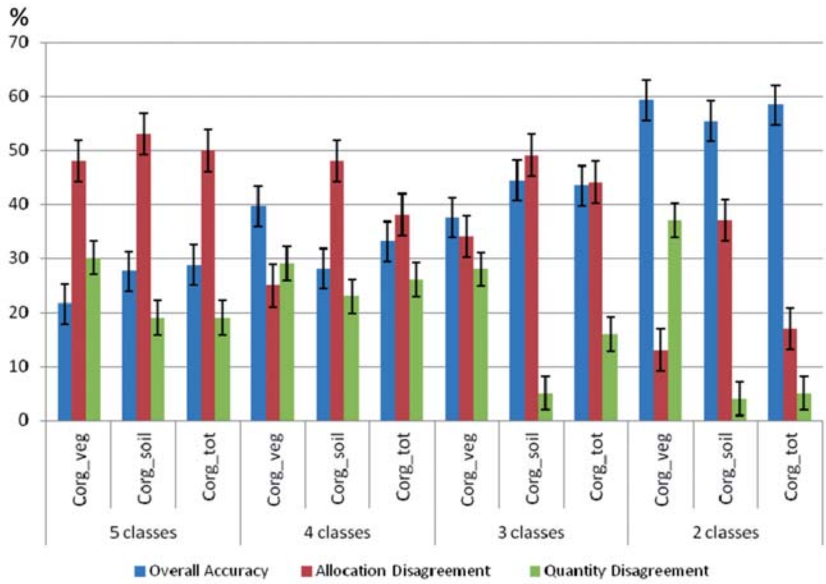

Fig. 3: Overall accuracy, allocation, and quantity disagreement in percent for classifications of C_{org_veg}, C_{org_soil}, C_{org_tot} based on five, four, three, and two classes.

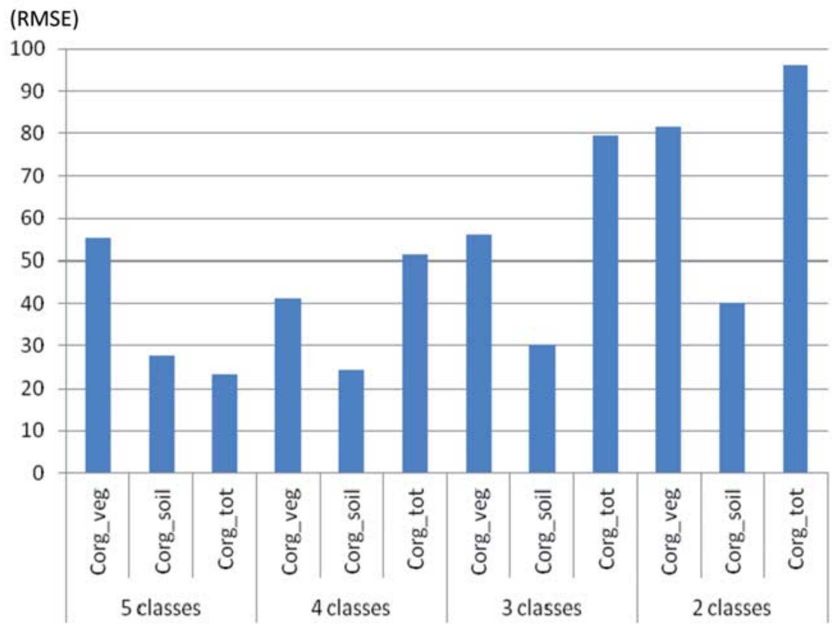

Fig. 4: Root-mean-square error for classifications of C_{org_veg}, C_{org_soil}, C_{org_tot} based on five, four, three, and two classes.

With regard to the model quality, we can examine Fig. 4. Classifications with fewer classes show higher RMSE values, e.g. more than 90 for C_{org_tot} two quantile classes, than classifications with more classes. The lowest RMSE values are below 25 for C_{org_soil} with four classes and C_{org_tot} with four classes.

3.2 *Parameter Relevance*

In the following we analyze the use frequency of the individual datasets and parameters. Tab. 3 shows the results for classifications with all quantile classes for C_{org_veg}, C_{org_soil} and C_{org_tot}.

For RapidEye parameters, the relevance ranged from 3.6 % (C_{org_soil} two classes) to 25.6 % (C_{org_tot} five classes). As the number of classes grows, the parameter relevance rises. For texture parameters, the relevance ranged from 4.6 % (C_{org_soil} 5 classes) to 29.5 % (C_{org_veg} four classes) with no clear indication of which number of classes provided the best results. The overall parameter relevance for Ikonos was lower. It ranged from 0 % (C_{org_soil} two or three classes) to 9.6 % (C_{org_veg} two classes) which could be explained by the acquisition date of April, when full leaf-out had not occurred yet.

