Modeling the Distribution of Organic Carbon Stocks in a Central European Floodplain with VHR Remote Sensing Data and Multiple Geodata

Vorgelegt von

Mag.rer.nat.

Leonhard Suchenwirth

aus Kassel

von der Fakultät VI - Planen Bauen Umwelt

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

Promotionsausschuss:
Vorsitzender: Prof. Dr. Stefan Heiland
Gutachterin: Prof. Dr. Birgit Kleinschmit
Gutachterin: Prof. Dr. Friederike Lang

Tag der wissenschaftlichen Aussprache: 05. Dezember 2013

Berlin 2014

D 83

Eidesstattliche Versicherung

I prepared this dissertation without illegal assistance. This work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. This dissertation has not been presented to any other University for examination, neither in Germany nor in another country.

Leonhard Suchenwirth

Berlin, September 2013

Abstract

In the wake of climate change, the release of CO_2 into the atmosphere caused by deforestation and destruction of natural ecosystems is predominantly in the focus of research. In particular, the determination of the carbon content and its monitoring by remote sensing and Geographic Information Systems (GIS) has been pushed forward in recent years, mainly in tropical areas, and on a small scale. For floodplains and wetlands, which have a very high content of carbon in soil and vegetation, also in temperate climates, the methods for large-scale mapping of carbon have yet been scarce. The main goal of this thesis was to determine the carbon content of a Central European floodplain, using very high resolution satellite data and additional spatial information (digital elevation model, topographic and historic maps, ground water model). Parameters were derived from the different datasets, and used for spatial modeling and compared in their significance. In particular, the remote sensing parameters, but also the additional data were to be analyzed for their importance to the modeling process. Three different approaches were used for modeling and mapping. In a first approach, vegetation types were classified with object-based image analysis, using varying classification rules. To each class, a specific value in vegetation and soil carbon content was assigned; hence, the carbon content of the study area was calculated. In a second approach, quantile classes with high, medium and low carbon content in vegetation, soil and in total were defined. A combined method of object-based image processing and machine learning techniques were used to generate rule sets; the individual parameters were compared and assessed in their importance for the carbon estimation. In a third approach, the performance of two machine-learning approaches (self-organising maps, and k-nearest-neighbor algorithm) with two different data combinations was evaluated. The various approaches differ in their methods, but they are all feasible for carbon assessment in a Central European floodplain. The use of additional spatial information improved the results compared to a pure remote sensing analysis. The methods are applicable for other areas on a comparable data basis and have potential for future applications. The work represents a contribution to the evaluation of floodplain systems and wetland systems in general.

Zusammenfassung

Im Zuge des Klimawandels ist die Freisetzung von CO_2 in die Atmosphäre durch Abholzung und Zerstörung natürlicher Ökosysteme verstärkt in den Fokus der Forschung geraten. Insbesondere die Bestimmung des Kohlenstoffgehalts und dessen Überwachung mittels Geoinformationssystemen (GIS) und Fernerkundung wurde in den letzten Jahren verstärkt vorangetrieben, jedoch meist für tropische Ökosysteme, und im kleinen Maßstab. Für Augebiete und Feuchtgebiete generell, die auch in gemäßigten Klimazonen einen sehr hohen Kohlenstoffgehalt in Boden und Vegetation aufweisen, fehlten bisher die Methoden für großmaßstäbige Bestimmungen und Kartierungen des Kohlenstoffs. Ziel dieser Dissertation war die Bestimmung des Kohlenstoffgehalts eines mitteleuropäischen Augebietes mit Hilfe von sehr hochauflösenden Satellitendaten und zusätzlichen Geodaten (Digitales Geländemodell, topographische und historische Karten, Grundwassermodell). Aus den verschiedenen Datensätzen wurden Parameter abgeleitet, zur räumlichen Modellbildung verwendet und in ihrer Bedeutung verglichen. Besonders die Parameter der Fernerkundung, aber auch der Zusatzinformationen sollten dabei auf ihre Wichtigkeit für den Modellbildungsprozess analysiert werden. Dabei wurden drei verschiedene Ansätze zur Modellierung und Kartierung verwendet. Im ersten Ansatz wurden mittels objektbasierter Bildanalyse Vegetationstypen mit variierenden Klassifikationsregeln klassifiziert. Den Klassen wurden bestimmte Kohlenstoffwerte in Vegetation und Boden zugewiesen, und der entsprechende Kohlenstoffgehalt des Gebietes errechnet. Im zweiten Ansatz wurden Quantilsklassen mit hohem, mittlerem und niedrigem Kohlenstoffgehalt in Vegetation, Boden und in der Gesamtmenge gebildet. Ein kombiniertes Verfahren aus objektbasierter Bildverarbeitung und maschinellen Lernen wurde verwendet; anschließend wurden die einzelnen Parameter in ihrer Bedeutung miteinander verglichen. Im dritten Ansatz wurde die Leistung von zwei Ansätzen maschinellen Lernens (Self-Organising-Maps und k-Nearest-Neighbour), mit zwei verschiedenen Datensatzkombinationen evaluiert. Die einzelnen Ansätze unterscheiden sich in ihrem Vorgehen, jedoch eignen sie sich alle für die Kohlenstoffabschätzung in einem mitteleuropäischen Augebiet. Die Verwendung zusätzlicher Geodaten hat die Ergebnisse im Vergleich zu einer reinen Fernerkundungsanalyse verbessert. Die Methoden sind übertragbar bei einer vergleichbaren Datengrundlage und haben Potential für zukünftige Anwendungen. Die Arbeit stellt einen Beitrag zur Bewertung von Ausystemen dar.

Contents

List of Figures

List of Tables

Abbreviations

AD	Allocation disagreement
BEV	*Bundesamt für Eich- und Vermessungswesen* (Federal Office of Metrology and Surveying; Austria)
BfN	*Bundesamt für Naturschutz* (Federal Office for Nature Conservation; Germany)
BMU	Best matching unit
C	Carbon
CA	Combine and Assign
CART	Classification and Regression Tree
CI	Confidence Intervall
CLDV	Gray-level difference vector
CO_2	Carbon dioxide
Corg	Organic carbon
C_{org_soil}	Soil organic carbon
C_{org_tot}	Total organic carbon
C_{org_veg}	Vegetation organic carbon
CW	Cottonwood forest
DEM	Degital elevation model
DFG	*Deutsche Forschungsgemeinschaft* (German Research Foundation)
DR	Direct remote sensing
GHG	Greenhouse ases
GIS	Geographic Information systems
GLCM	Gray-level co-occurrence matrix
GLDV	Gray-level difference vector
HW	Hardwood forest
IPCC	International Panel on climate change
IUCN	International Union for Conservation of Nature
K	Kappa overall statistic
KIA	Kappa per class
Knn	K-nearest-neighbour
LAI	Leaf Area Index
Laser	Light Amplification by Stimulated Emission of Radiation
Lidar	Light detection and ranging

MGW	Medium ground water
MRV	Monitoring, Reporting and Verification
MW	Mean water level
NDVI	Normalized Differenced Vegetation Index
NGO	Non-Government organisation
NIR	Near infrared
OA	Overall Accuracy
ÖBf	*Österreichische Bundesforste* (Austrian Federal Forest Agency)
OBIA	Object-Based Image Analysis
ÖK	Österreichische Karte
PA	Producer's Accuracy
QD	Qualitiy Disagreement
Radar	Radio Detection and Ranging
RE	RapidEye
REDD	Reduction of Emission from Deforestation and Degradation
RESA	RapidEye Science Archive
RMSE	Root Mean Square Error
SE	Standard Error
SM	Stratify and Mulitply
SOM	Self-Organising Maps
SP	Scale Parameter
SW	Softwood forest
UA	User's Accuracy
UN	United Nations
UTM	Universal Transversal Mercator
VHSR	Very High Spatial Resolution
WGS	World Geodetic System
ZAMG	*Zentralanstalt für Meteorlogie und Geodynamik* (Central Institution for Meteorology and Geodynamics; Austria)

Chapter I: Introduction

The introduction contains citations from the already published papers or submitted manuscripts in chapter II (Suchenwirth et al. 2012), chapter III (Suchenwirth et al. 2013)**, and chapter IV (Suchenwirth et al. submitted)***. Respective sections are marked in grey with subsequent asterisks.*

1 The role of carbon in the context of climate change

In the 20th century, especially in the last 50 years, global temperatures are rising, due to the massive setoffs of carbon into the atmosphere in the form of CO_2 and other greenhouse gases (GHG), such as H_2O, CH_4, N_2O and O_3. The rising temperatures have manifold of consequences, the majority of them negative. Even though there may be positive side effects, for instance areas formerly too cold and forbidding for agriculture will in the future be usable for crop plantations and specialized crops of higher value, and agricultural use may shift towards higher latitude and altitude. However, there will be a plethora of negative side effects. The change of regional and local precipitation and temperature regimes causes droughts, flooding, thunderstorms, heat-waves but also locally extended periods of extended frost and winter. All of these events have a high impact on global and local ecosystems, such as the reduction of biodiversity through the extinction of species, but also on primary sector economy such as fishing, farming, forestry but also the entire economy, health sector and human livelihood (IPCC 2000).

Among the contributors of CO_2 we have to mention human activities including the combustion of fossil fuels (such as coal, wood, petrol, and gases) to generate energy (for instance through traffic, heating or generation of electricity), but especially the deforestation and destruction of other natural habitats. The reasons of deforestation and degradation of ecosystems can be manifold and polycausal.

Deforestation may result from intended lumbering such as traditional transhumance or shifting cultivation in small scale in tropical and subtropical countries, which has been practiced by traditional indigenous societies since millennia. It may even be a sustainable way of agriculture without long-term impacts, if population numbers are not exceeding and the recreation period for forests is long enough. In contrast, modern agriculture uses intensive cultivation methods, e.g. soy plantations in Southern America, or palm oil plantations in South-East-Asia (Gibbs et al. 2010), have severe impacts onto primary rain forests. Also the conversion to crop forest or timber wood plantations with invasive species such as eucalyptus (Liao et al. 2010), or Monterey pines *pinus radiata*; (Wharton and Kriticos 2004; Farley et al. 2004) may cause a intense damage for water balances, soils and natural biodiversity, in tropical and non-tropical environments. While some practices are done with official license and a certain degree of control, deforestation resulting from wrong harvest techniques and exceeding demands of timber- and firewood, such as clear felling, additionally the collection of dead wood, and especially illegal logging practice is a further menace to primary forests. Besides voluntary damages, also unintended damages through pollution may lead to forest decline. Besides human impacts, deforestation (with smoldering long-term reforestation) may also be the outcome of natural disasters such as volcano eruptions, sceletonizing by animals and natural forest fires.

Besides the emissions of GHG, especially CO_2, and the thus triggered climate change, deforestation and degradation processes have severe impacts on ecosystem services such as biodiversity, soil and water balances, and thus the loss of forests can expedite erosion, landslides, droughts, desertification, reduced fertility and the more. Among the most destructive ones are inundations (Opperman et al. 2009). Therefore, the protection of natural forest habitats and the afforestation of cleared areas are strongly envisaged by the international community.

The issue of climate change is focused by a variety of conventions, programs and initiatives, among many others the United Nations Framework Convention on Climate Change (UNFCCC) in 1992, the Kyoto Protocol in 1997. Developed countries that have ratified the Kyoto Protocol commit to limit net GHG and to report their annual greenhouse gas balance according to the rules as elaborated in the IPCC Good Practice Guidance for Land Use, Land-use Change and Forestry (IPCC 2003).

For the sector of carbon sequestration in forestry, the programme REDD (Reducing emissions from deforestation and forest degradation) was established in 2007 at the UN conference in Bali. It shall serve as an instrument of climate protection in order to create attractive compensations for the protection of large-scale forests. Besides tropical rain forests, REDD aims at the protection of e.g. boreal taiga forests in the Northern hemisphere, as they provide the largest continuous forest zones on earth. While the basic idea of REDD is to pay compensations to national states or local organizations for forest protection measures, the overhauled concept "REDD+" additionally includes sustainable forestry management practices as well as the improvement of living conditions and the inclusion of so far unprotected forest areas into the REDD mechanisms. On a political base, REDD and REDD+ are seen as a model in the wake of the UNFCCC. However, sensitive issues and concepts of

REDD and REDD+ are still not defined, e.g. as to which degree a secondary forest or even a forest plantation can be regarded equal as a primary or protected forest. The missing definition in the REDD mechanism implies the high risk that non-sustainable plantations may not be regarded as a deforestation, but as a conventional forest, thus making the agreement futile.

While some non-government organizations (NGOs) support the implementation of REDD as a contribution to the mitigation of deforestation of primary forests, other NGOs criticize REDD and REDD+. In addition to unclear definitions, deforestation is only slowed down but not prevented, and the commercialization of forests inspires a knotty international carbon market with companies benefiting from cheap certificates and restrictions and bypasses for traditional indigenous societies that have traditional sustainable management practices and spiritual attachments to their forests. There is also a debate whether REDD strategies might increase pressure on areas with high biodiversity (Strassburg et al. 2010; Asner et al. 2010).

The successful implementation of policies (such as REDD or REDD+) to reduce GHG emissions caused by deforestation especially in developing countries requires effective and reproducible forest monitoring systems. They should provide consistent results, meet standards for mapping accuracy and shall be implementable from national inventories to pan-tropical levels (Achard et al. 2007).

In developed countries, the need for carbon inventories has been recognized. Yet there are no obligatory regulations on a national level or legal frameworks on C_{org} inventories in forests and specific habitats such as wetlands and floodplain forests. The issue of carbon sequestration is at stake in national forest inventories, e.g. in Germany (BVEL (Bundesministerium für Verbraucherschutz 2005), or in Austria (Federal Ministry of Agriculture 2008) or Canada (Kurz and Apps 2006). Within landscape planning, it is the aim to protect of climate (*Schutzgut Klima*) within the German environmental planning laws to reduce CO_2 emissions, yet there is no specification on the monitoring or mapping of C_{org} stocks (Haaren 2004).

2 Global research on carbon stocks in floodplains

Carbon stocks in forests, especially in tropical rain forests and boreal taiga forests, are of undeniable global importance. Nevertheless, also small-area ecosystems in temperate climates make an important contribution (on a local or regional base) to the sequestration of carbon. Comparing different ecosystems, floodplains (especially including riparian forests and their soils) exhibit a particularly high storage capacity for C_{org} (up to 474 Mg C ha^{-1}; Cierjacks et al. 2010, 2011)

In a comparison on a hectare base, storage capacities of other ecosystems (tropical rainforests in general: 243 Mg C ha^{-1}; temperate forests in general 147 Mg C ha^{-1}; boreal forests: 408 Mg C ha^{-1}; IPCC 2000) are even of minor importance in comparison to floodplains. As a consequence, floodplains, especially with the combination of riparian forests, can be regarded to have a crucial function in the global carbon cycle and climate change.