For DEM parameters relevance ranged from 0 % (C_{org_veg} three classes; C_{org_soil} five classes) to the highest overall share of 52.1 % (C_{org_soil} two classes). The MGW reached the highest parameter relevance for all classification runs (32.7 % / 18.3 % / 26.2 %), with the relevance ranging from 0 % (C_{org_soil} two and four classes; C_{org_tot} five classes) to 43.2 % (C_{org_tot} two classes). For the "distance to river" parameter, the relevance ranged from 0 % (C_{org_soil} two and four classes) to 50.4 % (C_{org_soil} five classes), with this parameter achieving greater relevance when greater numbers of classes are used. For the parameters based on the existence of historical riverbeds, the relevance ranged from 0 % (C_{org_veg} two, three and four classes; C_{org_soil} five classes; C_{org_tot} two, four and five classes) to 36.0 % (C_{org_soil} two classes), and was important only when classifying C_{org_soil} classes.

To illustrate the importance of single parameters, Figs 5a–c give an exemplary insight of the parameter relevance of classifications with four classes for C_{org_veg}, C_{org_soil}, and C_{org_tot}. For C_{org_veg}, there are 16 parameters (RapidEye: 6; texture: 4; Ikonos: 2; DEM: 2; MGW and distance: 1 each), where the index *b4db2* (i.e. RapidEye's RedEdge divided by green channel) is the most important with more than 23 %. For C_{org_soil}, there are eleven

Tab. 3: Dataset relevance for classifications of C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks.

		RapidEye	Texture	Ikonos	DEM	MGW	Distance to river	Historic maps
C_{org_veg}	5cl	14.5	22.5	5.5	6.3	16.5	31.7	3.0
	4cl	12.0	12.9	5.0	25.3	37.0	7.8	0.0
	3cl	21.8	20.6	3.0	0.0	34.3	20.2	0.0
	2cl	5.9	23.8	9.6	7.3	42.8	10.7	0.0
	Average	13.5	20.0	5.8	9.8	32.7	17.6	0.7
C_{org_soil}	5cl	4.1	4.6	1.7	0.0	39.2	50.4	0.0
	4cl	13.1	29.5	8.4	13.0	0.0	0.0	36.0
	3cl	5.2	18.4	0.0	6.6	33.9	16.6	19.3
	2cl	3.6	8.4	0.0	52.1	0.0	0.0	35.8
	Average	6.5	15.2	2.5	17.9	18.3	16.7	22.8
C_{org_tot}	5cl	25.6	9.8	5.0	11.6	0.0	48.0	0.0
	4cl	4.3	20.8	5.1	13.5	34.5	21.8	0.0
	3cl	9.8	7.6	8.2	8.4	27.0	35.9	3.0
	2cl	9.4	19.7	5.5	22.2	43.2	0.0	0.0
	Average	12.3	14.5	6.0	13.9	26.2	26.4	0.7

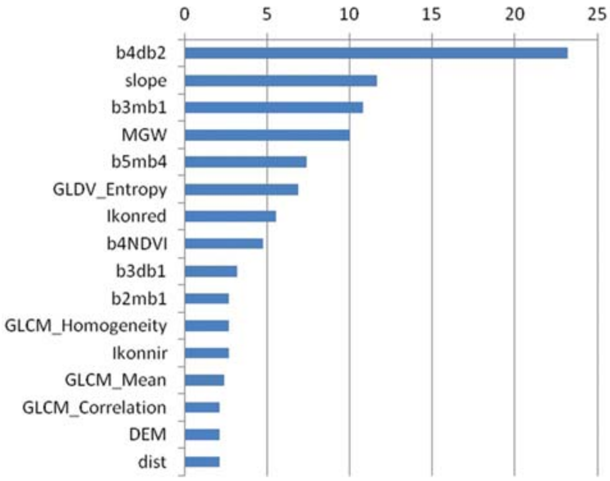

Fig. 5a: Parameter relevance for C_{org_yet} classifications based on 4 quantile classes (all abbreviations are explained in Tab. 1).

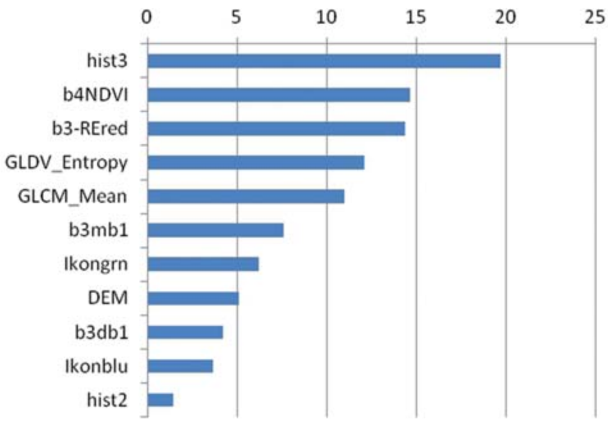

Fig. 5b: Parameter relevance for C_{org_soil} classifications based on 4 quantile classes (all abbreviations are explained in Tab. 1).

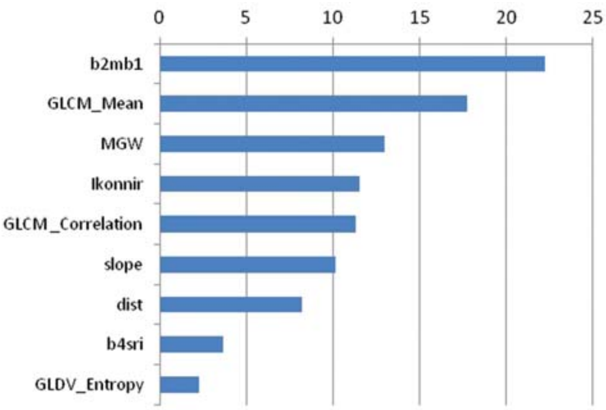

Fig. 5c: Parameter relevance for C_{org_tot} classifications based on 4 quantile classes (all abbreviations are explained in Tab. 1).

parameters (RapidEye: 4; texture: 2; Ikonos: 2; historical maps: 2; DEM: 1), of which *hist3* (existence of riverbed between 1868 to 1880) is the most relevant with almost 20 %. For C_{org_tot}, there are in total nine parameters (RapidEye: 2; texture: 3; Ikonos: 1; MGW, DEM and distance: 1 each), of which *b2mb1* (RapidEye's green channel minus blue channel) is the most important one with more than 22 %.

4 Discussion

4.1 *Classification Results and Accuracies*

Our study provides a novel technique for the estimation and mapping of C_{org} stocks in floodplains based on remote sensing and additional geodata. It could be used to generate C_{org} inventories in other temperate wetlands, especially forested floodplains where ground assessment is difficult or impossible. The visualization of the individual classes shows complex distribution patterns of C_{org} stocks. Despite of the cluttered structure and the heterogeneous distribution within the different classes, the majority of classifications show that higher C_{org_soil} stocks have developed at a certain distance to the main riverbed of the Danube and its side arms. This is best illustrated by classifications with two but also four classes of C_{org_soil}. These lateral gradients were also described by Cierjacks et al. (2010, 2011). In comparison, the patterns of C_{org_veg} and C_{org_tot} were less predictable. Classifications are very speckled for every model and a fully consistent classification is difficult due to the type of the terrain. This reflects the complexity of floodplain habitats in general, and the detailed intricacy of riparian C_{org} stocks in particular and also has been shown by Samaritani et al. (2011) and Suchenwirth et al. (2012). For the particular case of the Danube floodplain, this may also be related to the widespread planting of hybrid poplars in the 1960s, which altered the natural vegetation structure of hardwood and softwood forests.

Surprisingly, the accuracy of the C_{org_soil} stock models was similar to the accuracy of the C_{org_veg} stock models. Predictive variables derived from remote sensing and other geodata serve as proxies for recent environmental conditions that control vegetation properties. Soil organic matter, in contrast, can accumulate over hundreds of years. Thus relations of C_{org_soil} stocks to recent environmental conditions might not be expected. It is likely that the variations in C_{org_soil} stocks found in our study are mainly due to variations in the C_{org} stocks of the upper soil horizons, which in turn are affected by recent environmental conditions. Furthermore, the position of historic riverbeds, a parameter with strong and long-lasting influence on soil organic matter content, was considered (Figs. 3 and 5b).

Predictably, an increase in the number of classes goes along with a more speckled appearance of the classification and overall accuracy decreases. Here, we have to keep in mind that a classification with fewer classes will automatically result in higher accuracy, and therefore the differences simply reflect the higher chance of misclassifications.

Similarly to the overall accuracy, allocation disagreement as well as quantity disagreement values decreased, i.e., the accuracy improved, with fewer classes. An exception is the very high quantity disagreement value for C_{org_veg} based on two classes.

The RMSEs (Fig. 4) provides a measure independent of the number of used classes. The RMSEs "mirror" the results of accuracy assessment, with lower RMSEs for classifications with higher class numbers. Especially for C_{org_soil} accuracies.

For assessing the performance of the CART approach we also compared our results with a linear multiple regression analysis for estimating C_{org_soil}, C_{org_veg}, and C_{org_tot}. Results showed that for C_{org_soil} regression (model intercept p = 0.0069; F = 3.3789) groundwater level was the most important parameter (p = 0.0177; y = -11.275x + 1833.4; R^2 = 0.8657).

For C_{org_tot} regression (model intercept p = 2.3833^{-9}; F = 6.5114), the green RapidEye channel (p = 0.0145; y = -0.0756x + 584.28; R^2 = 0.5619) and the red Ikonos channel (p = 0.0188; y = -0.4198x + 426.33; R^2 = 0.5244) were the most important parameters.

For C_{org_veg} regression (model intercept p = 1.7728^{-6}; F= 7.7927), the green RapidEye channel (p = 0.0099; y = -0.0482x + 335.83; R^2 = 0.5301) and red Ikonos channel (p = 0.0081;

y = -0.3752x + 208.54; R^2 = 0.5562) have the highest importance among the parameters.

The regression confirms our findings that remote sensing parameters are more important for the classification of C_{org_veg}, whereas parameters from auxiliary geodata have more influence on the classification of C_{org_soil} stocks.