Besides focusing on C_{org} stocks in vegetation, it is essential to pay attention towards the sequestration in soils, which in many cases dominate C_{org} pools, as shown globally (Kooch et al. 2012; Lal 2005) and within Europe (Baritz et al. 2010; Harrison et al. 1995; Hofmann and Anders 1996). Soil sequestration is higher in terms of quantity, but also longer in terms of time, and crucial for the global climate (Stockmann et al. 2013). Comparing the global storage in vegetation, soil, and deadwood, there is a shift of the importance of organic matter and thus organic carbon in soils increasing from tropical climates to boreal climates, while the importance of organic matter and especially carbon in vegetation decreases from tropical forests to boreal forests.

In general, floodplains are among the most important providers of eco-system services in general, second only to estuaries in terms of value per hectare (Costanza et al. 1997). Natural floodplains are among the most biologically productive and diverse ecosystems on earth. Globally, there are about 2,000,000km² of floodplains, however they are highly threatened (Tockner and Stanford 2002; Ward et al. 2002).

In particular, floodplains are famous for their high biodiversity. They often form a habitat for rare and endangered species, regarding fauna (Bonn et al. 2002; Rothenbücher and Schaefer 2005; Adis and Junk 2002), and flora (Rotach 2004; Vogt et al. 2006). As riparian areas have the natural function to balance inundations after snow melts in spring or heavy rain falls, the landscape was shaped over millennia into old river arms, oxbow lakes, ponds and a high variety of biotopes and habitats. Shiel et al. (1998) investigated the biodiversity in floodplains in Australia, Agostinho et al. (2005) in Brazil and Schindler et al. (2013) in general.

Besides riverine floodplains, there also other wetland ecosystems are well-known for their high rank in ecosystem services (Costanza et al. 1997); the Ramsar convention differentiates between marine wetlands (including coastal lagoons, rocky shores and coral reefs), estuarine (including deltas, tidal marshes and mangrove swamps), lacustrine (i.e. wetlands associated with lakes), and palustrine wetlands (marshes, swamps and peat bogs). There are many studies of these semi-aquatic - semi-terrestrial ecosystems indicating a high sequestration potential for C_{org} (Mitra et al. 2005; Mitsch et al. 2012; Hoffmann et al. 2009). Values range up to 1023 Mg C ha^{-1} in mangrove areas (Donato et al. 2011) and even up to 1450 Mg C ha^{-1} in peat land areas (Parish et al. 2008).

Unfortunately, few floodplain habitats are left in their original state within temperate zones, especially in Central Europe (Krause et al. 2011). They have been (and still are) transformed by human activities since centuries, e.g. through the construction of settlements, conversion to farmland, exaggerated river regulation and stream straitening. The reasons are to be found in the high fertility of riparian soils, a location near the river favorable for human activities such as traffic, industry, commerce and settlements.

Consequently, the conversion of floodplains exposes settlements, infrastructure and agricultural areas at a high menace. There are many historic and current examples, where measures of imprudent site planning and floodplain conversion into other land uses have lead in the case of extreme rainfalls to natural flooding disasters, including human losses and considerable impacts onto local infrastructure and large-scale economies (Mitchell 2003). Additionally to the loss of the protector function from inundations, and the obvious loss of biodiversity, the alteration of floodplains also leads to high releases of carbon (Mitra et al. 2005; Jaramillo et al. 2003).

Floodplains are not only at risk through direct anthropogenic impacts such as land transformations; also the indirect effects of climate change may impose a threat to riparian zones. River discharge quantities may alter with declining precipitations and increasing temperatures; this may modify soil conditions and vegetation growth and cause a decline in the carbon storage capacity of floodplains.

Fortunately, there are endeavours to remediate the negative impacts (Zedler 2003; Kreibich and Thieken 2009; Petrow et al. 2006), first examples of successful dismantling measures to free the river bed again and to provide sufficient space for inundations and to expand natural inundation zones. as well as the planned management of floodplains (Baptist et al. 2004; Stammel et al. 2011; Tockner et al. 1998).

There are highly detailed studies that investigate the function of wetlands as carbon sinks, both on the C_{org} sequestration in specific floodplain habitats, and wetland ecosystems in general. Research has been performed based on the hydrological context with the soil carbon in coastal floodplains in South Carolina (Giese et al. 2000), or as comparison of N and C contents between riparian and upland forests in Ontario (Hazlett et al. 2005). In subtropical and tropical wetlands, there has been for instance research on increased carbon stocks in rehabilitated mangroves at shrimp farms in Thailand (Matsui et al. 2009), while Mitsch et al. (2010) compared carbon dynamics and regional hydrology in Botswana and Costa Rica. Grimm et al. 2008 used a random forest analysis to map the C_{org} potential in soils. Cierjacks et al. (2011) rendered statistical models for the spatial distribution of C_{org} stocks in vegetation and soils of the Danubian floodplain. Rheinhardt et al. (2012) used indicators based on the distance to the river to estimate biomass and C_{org} in a fluvial system in North Carolina.

Yet, these studies rely on data gathered by cost- and labor-intensive terrestrial surveys. Wetland and floodplain areas can be wide-ranging. In most cases they provide only a very limited road and path network, due to the periodic inundations, as well as an increased protection status. Therefore they may be difficult to access for comprehensive terrestrial surveys. To enhance and facilitate the monitoring, particularly for large and/or or less accessible wetland and riparian areas, especially with the aim to estimate C_{org} stocks, combined methods of remote sensing, geographic information systems (GIS) and approaches of machine learning are promising techniques.

There is already a number of studies using various types of remote sensing for the analysis of wetland and floodplain ecosystems, but not in the context of C_{org} monitoring. Lidar (light detection and ranging) data were used to classify the Mackenzie river delta (Mertes 2002), or

riparian cottonwoods in Arizona (Farid et al. 2008). HyMap imaging (spectrometry data) was used to assess and predict biodiversity (Kooistra et al. 2008). Landsat data have been used to classify Kafue floodplains in Zambia (Munyati 2000) and poplars along the Tarim river in Xinjiang, NW China (Thevs et al. 2008).
Object-based image analysis (OBIA) has been applied to map coastal marshes in Georgian Bay (Rokitnicki-Wojcik et al. 2011; Midwood and Chow-Fraser 2010), or the James Bay project (Dissanska et al. 2009) in Canada, the Danube floodplain in Austria (Wagner-Lücker et al. 2013), or mangrove forests in Senegal (Conchedda et al. 2008). The Amazon basin and its adjacent wetlands has been in the special focus, as it represents a huge tropical river system; there are examples from Brazil (Evans et al. 2010; Silva et al. 2010) and Peru (Asner et al. 2010). Several studies give an overall review for analysis methods of floodplains and wetlands based on remote sensing (Adam et al. 2010; Ozesmi and Bauer 2002).

Even though the importance of floodplains for carbon sequestration has been acknowledged and proved in terrestrial studies, and various studies about floodplain ecosystems applied remote sensing, yet there has been few research on C_{org} distributions in floodplains using remote sensing data, especially for large-scale mapping in smaller ecosystems, especially in non-tropical environments. The application of remote sensing, GIS, modelling and data mining techniques for the determination of C_{org} stocks has already a wide range of instruments and experiences.

3 Remote sensing and the estimation of carbon stocks

In order to analyze the carbon storage in forests and ecosystems in general, and trace changes in carbon storage, the MRV (monitoring, reporting, and verification) system is to ensure that guidelines such as REDD are implemented (Obersteiner et al. 2009; Böttcher et al. 2009). MRV systems are based on remotes sensing data and forest inventories.

As mentioned earlier, methods of conventional measurements by terrestrial surveying are limited due to time and financial restrictions, especially for large ecosystems such as tropical rain forests or taiga forests. In order to get information for large areas, terrestrial measurements have to be interpolated and reinforced by advanced measurements of remote sensing, statistics and geoinformation sciences. There have been various research studies, and results have already been applied for publications for a broader public such as atlases, policy papers etc. Patenaude et al. (2005) give a broad overview.

Gibbs et al. (2007) describe several data types to delineate carbon stocks:

-Biome averages (estimating average forest carbon stocks for broad forest categories based on a variety of input data sources),

-Forest inventory (relating ground-based measurements of tree diameters or volume to forest carbon stocks using allometric relationships),

-Optical remote sensors (using visible and infrared wavelengths to measure spectral indices and correlate to ground-based forest carbon measurements)

-Very high spatial resolution (VHSR) airborne optical remote sensors (using aerial photos or 3D digital aerial imagery images) to measure tree height, crown area and allometry to estimate carbon stocks)

-Radar (radio detection and ranging) data to measure forest vertically

-Laser remote sensors (using Lidar to estimate forest height/vertical structure)

Based on these data, Goetz et al. (2009) discern three approaches for the use of remote sensing data (and ancillary geodata), to represent spatial distributions of C_{org} stocks in maps:
In the simplest method, the '**stratify and multiply**' (SM) approach, a single value or a range of values is assigned to each class of vegetation type, land cover, or other site characteristics. This approach has constraints, owing to the fact that within a given thematic class, there is a wide range of biomass and organic carbon and the uncertainties concerning the recognition of given classes such as vegetation or land cover types.

The '**combine and assign'** (CA) approach expands the SM approach to a wider range of geodata to improve classifications. It has the benefits of using finer spatial units of aggregation and applying weighted data layers, as well as the possibility to aggregate values and provide them for specific political jurisdictions. But the approach is limited by the same restrictions as the SM approach, especially due to the debatable representativeness of class values and challenges of obtaining consistent information as the research area size enlarges.

The '**direct remote sensing**' (DR) approach applies techniques of statistics or machine learning and extends satellite measurements directly to maps. It uses the field measurements to train a classification algorithm through iterative repeated data analysis in order to develop an optimal rule set defining an apt combination of satellite observations for the estimation of biomass and carbon and the generation of large-scale maps. Once the rules are optimized for training data, they are applied for the entire satellite image. This approach results in continuous values for biomass based on easily understandable rules, that can be adapted and used for a monitoring framework.

Carbon distribution is mainly mapped at a national or global level (Groombridge and Jenkins 2002; UNEP-WCMC 2008), while regional validation is usually not available. Yet there are examples of how C_{org} in local or regional ecosystems have been estimated successfully, such as applications for the Amazonian rain forests by a radar approach (Neeff et al. 2005), Japanese spruce stands by high-spatial resolution imagery (Awaya et al. 2004), Scandinavian forests (Olofsson et al. 2008; Cao et al. 2010; Backéus et al. 2005) or in the UK using Lidar data (Patenaude et al. 2004) and several areas across North America (Turner et al. 2004). In particular, the use of high spectral resolution data indices such as the Normalized Differenced Vegetation Index (NDVI) or the Leaf Area Index (LAI) have proven to be useful tools to calculate the gross primary production of C_{org} (Hilker et al. 2008).

The detection of C_{org} stocks in soils by remote sensing methods has been described by McBratney et al. (2003) and applied for farming purposes (Wendroth et al. 2003), but the focus lay on bare soils or agricultural lands, not forested soils.

The well-introduced studies (e.g. Baccini et al. 2008, Conchedda et al. 2008, Goetz et al. 2009) of remote sensing of C_{org} stocks have their focus mainly on tropical and subtropical ecosystems or traditional timber forests. In accordance with the concept of carbon monitoring and frameworks such as REDD and MRV and local implementations, it is necessary to conceptualize and implement also methods to model C_{org} stocks in floodplains and riparian forest by means of remote sensing and using additional geodata.

Although the importance of floodplains for carbon sequestration has been acknowledged and proved in field-based studies (Cierjacks et al. 2011, 2012; Mitsch et al. 2012), and there are various remote sensing applications for floodplain ecosystems, several academic voids prevail; there has been few research on modelling C_{org} distributions in floodplains using remote sensing data. In particular, there has been very little research on large scale-maps, for smaller ecosystems in a non-tropical environment. Regarding the specific importance of C_{org} in soils described above, it is necessary to include soils in a comprehensive model along with floodplain vegetation at large scale.

4 Research area

4.1 The Danube Floodplains

Having a length of 2,860 km and a total watershed area of 817,000 km², the Danube river is the second largest river in Europe; it is the river with most adjacent countries in the world. The mean annual discharge into the Black Sea is about 6,500 $m^3 s^{-1}$. Our research has been performed in the Danube Floodplain National Park (German: *Nationalpark Donau-Auen*). This national park is situated in the Austrian states/provinces of Vienna and Lower Austria (political districts of Gänserndorf, Bruck an der Leitha and Wien-Umgebung); nearest cities are the Austrian capital of Vienna, and the Slovak capital Bratislava, both located at the shores of the Danube river (fig. I-1). The park stretches along the river for about 35 km and has a width of about 3 km, with the river width of about 350m.

The national park was established in 1996, after a planned hydropower plant in the Hainburg floodplain could be avoided by environmental NGOs and a growing public in the early 1980ies, including the occupation of the Hainburger Au in 1984, forcing the Austrian government to withdraw their plans. Previous human impacts were the use as an imperial hunting grounds before the 20th century, the construction of a levee or dike in the 19th century in order to protect the agriculturally important Marchfeld area north of the floodplain from inundations (locally known as *Marchfeld-* or *Hubertusdamm*), and the plantation of hybrid poplars (*Populus x canadensis*) in the 1960ies. In the past centuries, the Danube floodplains have been extending more northwards to the actual city of Vienna; there is a long intertwined history of the city and the river (Winiwarter et al. 2013).

For the Lobau (the Viennese section) there have been also several plans to build a bypass road for the city, either in form of a dam (in Napoleonic times) or a tunnel in a current version. After the implementation as National park, the area is mainly used for recreational purposes, while commercial enterprise inside the territory is banned and leaves the area in a natural state.

Haplic Fluvisols and Gleysols, both calcaric, are the main soil types in the area. The climate is continental, with a mean temperature of 9.8° C and mean precipitation of 533 mm [Schwechat climate station, 48°07'N, 16°34'E, 184 m above sea level (ZAMG 2002)]. The flow velocity of the Danube River ranges between 2 and 2.5 m/s with a mean discharge of 1,950 m^3/s.

The national park is characterized by wide range of environmental conditions such as the Danube's water body, various side channels and oxbow lakes, gravel banks on shores and islands, riparian forests, meadows and even xeric habitats (locally known as *Heißlände*). Its mixture of vegetation types includes hardwood forest, softwood forest, cottonwood forest (consisting of the hybrid poplars planted in the 1960ies), reed beds, and meadows. Among the tree species, we can find *Acer pseudoplatanus*, *Acer negundo*, *Acer campestre*, *fraxinus excelsior*, *Alnus incana*, *Alianthus altissima*, *Populus x canadensis*, *Populus nigra*, *Populus alba*, *Aesculus hippocastanum*, *Prunus padus subsp. Padus*, *Ulmus minor*, *Salix fragilis*, *Salix alba*, and *Malus sylvestris*.

In the fauna we can find prominent exponents of increasing biodiversity such as *Aquila heliaca* or *Haliaeetus albicilla*, *Merops apiaster* and *Alcedo atthis* among the birds, *Cricetus cricetus*, *Castor fiber*, *Sus scofra*, *Cervus elaphus*, *Barbastella barbastellus* and *Myotis daubentoni* among mamals, *Pelobates fuscus fuscus*, *Bufo calamita*, *Triturus dobrogicus*, and *Bufo viridis* among amphibians. Among reptiles, there are *Lacerta viridis viridis*, *Emys orbicularis*, *Zamenis longissimus* and *Natrix tessellata tessellata*, among fish species there are *Umbra krameri*, *Acipenser ruthenus*, *Cyprinus carpio carpio* and *Zingel zingel.*

The national park is one of the last large pristine floodplain areas in Central Europe (Krause et al. 2011; Wenger et al. 1990), despite the human interventions in the past. It has been recognized by the IUCN as a Riverine Wetlands National Park, complying with the category

II of the IUCN, meaning that the protected area has similar characteristics to wilderness areas (category I) in terms of size and the main objective of protection, but category II areas are more indulgent towards human visitation and supporting infrastructure. National parks promote moderate recreational and educational tourism, having barriers towards surrounding areas in order to protect native species and communities and their sustainable survival.

The Danube Floodplain National Park already has seen a whole range of research studies, among them studies about vegetation characteristics (Ellenberg 1986; Wagner 2009; Wagner-Lücker et al. 2013), soil characteristics (Lair et al. 2009; Sali-Bazze 1981; Zehetner et al. 2009), carbon stocks (Cierjacks et al. 2010; 2011; Rieger et al. 2013) and other (Tockner et al. 1998; Ward et al. 1999).

4.2 Study area

The study area is located in the center of the Danube Floodplain National Park, close to the town of Orth (16.66° E, 48.4° N). It contains all landscape features of the National park, comprised in a small area. The study site was chosen due to the previous collection of terrestrial data in this area, and has the advantage of being easily accessible in comparison to other areas of the national park. The research area is confined between the Hubertusdamm dike in the North, the river Danube in the south, and the towns of Eckartsau in the east and Schönau in the west (Fig. I-1). The detailed size was adapted during the application of the methods.

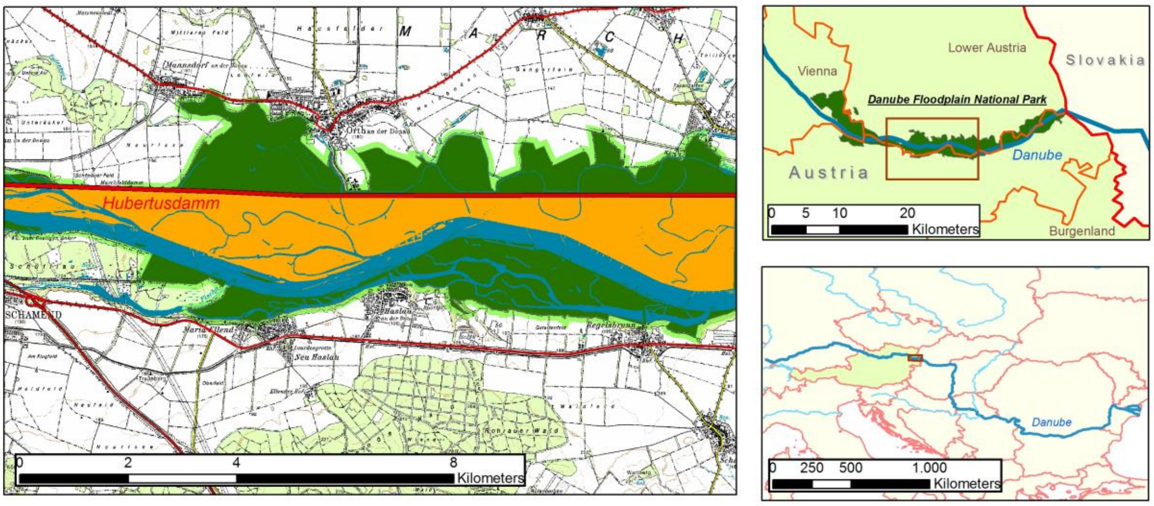

Fig. I-1: Overview map of the study area (orange) inside the floodplain national park (dark green)

Mean carbon storage in the area was estimated as 359.1 Mg C ha^{-1} (472,186 Mg in an area of 13.1 km²) by Cierjacks et al. (2010). Depending on the applied method, the area size is slightly adapted in the single methods.

5 Data framework

In order to estimate carbon stocks in a small research area, it is necessary to make use of large-scale geodata of the research area. We can distinguish between remote sensing data, additional geodata, and ground survey data on C_{org} stocks.

5.1 Remote Sensing data

In order to cover analyze the comparatively small research area, the use of very high spatial resolution (VHSR) satellite imagery is appropriate. The launch of new generations of earth observation satellites provides cheap and reliable imagery. Table I-1 shows the available remote sensing data for the area and time.

Table I-1: Available Satellite data and derived parameters

Data	Derived parameters	Usage
Ikonos (April 22 -2009)	Blue channel (445 -516 nm) Green channel (506-595 nm) Red channel (632-698 nm) Near infra red channel (757-853 nm) NDVI (Normalized Difference Vegetation Index)	Chapter II*, Chapter II**I
RapidEye (August 1-2009)	Blue channel (440 -510 nm) Green channel (520-590 nm) Red channel (630-685 nm) Red edge channel (690-730 nm) -RE Near infra red channel (760-850 nm) -NIR NDVI (Normalized Difference Vegetation Index) Transformed NDVI [((NIR+red)+0.5)$^{0.5}$] modNDVI [(NIR - RE)/ (NIR+RE -2*blue)] b4NDVI [(NIR-RE)/(NIR+RE)] Solar Reflectance Index [NIR/red] [Green - blue] [Red - blue] [Red - green] [NIR -RE] [Red/ blue] [RE/ green] [NIR/ green] Texture parameters : Gray-level co-occurrence matrix (GLCM) homogeneity GLCM mean GLCM correlation GLCM contrast Gray-level difference vector (GLDV) entropy	Chapter III**, Chapter IV***

Through the RapidEye Science Archive (RESA) project (contract 454), free satellite imagery from the sensor RapidEye was provided. It was launched in 2008, financed with public and private support. Due to a bankruptcy in 2011 the ownership shifted to the Canadian company

Iunctus Geomatics, but the cooperation with DLR in the RESA project was secured. The image was recorded on August 1, 2009 in level 3A with a spatial resolution of 5.0 m.

Additionally, commercial imagery was acquired from the satellite systems Ikonos-2. The image consists of a single-band panchromatic band (450 to 900 nm) with a spatial resolution of 1 x 1 m, and four spectral bands with a spatial resolution of 4 x 4 m covering the following intervals: blue, green, red, and near-infrared (Table I-1). The data have been geo-referenced by the provider to the World Geodetic System 1984 (WGS 84) datum, zone 33 N, with Universal Transversal Mercator (UTM) projection with an accuracy of 8 m RMS according to the data service provider (GeoEye 2009).

5.2 Auxiliary geodata

As Austria offers high-quality public data we have a good database for the study (table I-2). The following data was provided by the National Park Administration and other official authorities:

-A digital elevation model (DEM) derived from Lidar data with a spatial resolution of 2.4 x 2.4 m and a vertical resolution of 1 m was provided by the Danube Floodplain National Park administration and has been created by the Institute of Photogrammetry & Remote Sensing of the Vienna University of Technology. The model can be used to indicate height above sea level, river level or groundwater level. Additionally, slopes can be calculated. Increased slope values can be considered evidence of former riverbeds.

-A groundwater model indicating median ground water level was provided by via donau - Österreichische Wasserstraßen-Gesellschaft mbH through Prof. Paul Blaschke from Vienna University of Technology.

-Historical and current topographic maps from military mapping surveys in the 18th and 19th century 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). They provide valuable

information about historic riverbeds and extents of the Danube. The topographic map (*Österreichische Karte 1:50.000, ÖK 50*) is updated every seven years. The riverbeds, side channels and oxbow lakes of each map have been digitized.
We used the provided maps were in the original MGI 34 reference system (based on Bessel 1841 ellipsoid and Gauss-Krüger projection), in the wake of the accession of Austria to the NATO partnership for peace in 1995 all topographic maps are being updated to the WGS 1984 reference system with UTM projection.

-Forest inventory data from 1999 and 2009 were provided by the Austrian Forest agency (*Österreichische Bundesforste*, ÖBf). Forest inventory points are arranged in a grid of 100 x 400m; at each point, all trees within a radius of 8 m (total area of 201m²) are measured and tree species categorized.

Table I-2: Additional geodata and derived parameters

Data	Derived parameters	Usage
Digital elevation model	Elevation Slope	Chapter II*, Chapter III**, Chapter IV***
Historical and current topographic maps	Existence of historic riverbed during: First Military Mapping Survey (1773 - 1781) Second Military Mapping Survey (1806 -1869) Third Military Mapping Survey (1868 - 1880)	Chapter III**
Actual topographic map (ÖK 50) (1993)	Current distance to river based on current topographic map ÖK50	Chapter II*, Chapter III**,
Ground water model (provided in 2010)	Ground water level	Chapter III**, Chapter IV***
Forest inventory (2009) by ÖBf	Vegetation type	Chapter II*

5.3 Ground survey data

In a project by students of Landscape Architecture and Environmental Planning of the Technical University of Berlin, data on carbon stocks in the Danube Floodplain were collected. In order to cover all relevant vegetation types in the study area, polygons with more or less homogeneous vegetation were classified with visual interpretation of aerial color-infrared (CIR) photographs; of the polygons 76 were selected by random, and within each plot, a 10 x 10 m² study plot established. Within each plot, vegetation was classified according to dominant tree species and categorized into softwood forest, hardwood forest, reforestations with hardwood species, meadows and reeds. Data were collected in Summer 2008 and confirmed in 2009. Soil samples ranging from 0 to 100cm in depth were extracted from the centre of each plot using an auger. Horizon thickness and texture was determined, and for each horizon of each sample, carbonate concentration and organic carbon concentration was measured. For vegetation, organic carbon was calculated according to tree parameters such as circumference, tree height and leaf coverage. Dead biomass was calculated accordingly as 50% of living biomass. All biomass was extrapolated to hectare values. The design of the data sampling and the calculations are described in detail by Cierjacks et al. (2010).

Table I-3: C_{org} ground survey data

Available geodata	Derived parameters	Usage
Ground survey data (2008)	Above ground C_{org} stocks Below ground C_{org} stocks Total C_{org} stocks	68 points: Chapter II*, Chapter III**, Chapter IV***
Ground survey data (2010)	Above ground C_{org} stocks Below ground C_{org} stocks Total C_{org} stocks	36 points: Chapter III**, Chapter IV***

In 2010, additional data were collected (Rieger et al. 2013). Even though the collection approach slightly differed from the study project, as 48 tree samples were gathered within six zones in the floodplains, forming the basis for the analysis of carbon stocks. Apart from his data on fine root carbon stocks and vegetation carbon stocks, data on carbon stocks in the soil were collected and calculated. These soil samples enlarged our terrestrial survey data base and were used in the second and third approach.

Of both datasets combined, we could use 104 points for our analyses. We could not use all data, as some points had been sampled outside our research area, e.g. north of the Hubertusdamm dike.

6 Research objectives and structure of this thesis

The first overall objective of this thesis is to develop and evaluate methods to integrate existing very high resolution satellite imagery and additional geodata to a classification process with the aim to model and map the spatial distribution of carbon stocks in floodplain vegetation, soil, and in total on a local scale. This objective is approached by applying different techniques to estimate the carbon stocks ranging from indirect methods to more direct methods (based on the description of 'stratify and multiply', 'combine and assign' and 'direct remote sensing') and to assess the quality of the resulting classifications.

The second overall objective is the analysis of the significance of additional geo-data and knowledge of the classification success. In remote sensing digital image processing the preparation of an accuracy assessment based on validation samples is frequently utilized. Unfortunately, the derived values give no direct information about the contribution of single parts of the original data-set (e.g. spectral bands). Therefore, this objective is addressed by developing and adapting techniques using different classifications sets with different extents of applied data.
Based on these objectives the following major research questions are posed:

Research question 1: *Which remote-sensing based methods can be sufficiently applied to model* C_{org} *stocks in soil and vegetation in floodplains by remote sensing and additional geodata?*

Research question 2: *How can additional geodata be included and their significance for the model be measured?*

Research question 3: *What are the advantages of automated* C_{org} *mapping on local scale for operational monitoring purposes?*

Answers to these questions can be found in the following chapters of this thesis. Most of the research undertaken to answer specific research questions has been published or submitted to peer-reviewed scientific journals as listed in Appendix A. The respective chapters are kept in their original form.

The following provides brief chapter introduction. The headlines give a brief understanding of the essential finding with respect to the research objectives of this thesis. The publication information is given subsequently; the research is related to the overall context and specific research objectives are presented.

Chapters II to IV were written as stand-alone manuscripts to be published in international peer-reviewed journals. Each chapter is therefore structured into the subsections background, study area, methods, results, discussion, and conclusions, thereby resulting in a limited amount of recurring material throughout the thesis.

Chapter II: 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 (2012): 151-163

Leonhard Suchenwirth, Michael Förster, Arne Cierjacks, Friederike Lang, and Birgit Kleinschmit

This approach aims to map the C_{org} stocks based on a classification of vegetation type classes, by using an OBIA approach combined with a Monte Carlo simulation. It compares and discusses the results of several classification sets with varying data quantities. The research aims are:

2.1 To classify Central European floodplain habitats by OBIA.
2.2 To evaluate the use of OBIA to improve the classification accuracy of vegetation cover mapping in Central European floodplain habitats.
2.3 to generate reliable, integrated above- and below-ground carbon stock estimates in floodplains with OBIA by using VHSR remote sensing data and auxiliary data in a SM approach.