It is worth discussing whether and which other additional parameters should be taken into consideration for the detection and modelling of C_{org} distributions in floodplains. Data on forest management practices or local sinks may be considered but were not available on a spatially inclusive and comprehensive level.

In general, Rocchini et al. (2013) argue that the classification of remotely sensed images for the derivation of ecosystem-related maps which also includes the estimation of C_{org} is commonly based on clustering of spatial entities within a spectral space with the implication that it is possible to divide the gradual variability of the Earth's surface into a finite number of discrete non-overlapping classes, which are exhaustively defined and mutually exclusive. Given the continuous nature of many ecosystem properties this approach is often inappropriate; especially as standard data processing and image classification methods involve the loss of information as continuous quantitative spectral information is being degraded into a set of discrete classes. For wetlands, Ozesmi & Bauer (2002) pointed out the limitations of remote sensing for classification and suggest the use of multi-temporal data for an improvement of classification accuracy. For remote sensing in wetlands, Adam et al. (2010) attribute the frequently observed limitations to the low spatial and spectral resolution in comparison to narrow vegetation units that characterize wetland ecosystems.

There may also be concerns about the reliability of terrestrial data. Error propagation may always be a source of uncertainty for the mapping of ecosystems (Rocchini et al. 2013). Our basic survey data have been collected very densely and thoroughly, but transferability to other terrains may become challenging.

Overall, we can conclude that the detection of floodplain characteristics is a challenging task. As for the appropriate number of classes, we consider three or four to be optimal. The accuracy is higher in comparison to a model with five classes, but the complexity is better represented than in a plain dichotomy of data and space created by merely two classes. Dillabaugh & King (2008) found an optimal number of three classes for their classifications of biomass in riparian marshes in Ontario.

Regarding our first research aim, a model approach with four classes seems to perform best. However, the concept of applying segregative classes remains to a certain extent debatable. Therefore, an approach with classes based on fuzzy logic (Zadeh 1989) should be considered in future works to improve the predictive capability of the C_{org} model.

A general point of criticism might apply to the question of why to classify a continuous variable with separate classes. Even though a continuous regression may seem more appropriate, we wanted to create statistically set classes and to follow the concept of different C_{org} concentrations in different compartments of the floodplains. For further planning applications, the regional managers would always apply an ordinal scale, e.g. high, medium, low. The provision of an estimate about the optimal class size for C_{org} might be valuable in terms of its practical application.

A further point of debate remains the sampling design. The random division of terrestrial survey data into 70 % training data and 30 % test data and repeated analysis would probably provide a better estimate about the uncertainties within the calibration and validation data. Repeated measurements could give an insight into the quality of the cal/val information and, in consequence, provide knowledge about the optimal sampling size and spatial distribution of these data. In further analysing steps a repeated calculation with varying samples is envisaged.

4.2 *Use of Parameters*

Regarding the application of parameters and their use frequency, classification of C_{org_veg} relied to a higher percentage on remotely sensed parameters like RapidEye, Texture, and Ikonos than did the classification of C_{org_soil} or C_{org_tot} stocks.

The fact that remotely sensed parameters, especially RapidEye parameters, are the most important factors for the classification of C_{org_veg} provides further evidence of the relevance of satellite imagery for the estimation of biomass, including C_{org} (Gibbs et al. 2007, Neeff et al. 2005, Rheinhardt et al. 2012). Schuster et al. (2012) in particular proved the special relevance of the RedEdge channel for vegetation classification. It is nevertheless remarkable that MGW and the distance to the river played a more dominant role in the classification of C_{org_veg} and C_{org_tot} stocks than C_{org_soil} stocks, although one could assume that median groundwater would be a comparatively less decisive factor for vegetation than for soil biomass and resulting C_{org}. Still, similar findings for fine-root and above-ground biomass which also clearly reflected groundwater depths in the same study area support our results (Rieger et al. 2013). For the case of distance to river, the differences within the parameter relevance (Fig. 5b) for C_{org_soil} is a specific characteristic and shows the variability of classification models. While remotely sensed parameters play the dominant role in all classifications, it is striking that the most important parameter for the C_{org_soil} classification are the historical riverbeds (Figs. 5a–c).

The case is different for the classification of C_{org_soil} stocks, where remote sensing based rules had in some cases less than 50 % influence towards the classification. In contrast, the application frequency of DEM and historical riverbeds – parameters not derived from remote sensing – was more common for the classifications of C_{org_soil} compared to C_{org_veg}. These parameters have already been used successfully in other studies (Cierjacks et al. 2011, Samaritani et al. 2011) to determine C_{org} stocks. Concerning the use of historical maps, it should be kept in mind that our maps only provide information on roughly the last 250 years, whereas C_{org} stocks in soil are the consequence of geomorphologic and pedogenetic processes that have taken place over centuries and millennia.

In general, the assessment of the relevance of individual parameters for the C_{org} model showed that spectral information from remote sensing provides direct information about above ground biomass, while information on soil characteristics can only be explained indirectly through vegetation. This is due to the fact that C_{org_soil} reflects not only recent vegetation, but accumulations over centuries. This is reflected in the high relevance of historical maps for this factor (Fig. 5b) which emphasizes the potential of soils to serve as a memory of previous site conditions, such as historical inundations and changes in riverbeds that often occurred prior to present-day land management practices.

5 Conclusion and Outlook

Our study provides a machine learning approach to model C_{org} stock distributions in riparian forests. We aimed to evaluate a machine learning algorithm (CART) and determine the relevance of individual variables derived from the geodata for the estimation.

Overall, a spatial model of C_{org} in riparian forests could be generated using CART. With the use of geographic datasets, it was possible to show the spatial distribution in terms of a cartographic representation generated by classification. Yet, classification accuracy remains a challenge due to the high complexity of floodplains where patterns of C_{org} distribution are inherently difficult to define.

The evaluation of the relevance of the individual parameters derived from the geodata revealed that remote sensing parameters are more important for the classification of C_{org_veg}, than for the classification of C_{org_soil}. This is also the case for MGW and the distance to the river. In contrast, parameters derived from auxiliary geodata such as DEM and historical maps were more decisive for the classification of C_{org_soil} than C_{org_veg}. C_{org_tot} stocks fell in between in terms of application frequency of remote sensing and other parameters. Therefore, depending on the target (C_{org_soil} or C_{org_veg}), different parameters should be considered when analyzing the spatial distribution of carbon storage.

The application of data-mining approaches to remote sensing and other geodata is helping to automate and facilitate estimations of C_{org} in riparian forests. In addition, information on vegetation structure might improve the C_{org_soil} model. Each classification model highlights the complex interrelations between C_{org} stocks and the external geofactors. In particular, vegetation cover and resulting C_{org_veg} seems to reflect recent site conditions while C_{org_soil} reflects both recent conditions and past processes. In this way, our model contributes to a better understanding of the importance and relationships of C_{org} cycling in floodplain ecosystems. Consequently, this work may serve as a local case study for a well and densely-surveyed area and contribute to improve methods of C_{org} estimation and monitoring in other floodplain areas with similar conditions in temperate climates. It might help to improve formal frameworks such as European biomass inventory (Gallaun et al. 2010), REDD, and Kyoto protocols (Böttcher et al. 2009, IPCC 2000, Obersteiner et al. 2009, Paoli et al. 2010, UNEP-WCMC 2008).

Acknowledgements

This study was funded by the German Research Foundation (DFG; project number KL 2215/2-2). We acknowledge the DLR for the RapidEye image as part of the RapidEye Science Archive – proposal 454. We would like to thank the administrators of the Danube Floodplain National Park for the provision of data, the Austrian Forest Agency (ÖBf) for the provision of forest inventory data, and the TU Vienna for the provision of a ground-water model. We would like to thank Dr. Arne Cierjacks and Isaak Rieger for the provision of terrestrial survey data. We would like to thank Kelaine Vargas for improving the linguistic quality of the English text.

References

Adam, E., Mutanga, O. & Rugege, D., 2010: Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. – Wetlands Ecology and Management **18** (3): 281–296.

Awaya, Y., Tsuyuki, S., Kodani, E. & Takao, G., 2004: Potential of Woody Carbon Stock Estimation Using High Spatial Resolution Imagery: A Case Study of Spruce Stands. – Shiyomi, M., Kawahata, H., Koizumi, H., Tsuda, A. & Awaya, Y. (eds.): Global Environmental Change in the Ocean and on Land: 425–440, Terrapub, Tokyo, Japan.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S. & Houghton, R.A., 2012: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. – Nature Climate Change **2** (3): 182–185.

Backéus, S., Wikström, P. & Lämås, T., 2005: A model for regional analysis of carbon sequestration and timber production. – Forest Ecology and Management **216** (2005): 28–40.

Baritz, R., Seufert, G., Montanarella, L. & Van Ranst, E., 2010: Carbon concentrations and stocks in forest soils of Europe. – Forest Ecology and Management **260** (3): 262–277.

Behrens, T. & Scholten, T., 2006: Digital soil mapping in Germany – a review. – Journal of Plant Nutrition and Soil Science **169** (3): 434–443.

Böttcher, H., Eisbrenner, K., Fritz, S., Kindermann, G., Kraxner, F., McCallum, I. & Obersteiner, M., 2009: An assessment of monitoring requirements and costs of 'Reduced Emissions from Deforestation and Degradation'. – Carbon Balance and Management **4** (1): 7.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J., 1984: Classification and regression trees. – Wadsworth & Brooks / Cole Advanced Books & Software, Monterey, CA, USA.

Breiman, L., 2001: Random Forests. – Machine Learning **45** (1): 5–32.

Busse, L.B. & Gunkel, G., 2002: Riparian alder fens – source or sink for nutrients and dissolved organic carbon? 2. Major sources and sinks. – Limnologica – Ecology and Management of Inland Waters **32** (1): 44–53.