Chapter III: Estimation and mapping of carbon stocks in riparian forests by using a machine learning approach with multiple geodata

Photogrammetrie, Fernerkundung und Geoinformation 4 (2012): 333-349

Leonhard Suchenwirth, Michael Förster, Friederike Lang, and Birgit Kleinschmit

This approach aims to map the C_{org} stocks determined in quantile classes for above, below ground and total C_{org} stocks by using an OBIA-based classification and regression tree (CART) approach. Results are analyzed and discussed. The research aims are:

3.1 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.
3.2 to rank the parameters in terms of their ability to predict C_{org}.

Chapter IV: Large-scale mapping of carbon stocks in riparian forests with self-organizing maps and the k-nearest-neighbor algorithm

submitted to iForest- Biogeosciences and Forestry on 28 May 2013; under review

Leonhard Suchenwirth, Wolfgang Stümer, Michael Förster, and Birgit Kleinschmit

This approach uses two machine learning approaches, self-organizing maps, and k-nearest-neighbor. The approach retrieves direct estimations without image segments and classes. It works on the following research aims:

4.1 to create distribution maps of vegetation, soil and total C_{org} stocks in a riparian forest, based on SOM and kNN algorithms and compare the results.

4.2 to compare and evaluate results with previous estimation techniques.

4.3 to evaluate the influence of additional geodata on estimation quality.

Chapter V synthesizes the outcomes of the preceding chapters and provides recommendations for upcoming research in the future.

Chapter II: 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 (2012): 151-163

Leonhard Suchenwirth, Michael Förster, Arne Cierjacks, Friederike Lang, and Birgit Kleinschmit

doi: 10.1007/s11273-012-9252-8

Received: 28 April 2011 / Accepted: 7 February 2012 / Published online: 2 March 2012

Abstract

Floodplain forests play a crucial role in the storage of organic carbon (C_{org}). However, modeling of carbon stocks in these dynamic ecosystems remains inherently difficult. Here, we present the spatial estimation of C_{org} stocks in riparian woody vegetation and soils (to a depth of 1 m) in a Central European floodplain using very high spatial resolution remote sensing data and auxiliary geodata. The research area is the Danube Floodplain National Park in Austria, one of the last remaining wetlands with near-natural vegetation in Central Europe. Different vegetation types within the floodplain show distinct capacities to store C_{org}. We used remote sensing to distinguish the following vegetation types: meadow, reed bed and hardwood, softwood, and cottonwood forests. Spectral and knowledge-based classification was performed with object-based image analysis. Additional knowledge rules included distances to the river, object area, and slope information. Five different classification schemes based on spectral values and additional knowledge rules were compared and validated. Validation data for the classification accuracy were derived from forest inventories and topographical maps. Overall accuracy for vegetation types was higher for a combination of spectral- and knowledge-based classification than for spectral values alone.

While water, reed beds and meadows were clearly detectable, it remained challenging to distinguish the different forest types. The total carbon storage of soils and vegetation was quantified using a Monte Carlo simulation for all classified vegetation types, and the spatial distribution was mapped. The average storage of the study site is 428.9 Mg C ha-1. Despite certain difficulties in vegetation classification this method allows an indirect estimation of C_{org} stocks in Central European floodplains

.

1 Introduction

In the context of climate change, the sequestration of organic carbon (C_{org}) in ecosystems is an increasingly urgent issue (IPCC 2000; Mitra et al. 2005). Soils and vegetation are essential for the storage of C_{org}; although carbon storage capacity and mean residence time are higher for soils, forests in particular can sequester and store significant amounts of C_{org} as woody biomass (Harrison et al. 1995; Köhl et al. 2008). Comparing different terrestrial ecosystems, riparian forests in floodplains exhibit a particularly high storage capacity for C_{org}. Values up to 474 Mg C ha^{-1} of total above- and below-ground C_{org} have been reported from Central European hardwood (HW) forests (Cierjacks et al. 2010, 2011), exceeded only by estuarine mangroves with values up to 1,000 Mg C ha^{-1} (Donato et al. 2011), and substantially higher than other ecosystems such as tropical rainforests (243 Mg C ha^{-1}), boreal forests (408 Mg C ha^{-1}) and temperate forests in general (147 Mg C ha^{-1}) (IPCC 2000). Consequently, floodplains may be expected to have an important function in the global carbon cycle and climate change. On the other hand, floodplains are highly vulnerable to climatic change or anthropogenic impacts. River discharges can change with rising temperatures and decreasing precipitations, which may alter soil conditions and vegetation growth and cause a decline in the carbon storage capacity of floodplains. Moreover, floodplain areas have long been converted to other land uses, such as agriculture, due to their higher fertility, which may have led to relevant historic carbon releases (Mitra et al. 2005; Jaramillo et al. 2003).

Given the theoretical importance of riparian zones in the carbon cycle, methods for production of large scale maps showing the spatial distribution of C_{org} in floodplains are necessary. Various studies have focused on carbon stocks in floodplain soils (Busse and Gunkel 2002; Giese et al. 2000; Hazlett et al. 2005) and (sub-) tropical wetlands (Matsui et al. 2009; Mitsch et al. 2010). More recently Cierjacks et al. (2011) provided statistical models to describe the spatial distribution of C_{org} in floodplain soils and vegetation. However, these methods were field-based, where data had to be collected by cost-intensive terrestrial surveys. To facilitate the assessment of C_{org} stocks, for larger or less accessible wetland and floodplain areas, methods of remote sensing should be developed. There are several studies using remote sensing for the general analysis of wetland and floodplain ecosystems, e.g. the application of Lidar to classify cottonwoods (CW) in Arizona (Farid et al. 2008) or the Mackenzie delta in Canada (Mertes 2002), or the analysis of orthophotos in the Danube floodplains (Wagner

2009). Spectrometer data have been analyzed for dynamic vegetation systems in the Netherlands (Kooistra et al. 2008), and Landsat data have been used to classify Zambian Kafue floodplains (Munyati 2000) and riparian forests in Xinjiang, NW China (Thevs et al. 2008). Object-based image analysis (OBIA) has been applied to map coastal marshes in Georgian Bay in Canada (Rokitnicki-Wojcik et al. 2011; Midwood and Chow-Fraser 2010), Canadian peatlands (Dissanska et al. 2009), wetlands in Brazil (Evans et al. 2010; Silva et al. 2010), and mangrove forests in Senegal (Conchedda et al. 2008). Other studies have compared different approaches for the remote sensing based analysis of wetlands and floodplains (Adam et al. 2010; Ozesmi and Bauer 2002).

Additionally, in the context of carbon budgeting, remote sensing has been successfully applied to riparian ecosystems such as Amazonian rain forests (Neeff et al. 2005), Japanese spruce stands (Awaya et al. 2004), Scandinavian forests (Olofsson et al. 2008) and areas in North America (Turner et al. 2004).

In particular, the use of high spectral resolution data indices such as the Normalized Differenced Vegetation Index (NDVI) or the Leaf Area Index (LAI) have proven to be useful tools to calculate the gross primary production of C_{org} (Hilker et al. 2008). As well, the application of remote sensing for the detection of soil characteristics has successfully been applied (McBratney et al. 2003; Behrens and Scholten 2006), but these methods were predominantly applied to bare soils such as agricultural lands, and not to riparian soils. In addition, most of these studies have considered C_{org} either stored in (riparian) vegetation or in soil; a comprehensive quantification of both above- and below-ground carbon storage capacity is yet missing.

Different vegetation types of floodplains show distinct capacities to store total C_{org}. Cierjacks et al. (2010) have shown, for Danubian floodplains, that above- and below-ground carbon stocks were different and increasingly higher in young plantations, softwood (SW) forests, and CW and mature HW forests.

In addition, spatial information and forest structure data were closely related to soil and vegetation carbon stocks (Cierjacks et al. 2011). These biological and spatial relationships may be used to map total carbon stocks based on vegetation classification using remote sensing.

This study aims to provide a spatial estimation of C_{org} stocks in floodplain vegetation and soils in a Central European floodplain ecosystem, with the use of very high spatial resolution (VHSR) remote sensing data and auxiliary data. In Central Europe, only a few floodplain areas have been preserved in their natural state. Among these, the Danube Floodplain (*Donau-Auen*) National Park in Austria offers a good research area due to its high conservation status. Our study focuses on the classification of VHSR remote sensing data with OBIA to distinguish the following vegetation types: SW, HW, and CW forest, meadows, reed beds and water areas. Each vegetation type has a known carbon storage capacity that may be used for large-scale mapping of C_{org} stocks in floodplains.

Existing knowledge from soil science can be made regionally applicable; there is great demand for such information in the field of landscape and environmental mapping, and in particular for use in estimating the role of soils in climate change.

In particular, in this paper the following research objectives will be addressed:

1) Central European floodplain habitats will be classified by OBIA.

2) The use of OBIA to improve the classification accuracy of vegetation cover mapping in Central European floodplain habitats will be evaluated.

3) Reliable, integrated above- and below-ground carbon stock estimates in floodplains will be made with OBIA by using VHSR remote sensing data and auxiliary data.

2 Materials and Methods

2.1 Study Area

The study area has a size of 11.3 km2 and is located in the center of the Danube Floodplain National Park, close to the village of Orth (16.66° E, 48.4° N) in Lower Austria, between the Austrian capital Vienna and the Slovak capital Bratislava. The Danube River flows through the National Park for a distance of about 36 km with an average width of 350 m, without any restraints caused by barrage. Apart from the construction of the Marchfeld dike in the 19th century to protect the areas north of the river from flooding and to improve navigation on the Danube, the area has hardly been touched by human impact. Throughout its history, its only use has been as an imperial hunting ground. However, plantings of hybrid poplars (CW, *Populus x canadensis*) in the 1960s altered the forest structure, in particular on the southern riverbank. In 1996 the area was declared a national park, banning any commercial enterprise in its territory. Despite previous human interventions the area remains one of the last major wetland areas in Central Europe, recognized by the IUCN as a Riverine Wetlands National Park, category II.

The habitat types within the National Park comprise a wide range of environmental conditions (Danube River water body, various side channels and oxbow lakes, gravel banks on shores and islands, riparian forests, meadows and xeric habitats). Haplic Fluvisols (calcaric) and Gleysols (calcaric) are the main soil types in the area. The climate is continental with a mean temperature of 9.8°C and a mean precipitation of 533 mm [Schwechat climate station, 48° 07' N, 16° 34' E, 184 m above sea level (ZAMG 2002)]. The flow velocity of the Danube River ranges between 2 and 2.5 m/s; the mean discharge is 1,950 m3/s (Zehetner et al. 2009).

The area was selected due to its high conservation status and the spatial contiguity of the territory, the existence of a high quality geographical database for the region as well as the large number of previous studies that have been done on the soils and vegetation of the area (Cierjacks et al. 2010; Cierjacks et al. 2011; Ellenberg 1986; Lair et al. 2009; Wagner 2009; Zehetner et al. 2009).

2.2 Data

For this study, we used cloudless Ikonos-2 imagery recorded on 22nd of April 2009. By that time vegetation had sprouted completely. The Ikonos-2 image consists of a single-band panchromatic band (450 to 900 nm) with a spatial resolution of 1 x 1 m, and four spectral bands with a spatial resolution of 4 x 4 m covering the following intervals: blue (445 to 516 nm), green (506 to 595 nm), red (632 to 698 nm), and near-infrared (757 to 853 nm). The data have been geo-referenced by the provider to the World Geodetic System 1984 (WGS 84) datum, zone 33 N, with Universal Transversal Mercator (UTM) projection with an accuracy of 8 m RMS according to the data service provider (GeoEye 2009).

A digital elevation model (DEM) derived from Lidar data with a spatial resolution of 2.4 9 2.4 m and a vertical resolution of 1 m was used to create a slope model that informed some of the knowledge-based rules. The Lidar model was provided by the Danube Floodplain National Park administration and has been created by the Institute of Photogrammetry & Remote Sensing of the Vienna University of Technology. Increased slope values can be considered evidence of former riverbeds (an indicator of SW stands away from actual riverbeds) that cannot be detected directly through spectral values. Additionally, the height above ground was included in the knowledge-based approach.

Training data were derived from Cierjacks et al. (2010) using data from 2008. The data were collected in a stratified randomized sampling design throughout the research area in 10 9 10 m plots. For each sample plot, detailed information, collected on the ground, includes data on the C_{org} stored in the soil and in the trees. These data were used to select classification samples of the riparian vegetation types in the OBIA.

Table II-1: List of remote sensing and ground data used for the classification of floodplain vegetation cover, classification accuracy assessment and C_{org} stock estimation for the Danube Floodplain National Park, Austria

Type of geo-data	**Date**	**Relevant and derived information**	**Use of data**
Ikonos-2 imagery	2009/04/22	Vegetation type	Object-based classification
Digital elevation model (DEM)	2008	Height, slope	Knowledge-based classification; indicators of old riverbeds and softwood vegetation
Field survey data (76 points; Cierjacks et al. 2010a)	2008	Vegetation type, soil parameters	Training data
Forest inventory data (72 points)	2009	Tree species, vegetation type	Validation data
Austrian topographic map (ÖK 50, 1:50,000)	1993	Topographic features, meadows, reeds, water bodies	Validation data

To validate the classification, a validation mask was created from an up-to-date topographic map and forest inventory data from 2009 provided by the Austrian Federal Forest (ÖBf) agency. The topographic map was used to generate a water mask as well as masks for the classes reed bed and meadow for validation. The forest inventory points were arranged in a grid of 100 9 400 m. At each point, all trees within a radius of 8 m (total area of 201 m2) were measured and categorized. Tree species were grouped into the classes HW, SW and CW forest, as these are the main vegetation types of the area (Cierjacks et al. 2010). If a forest inventory point contained tree species of all three classes, the point was not used as a reference because it was located in a transition zone. In total, the validation dataset consisted of 80 validation areas, of which 7 were CW, 20 HW, 24 SW, 9 meadow, 7 reed bed and 13 water. Each validation area had a size of about 201 m2 and about 201 pixels, that were used for the calculation of classification accuracy. An overview of the data is given in Table II-1.

2.3 Segmentation and classification of vegetation types

In the classification of remote sensing data with VHSR data, such as Ikonos-2 data, OBIA has become an important and powerful tool. In comparison to traditional remote sensing classification techniques, which work with single pixels, the approach of OBIA is different. First, the image is split up into homogenous and spatially contiguous image objects. This process is known as image segmentation. The image objects are built by an algorithm based on their shape, colour, compactness, and smoothness (Baatz and Schäpe 2000). In this study, we used the commercial software package eCognition Developer 8.64. The complete description for the region growing algorithm can be found in Baatz and Schäpe (2000) and Benz et al. (2004). The most important formulas for image segmentation are as follows:

The an d-dimensional feature space the degree of fitting h can be calculated as Baatz and Schäpe (2000):

$$h = \sqrt{\sum_d (f_{1d} - f_{2d})^2}$$

The form heterogeneity *h* is defined as Baatz and Schäpe (2000):

$h = \frac{l}{\sqrt{n}}$ where *l* is the factual edge length of an object, and *n* is the object size in pixels.

The technique has been successfully applied to the detection of forests as well as other vegetation types (Chubey et al. 2006; Förster et al. 2008; Förster and Kleinschmit 2008; Wagner 2009). In this study, several segmentation scales were tested for their suitability to delineate vegetation types. We chose a scale parameter (SP) i.e., the maximum allowed heterogeneity within an object, of 100, a shape factor of 0.3, and a compactness factor of 0.8

as they represent best the regional forest conditions. The average image object size was 0.21 ha.

After the creation of image objects (segmentation), about 15 training samples based on spectral values were visually collected for each class. In a series of nearest neighbor classifications, water, meadow and reed bed objects were classified; impervious surfaces were classified as well, and then masked out. Afterwards, we classified the vegetation types CW, SW and HW forest.

All classes were described by fuzzy logic membership functions of their spectral values; for the more exact discrimination of different forest vegetation types, we used additional knowledge and information (Fig. II-1).

According to the theory of fuzzy logic, sets, or in this case, vegetation types, can often better be described by fuzzy boundaries determined by a membership function than by crisp boundaries. Fuzzy logic is used to implement linguistic, verbal descriptions (Zadeh 1989). The occurrence of different forest types depends on specific ecological, but also anthropogenic influences. These conditions allow or prevent species and habitats from existing and can be related to geo-factors. For the classification of riparian vegetation, the spatial distance to the water and slope were taken into account.

The following knowledge-based rules for the classes CW, HW and SW were implemented with the help of fuzzy logic (Fig. II-1):

(1) CW forest: according to personal communication with the National Park administration, at that time the CW plantings were carried out, they were only considered cost-effective when a minimum size of *0.2 ha could be cultivated. Therefore, homogeneous areas under 0.2 ha were unlikely to be CW. For this classification the size was defined as 2,000–4,000 m2.

(2) SW forest: a threshold of 48 m for the mean distance to the class water was introduced (Cierjacks et al. 2010, 2011). As SW forests can also be found along previous riverbeds, the slope derived from the digital elevation model was used to derive the course of former streams.

(3) CW, SW and HW forest: the first-order classification was further refined with a reclassification step. This step was necessary, as the multiresolution segmentation process tends to subdivide large patches of equal vegetation type up to a maximum object size. This causes structures with smaller image objects especially in homogeneous classes, which sometimes tend to be misclassified. In order to straighten the boundaries between the

vegetation types, single image objects of a certain class surrounded by other classes were classified according to their surrounding classes.

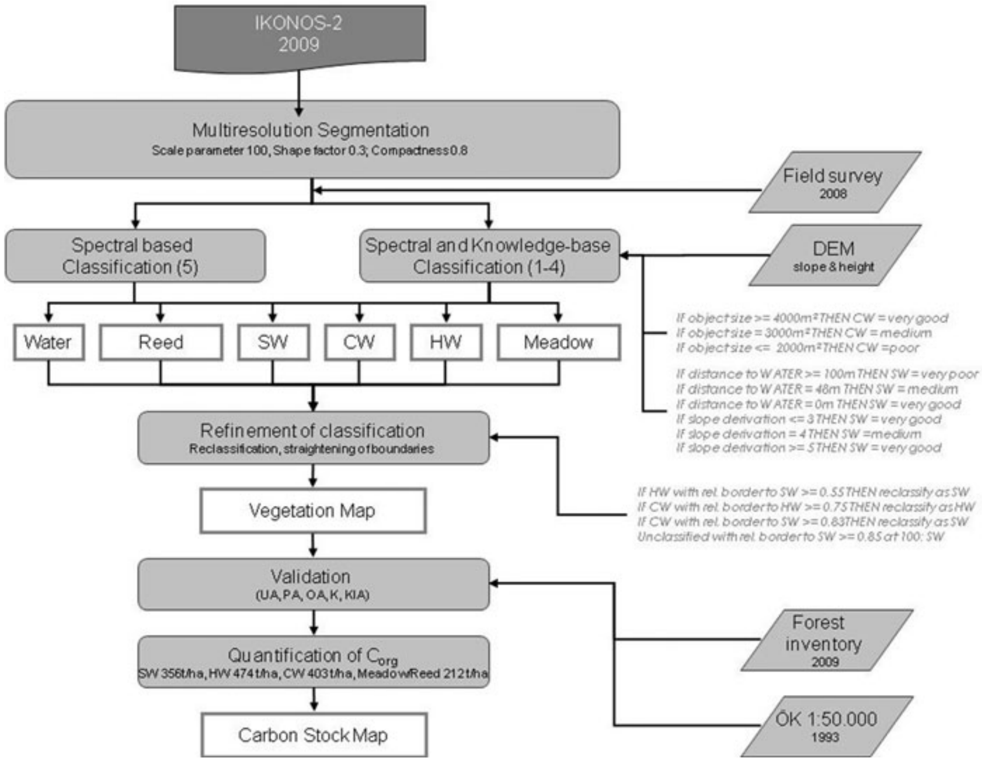

Fig. II-1:Schematic overview of the developed approach for combining high resolution remote sensing data with additional geodata (DEM digital elevation model, CW cottonwood forest, SW softwood forest, HW hardwood forest, UA user's accuracy, PA producer's accuracy, OA overall accuracy, K overall kappa statistic, KIA kappa per class, C_{org} organic carbon, ÖK Austrian topographic map)

The influence of the implemented rules was determined by running five classification iterations (classification sets) with different rule combinations. The accuracy assessment was conducted for classification based on (1) spectral and knowledge-based classification including all rules, (2) spectral and knowledge based classification without features for CW class, (3) spectral and object-feature classification without distance features for SW class, (4) spectral and object feature classification without slope features for SW class, and (5) spectral classification without any knowledge-based rules.

To evaluate the classification accuracy, the user's accuracy (UA; also known as commission errors), producer's accuracy (PA; also known as omission errors), and the overall accuracy (OA) were calculated as well as the kappa (K) and the kappa per class (KIA) statistics. The kappa coefficient serves as an additional measure of agreement between the classes represented in the satellite image and on the ground. The measure describes what level of agreement is due to chance; a kappa value of 1 describes a very high classification accuracy, a kappa value of 0 a very low accuracy (Congalton 1991; Fitzgerald and Lees 1994; Landis and Koch 1977; Lillesand et al. 2004).

2.4 Estimation of C_{org} stocks

After the different vegetation types were classified, the information obtained was used to estimate the carbon stocks. The following above- and below-ground amounts of C_{org} (mean and standard error, SE) were used from the studies of Cierjacks et al. (2010, 2011):SW 356 Mg C ha^{-1} (SE 35), CW 403 Mg C ha^{-1} (SE 29), HW 474 Mg C ha^{-1} (SE 61), meadows 212 Mg C ha^{-1} (SE 21) and reed beds 212 Mg C ha^{-1} (SE 21). The data of these studies are in line with other publications (Fierke and Kauffman 2005; Hofmann and Anders 1996). There were no data on C_{org} stocks underneath the water bodies, therefore these classes were not taken into consideration.

A Monte Carlo simulation was performed with the statistics software package R for the stock estimates for the classes HW, SW and CW. Reed beds and meadows were not considered for the simulation for several reasons: their clear-cut delineation from forest classes, their similar values of C_{org} stocks, and their lack of any above-ground woody C_{org} (Cierjacks et al. 2010). The calculation steps of the Monte Carlo simulation were as follows:

(1) The probabilities for the vegetation types classified through the OBIA process were computed for the classes CW, SW and HW.

(2) The meadow and reed bed areas were subtracted from the entire study area, leaving the remaining areas as CW, SW or HW.

(3) The standard deviation (σ) for each class (including meadow and reed bed) was determined based on the published mean (l) carbon stock estimates and the standard errors (Cierjacks et al. 2010), where σ = standard error x √(number of study plots).

(4) The total areas of CW, SW, and HW were determined based on their probabilities.

(5) Carbon stocks for each pixel in a class (meadow, reed bed, CW, SW and HW) were created at random based on the l and r values.

(6) The total carbon stock for each class was calculated by summing all pixels of that class.

(7) Steps 1–6 were repeated 1,000 times.

(8) The mean value, median value, 95% confidence interval (CI), standard deviation, and standard error were calculated for the generated carbon stocks.

3 Results

3.1 Segmentation and classification of vegetation types

The five classification sets showed pronounced differences in the different accuracy measures (Table II-2). Overall accuracies for the different knowledge and spectral-based classifications ranged from 0.70 (classification sets 1 and 3) to 0.69 (classification set 4) to 0.60 (classification sets 2 and 5). The confusion matrix for classification set 1 highlights the high accuracy values for this set (Table II-3).

Kappa values ranged from 0.64 (classification set 1: knowledge-based including all rules) to 0.52 (classification set 2: spectral and knowledge-based without features for CW class). For quality control, there are considered to be five categories of agreement (Congalton 1991; Landis and Koch 1977): almost perfect (0.81–1.00), substantial (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40) and slight (0.00–0.20). Consequently, our overall accuracies were moderate to substantial.

Classification accuracy for water was very high in all classification sets and accuracy measures. While the user’s accuracy for the class meadow was also very high, the producer’s accuracy was lower (0.71 for all sets) as was the KIA (0.69 for all sets). Also for the class reed bed, the accuracy measures ranged from 0.71 to 0.86 for user’s accuracy, 0.86–0.88 for producer’s accuracy, and 0.84–0.87 for KIA values.

The OA and KIA values were the same for classification sets 2, 3, 4, and 5, which only implemented knowledge-based classification rules in part and differed from classification set 1 which implemented all rules.

The classification sets exhibited clear differences in their delineation of the forest classes. Comparing the classification that used only spectral values (classification set 5) with the one that used both spectral values and additional knowledge (classification set 1), the knowledge-based classification led to a much higher percentage of SW forests and a more pronounced agglomeration of HW patches (Fig. II-2).

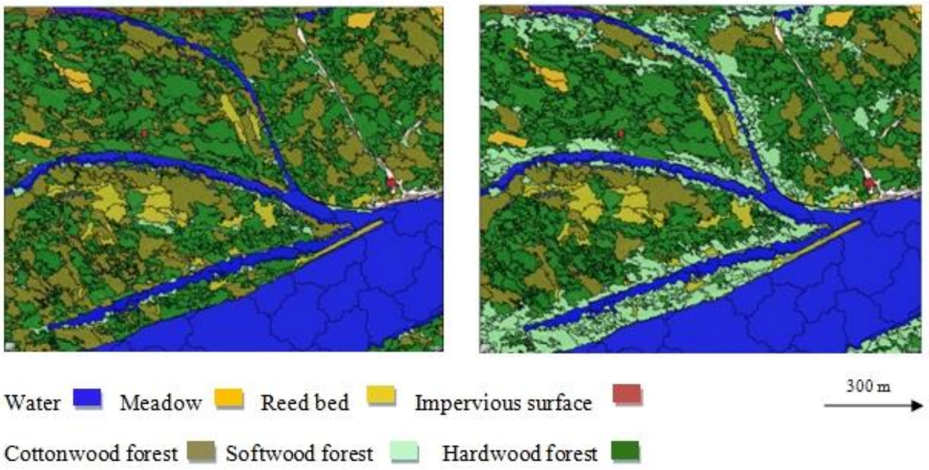

Fig. II-2: Classification based on spectral values only (left) and classification based on spectral values and additional knowledge (right)

Table II-2: Accuracy Assessment for classification based (1) spectral and knowledge-based classification including all rules, (2) spectral and knowledge-based classification without features for cottonwood class, (3) spectral and object-feature classification without distance features for softwood class, (4) spectral and object-feature classification without slope features for softwood class, and (5) spectral classification without any rules

Classification set		1			2			3			4			5	
OA		0.70			0.60			0.70			0.69			0.60	
K		0.64			0.52			0.63			0.61			0.53	
Accuracy/ Reference class	UA	PA	KIA	UA	PA	KIA	UA	PA	KIA	UA	PA	KIA	UA	PA	KIA
Water	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Meadow	1	0.71	0.69	1	0.71	0.69	1	0.71	0.69	1	0.71	0.69	1	0.71	0.69
Reed	0.86	0.88	0.87	0.71	0.86	0.84	0.71	0.86	0.84	0.71	0.86	0.84	0.71	0.86	0.84
Cottonwood	0.47	0.75	0.71	0.23	0.87	0.82	0.47	0.75	0.71	0.47	0.75	0.71	0.23	0.87	0.82
Softwood	0.76	0.51	0.39	0.83	0.29	0.21	0.76	0.5	0.39	0.58	0.87	0.78	0.64	0.52	0.38
Hardwood	0.51	0.63	0.47	0.53	0.41	0.27	0.53	0.64	0.48	0.66	0.17	0.11	0.66	0.17	0.11

OA= overall accuracy, K= overall kappa statistic, UA= user's accuracy, PA= producer's accuracy, KIA= kappa per class

Incorporating the CW rule (classification sets 1, 3, and 4) increased the user's accuracy (0.47 vs. 0.23 for classification sets 2 and 5) but lowered the producer's accuracy (0.75 for classification sets 1, 3 and 4 vs. 0.87 for classification sets 2 and 5) and KIA (0.71 vs. 0.82).
Classification accuracy for SW forest differed widely among the different classification sets, with user's accuracy ranging from 0.58 (classification set 4) to 0.83 (classification set 2), producer's accuracy ranging from 0.29 (classification set 2) to 0.87 (classification set 4) and KIA ranging from 0.21 (classification set 2) to 0.78 (classification set 4).
For HW forests, user's accuracy ranged from 0.51 (classification set 1) to 0.66 (classification sets 4 and 5), producer's accuracy from 0.17 (classification set 5) to 0.64 (classification set 3) and KIA from 0.11 (classification sets 4 and 5) to 0.48 (classification set 3).