Cao, T., Valsta, L. & Mäkelä, A., 2010: A comparison of carbon assessment methods for optimizing timber production and carbon sequestration in Scots pine stands. – Forest Ecology and Management **260** (10): 1726–1734.

Chubey, M.S., Franklin, S.E. & Wulder, M.A., 2006: Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. – Photogrammetric Engineering and Remote Sensing **72** (4): 383–394.

Cierjacks, A., Kleinschmit, B., Babinsky, M., Kleinschroth, F., Markert, A., Menzel, M., Ziechmann, U., Schiller, T., Graf, M. & Lang, F., 2010: Carbon stocks of soil and vegetation on Danubian floodplains. – Journal of Plant Nutrition and Soil Science **173** (5): 644–653.

Cierjacks, A., Kleinschmit, B., Kowarik, I., Graf, M. & Lang, F., 2011: Organic matter distribution in floodplains can be predicted using spatial and vegetation structure data. – River Research and Applications **27:** 1048–1057.

DATT, B., 1999: A New Reflectance Index for Remote Sensing of Chlorophyll Content in Higher Plants: Tests using Eucalyptus Leaves. – Journal of Plant Physiology **154** (1): 30–36.

DE'ATH, G. & FABRICIUS, K.E., 2000: Classification and regression trees: a powerful yet simple technique for ecological data analysis. – Ecology **81** (11): 3178–3192.

DEERING, D.W., ROUSE, J.W., HAAS, R.H. & SCHELL, J.A., 1975: Measuring "forage production" of grazing units from Landsat MSS data. – 10th International Symposium Remote Sensing of Environment **II:** 1169–1178.

DILLABAUGH, K.A. & KING, D.J., 2008: Riparian marshland composition and biomass mapping using Ikonos imagery. – Canadian Journal of Remote Sensing **34** (2): 143–158.

ECOGNITION, 2012: eCognition Developer Reference Book 8.8. – Trimble Germany GmbH, Munich.

ERASMI, S., RIEMBAUER, G. & WESTPHAL, C., 2013: Mapping habitat diversity from multi-temporal RapidEye and RADARSAT-2 data in Brandenburg, Germany. – BORG, E., DAEDELOW, H. & JOHNSON, R. (eds): 5th RESA Workshop, Neustrelitz, March **2013:** 75–89, GITO Berlin.

FARID, A., GOODRICH, D.C., BRYANT, R. & SOROOSHIAN, S., 2008: Using airborne lidar to predict Leaf Area Index in cottonwood trees and refine riparian water-use estimates. – Journal of Arid Environments **72** (1): 1–15.

GALLAUN, H., ZANCHI, G., NABUURS, G.-J., HENGEVELD, G., SCHARDT, M. & VERKERK, P.J., 2010: EU-wide maps of growing stock and above-ground biomass in forests based on remote sensing and field measurements. – Forest Ecology and Management **260** (3): 252–261.

GIBBS, H.K., BROWN, S., NILES, J.O. & FOLEY, J.A., 2007: Monitoring and estimating tropical forest carbon stocks: making REDD a reality. – Environmental Research Letters **2** (4), doi 10.1088/1748-9326/2/4/045023.

GIESE, L.A., AUST, W.M., TRETTIN, C.C. & KOLKA, R.K., 2000: Spatial and temporal patterns of carbon storage and species richness in three South Carolina coastal plain riparian forests. – Ecological Engineering **15** (Supplement 1): 157–170.

GITELSON, A. & MERZLYAK, M.N., 1994: Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves – Spectral Features and Relation to Chlorophyll Estimation. – Journal of Plant Physiology **143** (3): 286–292.

GOETZ, S., BACCINI, A., LAPORTE, N., JOHNS, T., WALKER, W., KELLNDORFER, J., HOUGHTON, R. & SUN, M., 2009: Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. – Carbon Balance and Management **4** (1): 2.

GRIMM, R., BEHRENS, T., MÄRKER, M. & ELSENBEER, H., 2008: Soil organic carbon concentrations and stocks on Barro Colorado Island – Digital soil mapping using Random Forests analysis. – Geoderma **146** (1–2): 102–113.

GROOMBRIDGE, B. & JENKINS, M.D., 2002: World atlas of biodiversity: Earth's living Resources in the 21st century. – Prepared by UNEP World Monitoring Centre University of California Press, Berkeley, CA, USA.

HARALICK, R.M., SHANMUGAM, K. & DINSTEIN, I.H., 1973: Textural Features for Image Classification. – IEEE Transactions on Systems, Man and Cybernetics **SMC-3** (6): 610–621.

HARRISON, A.F., HOWARD, P.J.A., HOWARD, D.M., HOWARD, D.C. & HORNUNG, M., 1995: Carbon storage in forest soils. – Forestry **68** (4): 335–348.

HAZLETT, P.W., GORDON, A.M., SIBLEY, P.K. & BUTTLE, J.M., 2005: Stand carbon stocks and soil carbon and nitrogen storage for riparian and upland forests of boreal lakes in north-eastern Ontario. – Forest Ecology and Management **219** (1): 56–68.