Table II-3: Confusion matrix for spectral and knowledge-based classification that includes all rules (classification set 1) based on pixels

User \ Reference Class	Water	Meadow	Reed	Cottonwood	Softwood	Hardwood	Sum
Water	2612	0	0	0	0	0	2612
Meadow	0	1277	0	0	0	1	1278
Reed	0	199	1244	0	0	0	1443
Cottonwood	0	0	1	931	310	737	1979
Softwood	0	0	0	18	2102	629	2749
Hardwood	0	115	167	297	1743	2375	4697
Unclassified	0	199	0	0	0	0	199
Sum	2612	1790	1412	1246	4155	3742	
Producer	1	0.71	0.88	0.75	0.51	0.63	
User	1	1	0.86	0.47	0.76	0.51	
KIA Per Class	1	0.69	0.87	0.71	0.39	0.47	
Overall Accuracy	0.70						
KIA	0.64						

The confusion matrix for classification set 1 shows the number of pixels that were correctly classified or misclassified; meadow tended to be misclassified as reed bed or HW forest; reed bed was misclassified as HW forest (Table II-3). Moreover, there was confusion within the three forest classes HW, SW and CW. As this classification set produced the best OA and kappa values, it was used for the estimation of carbon stocks.

3.2 Estimation of C_{org} stocks

The Monte-Carlo simulation revealed the following overall carbon stocks for the study area: mean and median carbon stocks were similarly at 483.6 Gg (standard deviation: 0.074 Gg and standard error: 0.002 Gg; Fig. II-3). The 95% CI ranged less than 10-2 Gg from the mean for the total amount of C_{org} as well as for each vegetation type class.

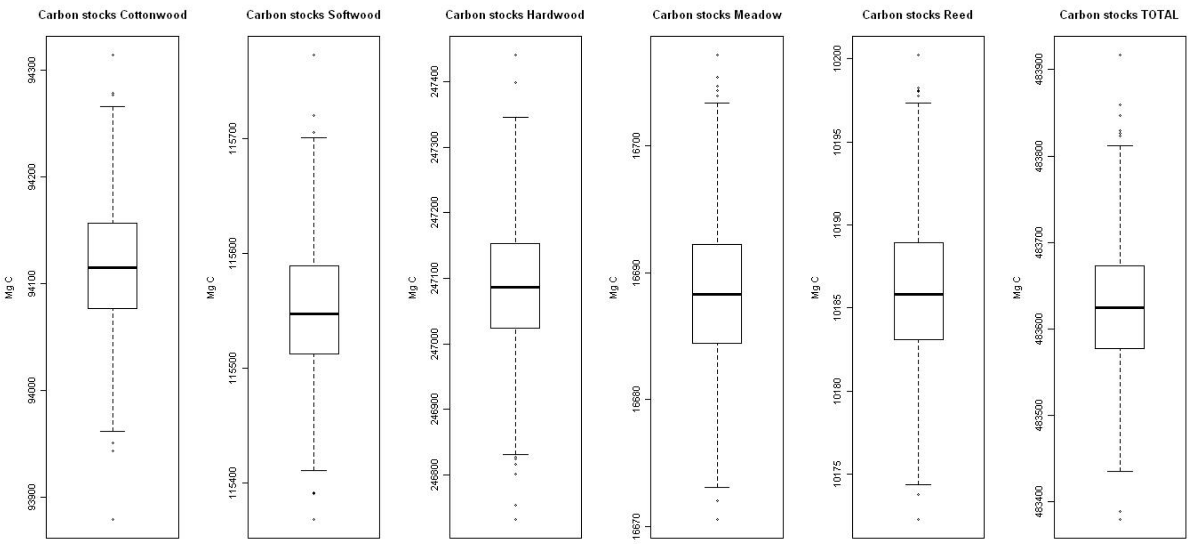

Fig. II-3: C_{org} stocks for all vegetation classes and total amount in the study area

The spatial distribution of Corg stocks corresponds to the distribution of the single vegetation types (Fig. II-4). In patches of reed beds or meadows, the amount of C_{org} was clearly lower than in the patches of HW or CW vegetation. Almost 62% of the terrestrial research area (not including the water bodies) has carbon stocks higher than 400 Mg C ha^{-1} (CW and HW), and about 26% of the area has carbon stocks of 356 Mg C ha^{-1} (SW). Over the whole study area, the average value of stored above- and below-ground C_{org} is 428.9 Mg ha^{-1}.

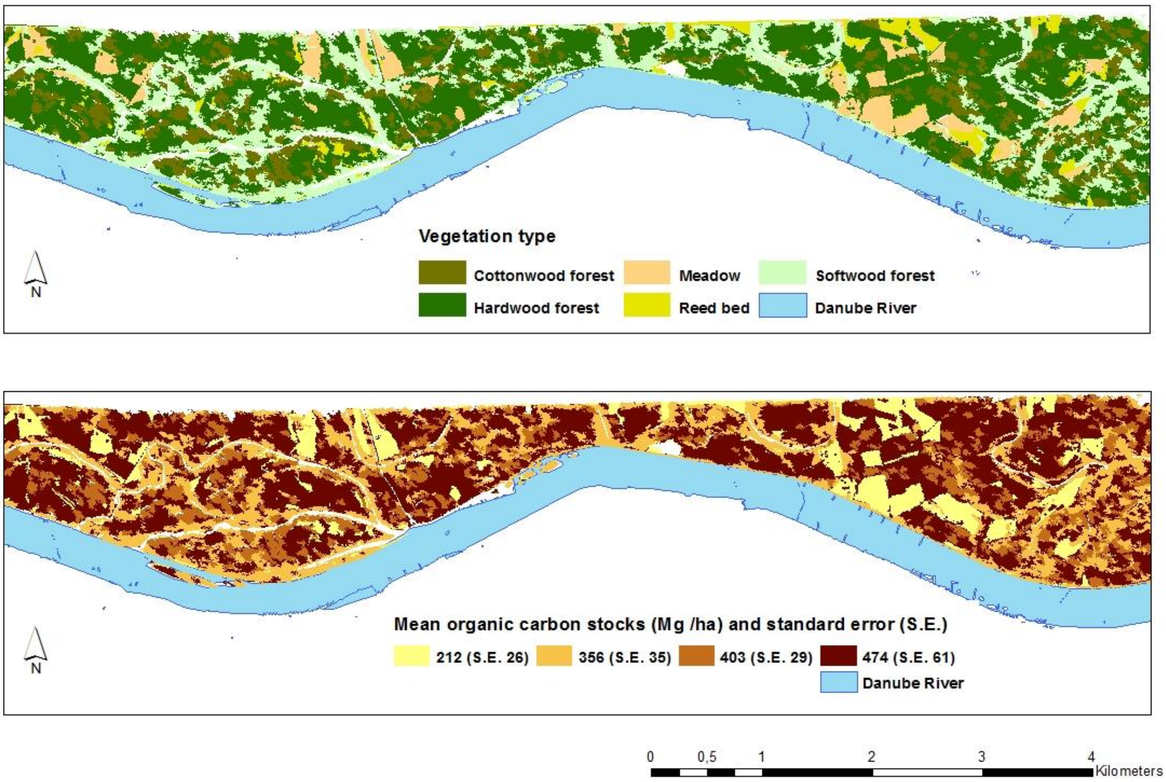

Fig. II-4: Classification of vegetation types and estimation of total (above- and below-ground) C_{org} stocks

4 Discussion

4.1 Segmentation and classification of vegetation types

Our results clearly show that the classification accuracy for all classes can be described as substantial (Congalton 1991; Landis and Koch 1977). In a study in the same research area based on OBIA of orthophotos (Wagner 2009), the overall kappa value reached 0.87; it can be argued that their accuracy was higher due to the higher spatial resolution of the orthophotos (0.2 m) in comparison to Ikonos data (1 m).

Concerning the accuracy of the particular classes of the classification (Table II-2), it is obvious that the correct classification of water bodies does not raise any problems, shown by an accuracy of 100% in each of the sets. The distinction of vegetation classes was more difficult. Meadows and reed beds were spectrally similar and tended to be confused with each other (and to a lesser extent with forest classes), therefore the producer’s accuracy and kappa values did not reach 100%, whereas the class meadow showed a user’s accuracy of 100%. In a study on coastal marshlands in Canada (Rokitnicki-Wojcik et al. 2011), the class meadow/shrub had one of the highest accuracies, but the neighboring classes were distinguished by their age in Rokitnicki-Wojcik’s study, not by their vegetation type. It is possible that a more refined analysis of the texture could create better results.

The classification accuracy for the classes SW, HW, and CW forest varied with the inclusion of additional knowledge. Our results provide evidence that a knowledge-based approach can improve classification results in comparison to a classification based only on spectral values, although this was not true for every class. In general, the more knowledge rules applied, the better the classification accuracy becomes, and potential errors through over- or underestimation are omitted.

In particular, the rule for the minimum size of a CW patch (classification set 2) improved the classification. This could be easily related to the fact that an economic silvicultural use of CW in this region required that forests stands had a certain minimum area. Therefore, for the purpose of our study, this knowledge-based classification rule was very valuable.

Distance to water (classification set 3) had been assumed to be a very important factor as proximity to the river is related to the presence of SW (Cierjacks et al. 2010; 2011; Ellenberg 1986). However, the inclusion of this information only slightly increased the overall classification accuracy. A convincing reason for this fact may be the application of the slope feature, which even without further rules provided a sufficient indicator of riparian ecosystems.

Consequently, when omitting the slope derived from a digital elevation model (classification set 4), the overall classification accuracy was slightly impaired. However, SW forest itself was overestimated. It appears that the two knowledge rules regarding distance to water and slope for the SW class were interchangeable and can be substituted. Overall, the changes in overall classification accuracy were rather low when the distance or surface slope rules were included. But they remained important for the individual accuracy of the SW class as illustrated in Fig. II-2.

In general, the detection of different forest vegetation types, especially in a mixed deciduous area, remains a challenging task, as shown by other studies classifying deciduous forests (Förster et al. 2008). It may be possible to use additional knowledge-based rules derived from other data to avoid severe overestimations. One possibly would be the use of a digital surface model (DSM) derived from Lidar for the determination of the vegetation age (Farid et al. 2008).

4.2 Carbon stocks

The estimation of above- and below-ground carbon stocks in the study area revealed a total amount of 483.6 Gg C_{org}. The value is comparable to the estimation of Cierjacks et al. (2010) of 472.2 Gg C_{org}. This is not within the range of our standard deviation, but the values are comparatively close.

The Monte Carlo approach to estimate the total amount of C_{org} is not very common, yet there are examples for its use (Maeda et al. 2010). We used a Monte Carlo simulation because of two major uncertainty factors. One was the accuracy of the classification, the other one refers to the standard error rates for each vegetation type, based on the values by Cierjacks et al. (2010) for the carbon stocks. The results had low standard deviation and standard error rates, so the result seems to be adequate.

There are several methods for mapping of carbon stocks (Goetz et al. 2009): the stratify and multiply approach (the area of a certain land cover class is multiplied by the C stock value for the landcover class), the combine and assign approaches (a wider range of geodata apart from remote sensing imagery are used for the calculation) and the direct remote sensing approaches (repeated machine learning algorithms train an algorithm to develop rule sets).

For our approach, we used a synthesis of the first two, as there was a multiplication of class area sizes (as in the stratify and multiply approach) for the classes of reed bed and meadow. For the classes CW, HW and SW forest, we integrated further geodata (combine and assign approach). The direct remote sensing approach for the estimation of carbon stocks has not yet been implemented.

Overall, the method for estimating overall carbon stock using VHSR remote sensing and additional geodata proved to be applicable. Still, this method does not allow for the C_{org} estimation of different ecosystem compartments (soil, vegetation) independently.

Moreover, the critical role of riparian ecosystems in the global carbon cycle is expected to reflect carbon dynamics more than just the carbon stocks present in the ecosystem. Here, additional information such as spatial and vegetation structure data will be important for more realistic estimations.

5 Conclusions and outlook

This study represents an approach to estimate the amount of C_{org} in soil and vegetation of riparian ecosystems by using remote sensing and additional geodata for an OBIA. The calculation of carbon stocks based on the values for forest classes (Cierjacks et al. 2010) revealed similar results as the on-the-ground assessment. Therefore, our study may contribute to the development of methods to automate the creation of large-scale carbon maps for floodplains.

Future steps to improve the model may include the application of additional geographic data, such as biotope type or geomorphologic maps or historic topographic maps from the 18th and 19th centuries to identify former river courses and to incorporate data on various river stages. Also a fusion with Lidar should be considered as a Lidar-derived DSM can be helpful to determine vegetation age (Farid et al. 2008).

For the implementation of these additional knowledge factors, further research will be necessary. Here, machine learning algorithms, described by Goetz et al. (2009) can be useful.

In summary, the application of remote sensing methods for the quantification of C_{org} stocks in floodplains and other ecosystems may be helpful for the specification of the spatial estimation of global carbon stock inventories, which may be important for global carbon and climate change monitoring schemes.

Acknowledgments

We would like to thank the administration of the Danube Floodplain National Park, especially Christian Baumgartner and Christian Fraissl, as well as the Austrian Federal Forests Agency (ÖBf) for the support and for supplying us with geographic data, especially with data for the validation of results. We would like to thank Tobias Schmidt for the data processing and statistical support, and Kelaine Vargas for improving our English. We would also like to thank two anonymous reviewers for their helpful and detailed comments on our manuscript. The German Research Foundation (DFG): project-number 2215/2-1, provided support for this project.

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

Photogrammetrie, Fernerkundung und Geoinformation (PFG) 4 (2013): 333-349

Leonhard Suchenwirth, Michael Förster, Friederike Lang, and Birgit Kleinschmit

doi: 10.1127/1432-8364/2013/0181

Received: 26 February 2013 / Accepted: 22 April 2013 / Published online: 1 August 2013

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 Rapid-Eye), 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.

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 and 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 and 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 focused on C_{org} stocks in ecosystems, such as in alder fens (Busse and 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 and 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 and 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 km2 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 350m, 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 differentiate 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).

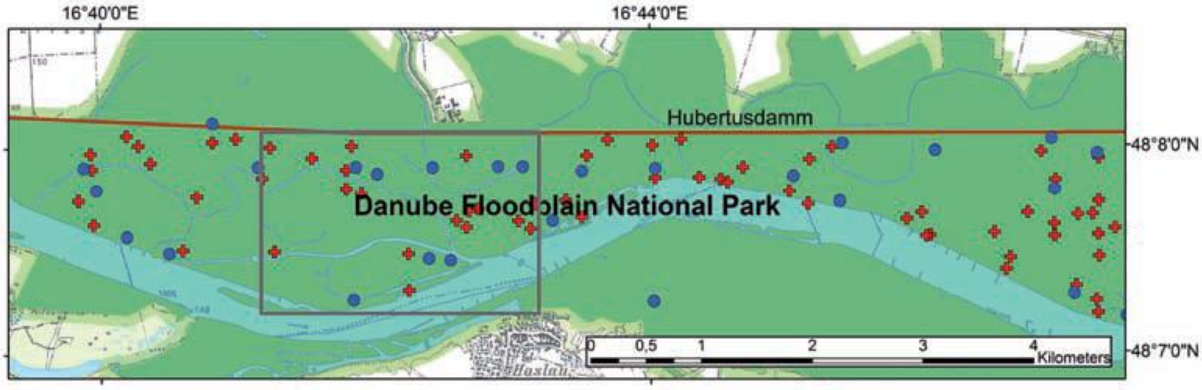

Fig. III-1: Research Area, green: Danube Floodplain National Park, red cross: locations of the terrestrial sample points training data, blue dot: test data. The red line represents the Hubertusdamm dike. The grey box represents the outline of the subsets in Fig. 2

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 km2) 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.

Tab. III-1**:** Available geodata and derived parameters.

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
	Red edge channel	b4 -REededge
	Near infra red 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 and Merzlyak 1994)	b4NDVI
	Solar Reflectance Index [b5/b3] (Rouse et al. 1973)	SRI
	[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)	hist 2
	Third Military Mapping Survey (1868 -1880)	hist 3
	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}

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. 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.III-1]. All data were collected in a stratified randomized sampling design throughout the research area in 10 x 10m 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 range of response types, such as numeric and categorical data, the ease and robustness of construction, and the ease of interpretation (De'ath and 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 data set 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.

Tab. III-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									

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

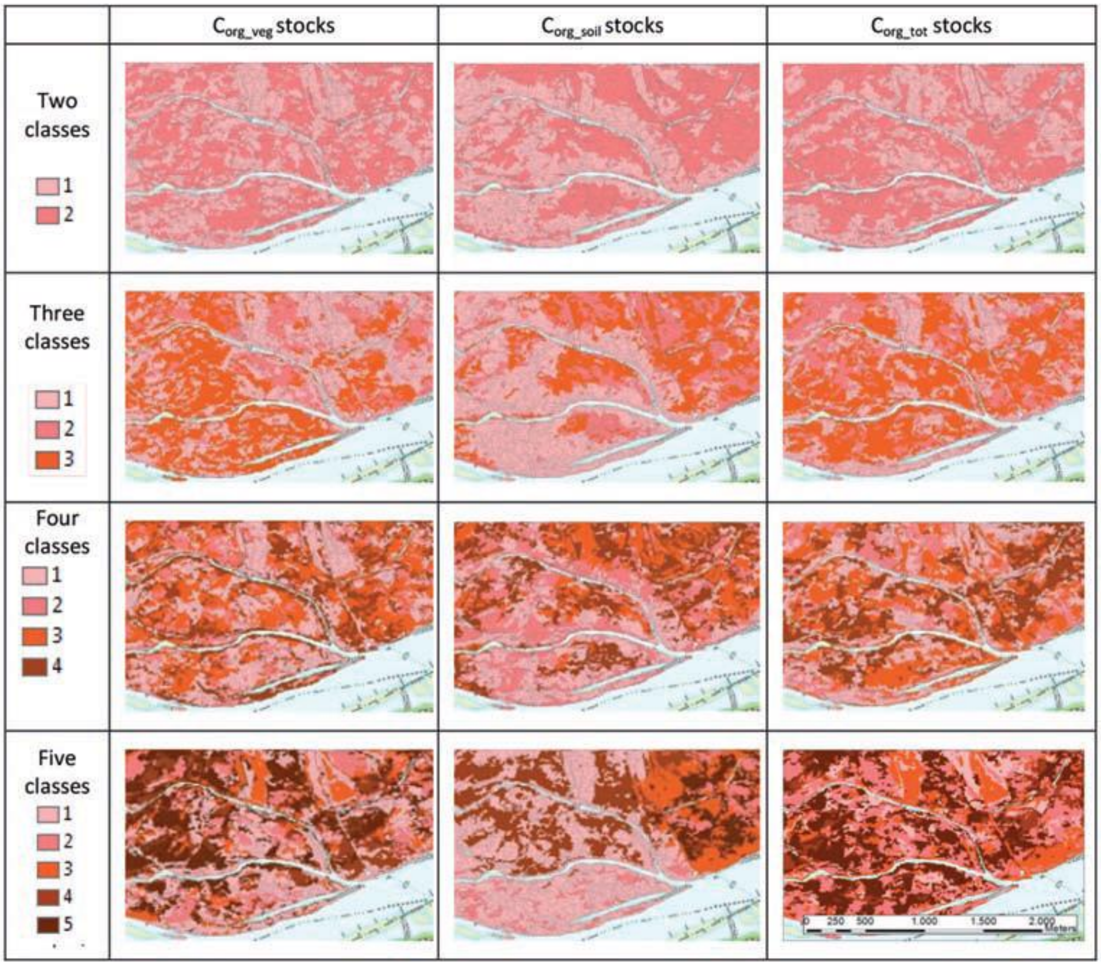

Fig. III-2: Modelled distribution of C_{org_veg}, C_{org_soil}, and C_{org_tot} stocks for different numbers of classes. The increasing amount of stored C_{org} is represented by colour graduations increases from pink to red to brown.

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. III-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. III-3). Models with three or four classes range in between and represent a good compromise between complexity and acceptable accuracy.
With regard to the model quality, we can examine Fig. III-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.

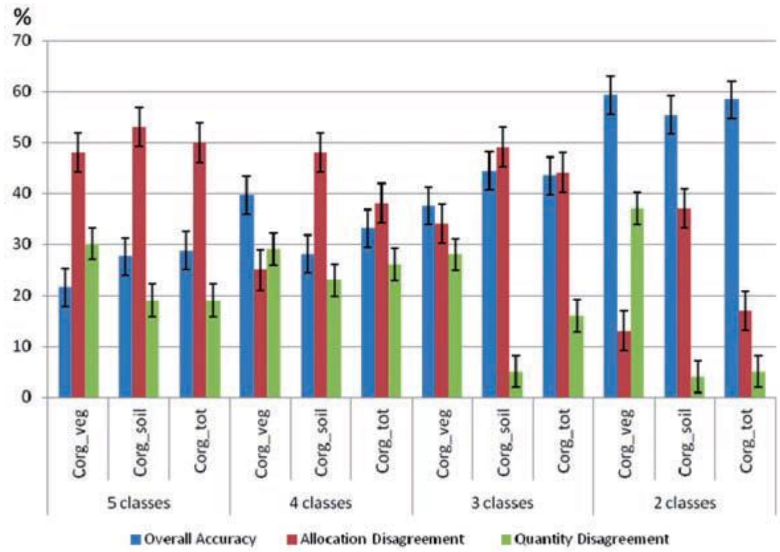

Fig. III-3: Overall accuracy, allocation, and quantity disagreement in percent for classifications of C_{org_veg}, C_{org_soil}, C_{org_tot} based on five, four, three, and two classes.

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.

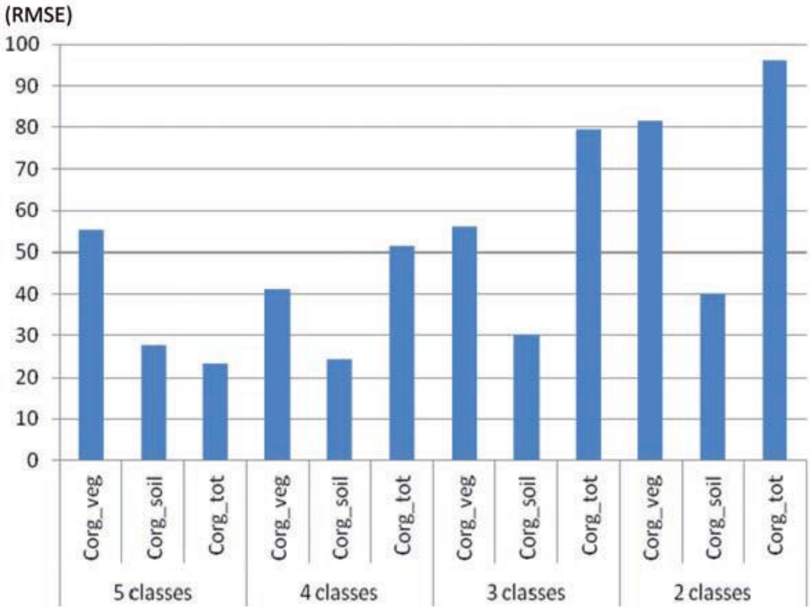

Fig. III-4: Root-mean-square error for classifications of C_{org_veg}, C_{org_soil}, C_{org_tot} based on five, four, three, and two classes.

To illustrate the importance of single parameters, Figs III-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 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 (Rapid-Eye: 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 %.

Tab. III-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

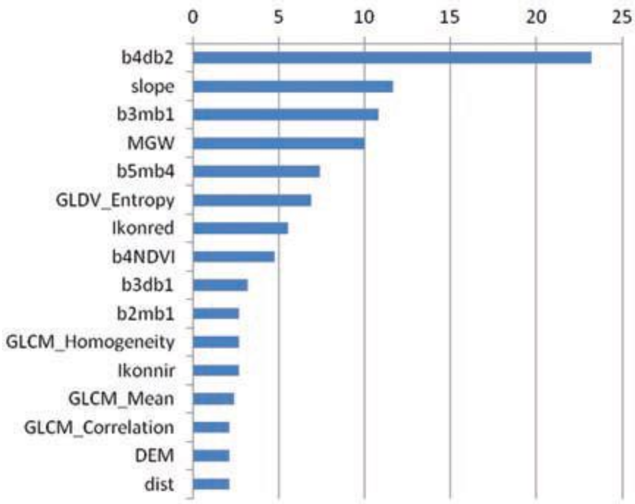

Fig. III-5a: Parameter relevance for C_{org_vet} classifications based on 4 quantile classes (all abbreviations are explained in Tab. III-1).

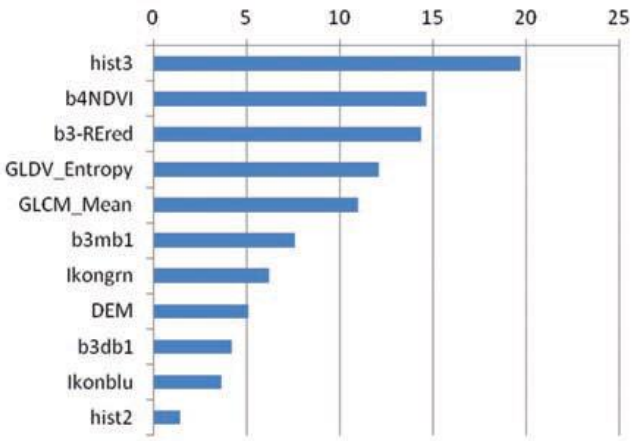

Fig. III-5b: Parameter relevance for C_{org_soil} classifications based on 4 quantile classes (all abbreviations are explained in Tab. III-1).

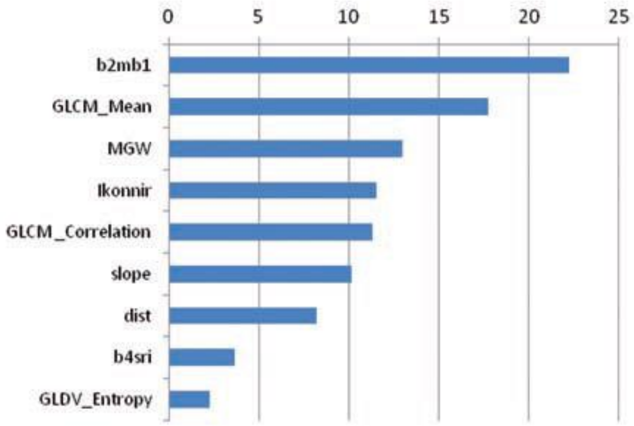

Fig. III-5c: Parameter relevance for C_{org_tot} classifications based on 4 quantile classes (all abbreviations are explained in Tab.III- 1).

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. III-3 and III-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. III-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; $R2$ = 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; $R2$ = 0.5619) and the red Ikonos channel (p = 0.0188; y = -0.4198x + 426.33; $R2$ = 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; $R2$ = 0.5301) and red Ikonos channel (p = 0.0081; y = -0.3752x + 208.54; $R2$ = 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 and 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. III-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. III-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. III-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.

Chapter IV: Large-scale mapping of carbon stocks in riparian forests with self-organizing maps and the k-nearest-neighbor algorithm

Leonhard Suchenwirth, Wolfgang Stümer, Michael Förster, and Birgit Kleinschmit

Submitted to iForest- Biogeosciences and Forestry on 28 May 2013; under review

Abstract

Among the machine learning tools being used in recent years for environmental applications such as forestry, self-organizing maps (SOM) classifications and the k-nearest neighbor (kNN) algorithm have been used with success. We applied the two methods for the mapping of organic carbon (C_{org}) in riparian forests as they have a considerably high carbon storage capacity. Despite the importance of floodplains for carbon sequestration, a sufficient scientific foundation for creating large-scale maps showing the spatial distribution of C_{org} is still missing. We estimated organic carbon in a test site in the Danube Floodplain based on RapidEye remote sensing data and additional geodata. As a result, carbon distribution maps of vegetation, soil and total C_{org} stocks were derived. The results were compared and statistically evaluated with terrestrial survey data for outcomes with pure remote sensing data as well as for the combination with additional geodata using bias and the Root Mean Square Error (RMSE). Results show that SOM and kNN approaches enable us to reproduce spatial patterns of riparian forest C_{org} stocks. While C_{org} from vegetation has very high RMSEs, outcomes for soil C_{org} and total C_{org} stocks are less biased with a lower RMSE, especially when remote sensing and additional geodata are conjointly applied. SOMs show similar percentages of RMSE to the kNN classifications.

1 Introduction

In recent decades, machine learning approaches have been introduced to manage the vast amount of data produced by various scientific disciplines, including environmental sciences such as forestry. One of the most intricate neural networks techniques are self-organizing maps (SOM), first described by Kohonen (1982). This unsupervised learning technology combines a high level of biological plausibility with applicability to numerous information processing and optimization problems. It allows one to reduce high dimensional information. The term 'maps' refers to the low dimensionality and does not necessarily imply a spatial or geographical application; in fact the technique emerged from neurosciences and there are a many examples from biosciences and engineering applications (Breijo et al. 2013; Wang et al. 2013; Xuan et al. 2013).