HILKER, T., COOPS, N.C., WULDER, M.A., BLACK, T.A. & GUY, R.D., 2008: The use of remote sensing in light use efficiency based models of gross primary production: A review of current status and future requirements. – Science of the Total Environment **404** (2–3): 411–423.

HOFFMANN, T., GLATZEL, S. & DIKAU, R., 2009: A carbon storage perspective on alluvial sediment storage in the Rhine catchment. – Geomorphology **108** (1–2): 127–137.

HOFMANN, G. & ANDERS, S., 1996: Waldökosysteme als Quellen und Senken für Kohlenstoff – Fallstudie ostdeutsche Länder. – Beiträge Forstwirtschaft und Landschaftsökologie **30** (1): 9–16.

HOUGHTON, R.A., BUTMAN, D., BUNN, A.G., KRANKINA, O.N., SCHLESINGER, P. & STONE, T.A., 2007: Mapping Russian forest biomass with data from satellites and forest inventories. – Environmental Research Letters **2** (4), doi 10.1088/1748-9326/2/4/045032.

IPCC, 2000: Special report on land use, land-use change and forestry. – Cambridge University Press, Cambridge, UK.

KANEVSKI, M., TIMONIN, V. & POZDNUKHOV, A., 2009: Machine learning algorithms for spatial data analysis and modelling. – EFPL Press, Lausanne, Switzerland.

KOLLÁR, S., VEKERDY, Z. & MÁRKUS, B., 2011: Forest Habitat Change Dynamics in a Riparian Wetland. – Procedia Environmental Sciences **7** (0): 371–376.

Kooch, Y., Hosseini, S.M., Zaccone, C., Jalilvand, H. & Hojjati, S.M., 2012: Soil organic carbon sequestration as affected by afforestation: the Darab Kola forest (north of Iran) case study. – Journal of Environmental Monitoring **14** (9): 2438–2446.

Lal, R., 2005: Forest soils and carbon sequestration. – Forest Ecology and Management **220** (2005): 242–258.

Laliberte, A.S., Rango, A., Herrick, J.E., Fredrickson, E.L. & Burkett, L., 2007: An object-based image analysis approach for determining fractional cover of senescent and green vegetation with digital plot photography. – Journal of Arid Environments **69** (1): 1–14.

Loh, W.-Y., 2011: Classification and regression trees. – Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **1** (1): 14–23.

Matsui, N., Suekuni, J., Nogami, M., Havanond, S. & Salikul, P., 2009: Mangrove rehabilitation dynamics and soil organic carbon changes as a result of full hydraulic restoration and re-grading of a previously intensively managed shrimp pond. – Wetlands Ecology and Management **18** (2): 233–242.

Mayaux, P., Bartholomé, E., Fritz, S. & Belward, A., 2004: A new land-cover map of Africa for the year 2000. – Journal of Biogeography **31** (6): 861–877.

McBratney, A.B., Mendonça, Santos, M.L. & Minasny, B., 2003: On digital soil mapping. – Geoderma **117** (1-2): 3–52.

Mitra, S., Wassmann, R. & Vlek, P.L.G., 2005: An appraisal of global wetland area and its organic carbon stock. – Current Science **88**, Bangalore, India.

Mitsch, W., Nahlik, A., Wolski, P., Bernal, B., Zhang, L. & Ramberg, L., 2010: Tropical wetlands: seasonal hydrologic pulsing, carbon sequestration, and methane emissions. – Wetlands Ecology and Management **18** (5): 573–586.

Munyati, C., 2000: Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. – International Journal of Remote Sensing **21:** 1787–1806.

Neeff, T., de Alencastro Graça, P.M., Dutra, L.V. & da Costa Freitas, C., 2005: Carbon budget estimation in Central Amazonia: Successional forest modeling from remote sensing data. – Remote Sensing of Environment **94** (4): 508–522.

Obersteiner, M., Huettner, M., Kraxner, F., McCallum, I., Aoki, K., Bottcher, H., Fritz, S., Gusti, M., Havlik, P., Kindermann, G., Rametsteiner, E. & Reyers, B., 2009: On fair, effective and efficient REDD mechanism design. – Carbon Balance and Management **4** (1): 11.

Olofsson, P., Lagergren, F., Lindroth, A., Lindström, J., Klemedtsson, L., Kutsch, W. & Eklundh, L., 2008: Towards operational remote sensing of forest carbon balance across Northern Europe. – Biogeosciences **5** (3): 817–832.

OpenCV-Wiki, 2010: Decision Trees. – http://opencv.willowgarage.com/documentation/cpp/ml_decision_trees.html (21.1.2013).

Ozesmi, S.L. & Bauer, M.E., 2002: Satellite remote sensing of wetlands. – Wetlands Ecology and Management **10** (5): 381–402.

Paoli, G., Wells, P., Meijaard, E., Struebig, M., Marshall, A., Obidzinski, K., Tan, A., Rafiastanto, A., Yaap, B., Ferry Slik, J., Morel, A., Perumal, B., Wielaard, N., Husson, S. & D'Arcy, L., 2010: Biodiversity Conservation in the REDD. – Carbon Balance and Management **5** (1): 7.

Pontius, R.G. & Millones, M., 2011: Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. – International Journal of Remote Sensing **32** (15): 4407–4429.

Quinlan, J.R., 1986: Induction of decision trees. – Machine Learning **1** (1): 81–106.

Rheinhardt, R., Brinson, M., Meyer, G. & Miller, K., 2012: Integrating forest biomass and distance from channel to develop an indicator of riparian condition. – Ecological Indicators **23** (0): 46–55.

Richter, K., Atzberger, C., Hank, T.B. & Mauser, W., 2012: Derivation of biophysical variables from Earth observation data: validation and statistical measures. – Journal of Applied Remote Sensing **6** (1): 063557–063551.

Rieger, I., Lang, F., Kleinschmit, B., Kowarik, I. & Cierjacks, A., 2013: Fine root and aboveground carbon stocks in riparian forests: the role of diking, environmental gradients and dominant tree species. – Plant and soil: 1–13, Springer, doi 10.1007/s11104-013-1638-8.

Rocchini, D., Foody, G.M., Nagendra, H., Ricotta, C., Anand, M., He, K.S., Amici, V., Kleinschmit, B., Förster, M., Schmidtlein, S., Feilhauer, H., Ghisla, A., Metz, M. & Neteler, M., 2013: Uncertainty in ecosystem mapping by remote sensing. – Computers & Geosciences **50:** 128–135, Elsevier.

Rokitnicki-Wojcik, D., Wei, A. & Chow-Fraser, P., 2011: Transferability of object-based rule sets for mapping coastal high marsh habitat among different regions in Georgian Bay, Canada. – Wetlands Ecology and Management: 1–14.

Rouse, J.W., Haas, R.H., Schell, J.A. & Deering, D.W., 1973: Monitoring vegetation systems in the Great Plains with ERTS. – Third ERTS Symposium **1973:** 309–317, Washington, DC, USA.

SAATCHI, S.S., HOUGHTON, R.A., DOS SANTOS ALVALÁ, R.C., SOARES, J.V. & YU, Y., 2007: Distribution of aboveground live biomass in the Amazon basin. – Global Change Biology **13** (4): 816–837.

SAMARITANI, E., SHRESTHA, J., FOURNIER, B., FROSSARD, E., GILLET, F., GUENAT, C., NIKLAUS, P.A., PASQUALE, N., TOCKNER, K., MITCHELL, E.A.D. & LUSTER, J., 2011: Heterogeneity of soil carbon pools and fluxes in a channelized and a restored floodplain section (Thur River, Switzerland). – Hydrology and Earth System Sciences **15** (6): 1757–1769.

SCHUSTER, C., FÖRSTER, M. & KLEINSCHMIT, B., 2012: Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. – International Journal of Remote Sensing **33** (17): 5583–5599.

SUCHENWIRTH, L., FÖRSTER, M., CIERJACKS, A., LANG, F. & KLEINSCHMIT, B., 2012: Knowledge-based classification of remote sensing data for the estimation of below- and above-ground organic carbon stocks in riparian forests. – Wetlands Ecology and Management **20** (2): 151–163.

TUCKER, C.J., 1979: Red and photographic infrared linear combinations for monitoring vegetation. – Remote Sensing of Environment **8** (2): 127–150.

UNEP-WCMC, 2008: Carbon and biodiversity: a demonstration atlas. – UNEP-WCMC, Cambridge, UK.

WAGNER, I., 2009: The Danube Floodplain Habitats – application of the Object-based Image Analysis approach. – CAR, A., GRIESEBNER, G. & STROBL, J. (eds): Geospatial Crossroads @ GI_Forum '09: 218–227, Wichmann, Heidelberg.

WILLIAMS, C., HANAN, N., NEFF, J., SCHOLES, R., BERRY, J., DENNING, A.S. & BAKER, D., 2007: Africa and the global carbon cycle. – Carbon Balance and Management **2** (1): 3.

ZADEH, L.A., 1989: Knowledge Representation in Fuzzy Logic. – IEEE Transactions on Knowledge and Data Engineering **1** (1): 89–100.

ZAMG, 2002: Klimadaten von Österreich 1971–2000. – Zentralanstalt für Meterorologie und Geodynamik, Vienna, Austria.

Addresses of the Authors:

LEONHARD SUCHENWIRTH, MICHAEL FÖRSTER & BIRGIT KLEINSCHMIT, Technical University of Berlin, Geoinformation in Environmental Planning Lab, Straße des 17. Juni 145, 10623 Berlin, Germany, e-mail: {leonhard.suchenwirth}{michael.foerster}{birgit.kleinschmit}@tu-berlin.de

FRIEDERIKE LANG, University of Freiburg, Institute of Soil Science and Forest Nutrition, Bertoldstraße 17, 79098 Freiburg, Germany, e-mail: fritzi.lang@bodenkunde.uni-freiburg.de

Manuskript eingereicht: Februar 2013
Angenommen: April 2013