A different approach to the spatial classification of data is the k-nearest neighbor (kNN) technique; this so-called instance-based, 'lazy' learning algorithm often serves as a benchmark for other methods (Kanevski et al. 2009). It has frequent applications in forestry; the kNN method has been applied in a number of forest inventories, e.g. in Finland (Tomppo 1991; Tomppo and Halme 2004), New Zealand (Tomppo et al. 1999), Austria (Koukal et al. 2007) or Ireland (McInerney and Nieuwenhuis 2009). Some studies explicitly used kNN to estimate C_{org} (Fuchs et al. 2009; Magnussen et al. 2009; Stümer et al. 2010). The majority of studies are based on the use of Landsat data, few of them used VHSR (very high spatial resolution) satellite data.

Lek and Guégan (Lek and Guégan 1999) give a broad overview of applications in ecological and environmental sciences; recent applications include monitoring of river quality (Astel et al. 2007; Shanmuganathan et al. 2006), urban modelling (Arribas-Bel et al. 2011) and forestry applications (Adamczyk et al. 2013; Giraudel and Lek 2001). For the estimation of C_{org}, Stümer et al. (2010) successfully applied SOM and compared it with the k-nearest neighbor (kNN) algorithm for the assessment of biomass (and thus C_{org}) in Thuringian forests.

In the wake of the climate change discussion, it has become an essential task not only to decrease carbon emissions but also to identify natural carbon sinks in ecosystems all over the globe. Among terrestrial ecosystems, mangroves, peat lands and wetlands have especially shown an increased potential to sequester organic carbon in addition to other ecosystem services. For the case of riparian wetlands, several studies have underlined the high storage capacity (Hoffmann et al. 2009; IPCC 2000; Mitra et al. 2005).

The sequestration potential of floodplains is dependent both on vegetation (including forests, reed beds, and meadows), and soils. The important link between C_{org} stocks of forests and underlying soils has been demonstrated by a whole range of studies inside (Baritz et al. 2010; Harrison et al. 1995; Hofmann and Anders 1996) and outside Europe (Kooch et al. 2012; Lal 2005).

Even though the value of riparian ecosystems has been recognized, the scientific underpinning for mapping large-scale carbon stocks is yet to be established. On a global scale, as well as on the national level, C_{org} maps have been produced and validated; however a regional or local validation of results is typically not obtainable. Various remote sensing analyses of C_{org} stocks have been utilized for non-floodplain habitats, especially forests (Olofsson et al. 2008; Patenaude et al. 2005), but most of these studies have focused either on C_{org} stocks in soil or vegetation. Detailed C_{org} maps of floodplain areas have seldom been produced, apart from Suchenwirth et al. (2012).

In the presented study, we estimate organic carbon above and below ground in a test site in the Danube Floodplain based on a SOM and kNN classification of VHSR RapidEye data and ancillary geodata. Both results are compared to field survey data. In contrast to Stümer's application of SOM and kNN for Thuringian forests (2010), we consider the vegetation (above ground), soil (below ground) and total C_{org} in a floodplain area. Moreover, we introduce additional auxiliary geodata as input source data for both algorithms. In this way we compare the outcomes for remote sensing (RS) input information with the results for RS and additional information. We decided to apply SOM and kNN for the C_{org} models, as previous methods such as the derivation of C_{org} stocks from classified vegetation types (Suchenwirth et al. 2012) or the derivation via quantiles in a classification and regression tree (CART) approach (Suchenwirth et al. 2013) had only limited success.

The specific aims of this paper are as follows:

(1) to create distribution maps of vegetation, soil and total C_{org} stocks in a riparian forest, based on SOM and kNN algorithms and compare the results,

(2) to compare and evaluate results with previous estimation techniques,

(3) to evaluate the influence of additional geodata on estimation quality.

2 Material and methods

2.1 Study area

The research area is located inside the Danube Floodplain National Park (*Nationalpark Donauauen)* in Austria (16.66° E, 48.14°N). The area is a pristine floodplain area with few human impacts. Human activities included hunting in previous centuries, the construction of the Marchfeld dike in the 19th century, and the plantings of hybrid poplars (*Populus x canadensis*). Apart from these cottonwood plantations, the area is characterized by softwood forests (dominated by *salix alba, acer negundo*), hardwood forests (dominated by *quercus robur, fraxinus excelsior* and *acer campestre*), as well as meadows and reed beds. Our study area (11.7 km²) is limited by the Marchfeld dike (locally named Hubertusdamm dike) in the north, and the main river course towards the south. Geographic coordinates are given in Figure IV-1.

The area was chosen for our study due to its high protection status, a good base of geographic data, and previous research in the area (Lair et al. 2009; Wagner-Lücker et al. 2013; Zehetner et al. 2009). Mean C_{org} storage in the area was estimated at 359.1 Mg C_{org} ha^{-1} by Cierjacks et al. (2010), and as 428.9 Mg C_{org} ha^{-1} by Suchenwirth et al. (2012). Figure 1 presents a RapidEye scene of the Danube Floodplain Area. Red color indicates pixels with high content of active biomass.

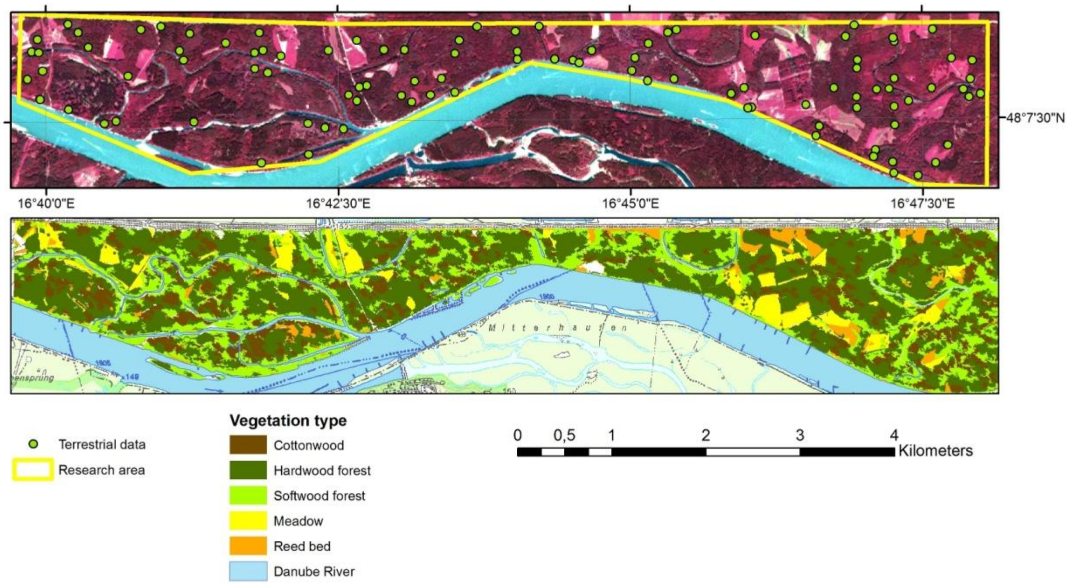

Fig. IV-1: Research area depicted as RapidEye RGB composite with terrestrial survey data (green dots; above) and vegetation classification (derived from Suchenwirth et al. 2012; below)

2.2 Data

We obtained a cloudless satellite image from RapidEye (recorded on August 1, 2009 in level 3A with a spatial resolution of 5.0 m (Sandau 2010)). The image was provided by the German Aerospace Center, in the UTM WGS 1984 reference system. We reprojected the image into the Austrian MGI M34 projected coordinate system, as local data were mainly available in the local reference system. Atmospheric correction was not performed as we did not work with time series. RapidEye data were used as they reflect spatial heterogeneity of carbon distribution in floodplains. Notably the RedEdge channel has already been successfully applied to improve classifications of vegetation (Schuster et al. 2012).

A digital elevation model (DEM) derived from Lidar data was used to compute altitude above river level; a groundwater model indicating median ground water depth was provided by the Vienna University of Technology. Distance to river (main stream) was derived from a topographic map. The topographic map is issued and updated every seven years by the Austrian Federal Office of Metrology and Surveying (*Bundesamt für Eich- und Vermessungswesen*).

In two terrestrial surveys in 2008 and 2010, a total of 104 samples from vegetation and soil were taken. C_{org} content of soil and vegetation was measured and calculated for each sample point (Cierjacks et al. 2010; Rieger et al. 2013). Total C_{org} consists of C_{org} in soil, vegetation, and dead wood on the ground.

Table IV-1 Available geodata and derived parameters

Available geodata	Derived parameters	Abbreviations
RapidEye image (August 1-2009)	Blue channel (440 -510 nm)	B
	Green channel (520-590 nm)	G
	Red channel (630-685 nm)	R
	Red edge channel (690-730 nm)	RE
	Near infra red channel (760-850 nm)	NIR
Digital elevation model	Elevation above river level	altitude
Ground water model	Ground water level	MGW
Topographic map 1:50.000 (ÖK 50)	Distance to river	distance
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}

2.3 Self-organizing maps (SOM)

The SOM approach is used to produce maps of C_{org} stocks in riparian forests of the Danube Floodplain. The method has been described in detail by Kohonen (1982, 2001), and has frequently been explained by other authors (Giraudel and Lek 2001; Kanevski et al. 2009; Stümer et al. 2010).

The application of SOMs is generally divided into two modes or phases: a learning (or training) phase and a classification or mapping phase. SOMs structure the neurons in the form of rectangular or hexagonal arrays or grids of nodes with n dimensions, with an associated weight vector attached to each node. The procedure of placing a vector from the high-dimensional data space into the two dimensional map space is performed by identifying the node with the closest associated distance to the presented data space vector, i.e. the winner pixel or best matching unit (BMU) is selected; its position within the grid is the excitation centre. Subsequently, differences between the weight vector and the data space vector are reduced. Afterwards, vectors in the neighborhood are adapted. The distance of the feature space is defined as the Euclidean distance. The learning process of the winner selection and adaption process is iteratively repeated until no further adaption is necessary, as the initial learning rate is much smaller than in the first stage and a stable state is reached. At this moment the learning phase is completed.

In the mapping phase, the input vector for which the prediction is necessary is presented to the map; distances from this location to all neurons are calculated. As a result, the BMU of the map is selected, providing a representative group of data samples to which the predicted input is most similar.

In our approach, we applied the algorithm programmed by Stümer et al. (2010). We use the RapidEye scene with additional geodata (see data section) for our classification as the initial layer. For the analysis, we used the following standard parameters: A feature space distance of five or eight (depending on the number of used channels/parameters), a start distance δ_{start} of 100,000 and an end distance δ_{end} of 100, and five iterations (t_{max}) were applied. It is necessary to set the start distance high in order to sufficiently consider the terrestrial samples, while at the end only the necessary neighbors shall be regarded.

2.4 k-nearest neighbor (kNN)

To compare the operational applicability of SOM we use the kNN method to provide spatially explicit results. It is described as the simplest, intuitively understandable und purely data-driven algorithm and is applied frequently for classification or regression tasks, or to provide a quick visualization or benchmark. It classifies a point by calculating the distances between the point and the points in the training data set. Then the point is assigned to the class which is most common among its k-nearest neighbors (with k being an integer number). There is no learning phase, since all training examples are simply stored in the memory for further predictions. The method was described by (Hall et al. 2008; Kanevski and Maignan 2004; Kanevski et al. 2009). We follow the method applied by Stümer et al. (2010).

For our kNN classification, we used standard settings to compare classifications: k = 5 neighbors; Euclidean distance $d_{(x1,x2)}$, of 2, and a distance weight $w_{(i),p}$ of 2. These parameter settings were often described as a compromise between a limited number of neighbors and a sufficient accuracy in other studies (Fuchs et al. 2009; McInerney and Nieuwenhuis 2009).

2.5 Validation

The reliability of C_{org} estimates obtained by the SOM- and kNN- approach is quantified by the bias and the root mean square error (RMSE). The bias is calculated as the difference between measured and estimated C_{org} stock; the RMSE includes variance of estimated C_{org} stock and the bias. The % RMSE facilitates comparisons between C_{org} measurements. In order to use terrestrial samples for both calibration and validation we used the Leave-one-out (L1o) cross–validation (Richter et al. 2012).

3 Results

The SOM and the kNN approach were used in the Danube Floodplain National Park. We produced two types of results: (1) spatially explicit maps of the vegetation, soil and total C_{org} stocks per unit (Mg C_{org} ha^{-1}), and (2) statistical estimates for vegetation, soil and total C_{org} stocks. The maps obtained by the SOM-approach were compared to alternative maps based on the kNN-approach. Terrestrial data were used as a basis for comparison of the statistical estimates obtained by the SOM-and kNN-approaches.

3.1 C_{org} stock estimations

C_{org} stock maps of vegetation, soil and total carbon, -based on the SOM method are displayed in Fig. IV-2.a and IV-2b. C_{org} maps based on the kNN method are presented in Fig. IV-3.a and IV-3.b. C_{org} stocks in the maps are displayed in a color range from yellow to red where lower stocks are indicated in light yellow, higher stocks in dark red, and for total C_{org} stocks color tones with higher values are in brown tones. All figures show the same detail of the area. The C_{org} stocks is given in tons per ha (Mg C_{org} ha^{-1}).

We can see from the satellite image (Fig. IV-1) that the wooded area has a dispersed distribution, with a high variation of vegetation within a small scale. This results in fragmented C_{org} stock maps. It is apparent that C_{org} stocks in soils are generally classified higher and with less divisions than those in vegetation. A comparison of the maps shows that forest areas are indirectly classified by both approaches due to higher concentrations of vegetation C_{org}. Stocks over 100 Mg C_{org} ha^{-1} are found mainly in areas recognizable as forests in the satellite imagery.

Comparing outcomes from SOM and kNN, we can identify a more distinct spatial pattern in SOM classifications. This is evident in the classification of soil C_{org}, where, while kNN classifications show a highly homogeneous surface with tiny differences, SOM classifications exhibit clear differences between forested areas and meadows and reed beds.

Comparing the maps generated by pure remote sensing data and the combination of remote sensing and auxiliary data, we can observe greater details for classifications with combined data, which is especially visible for classifications of total C_{org} stocks, where the range of possible values is much more highlighted.

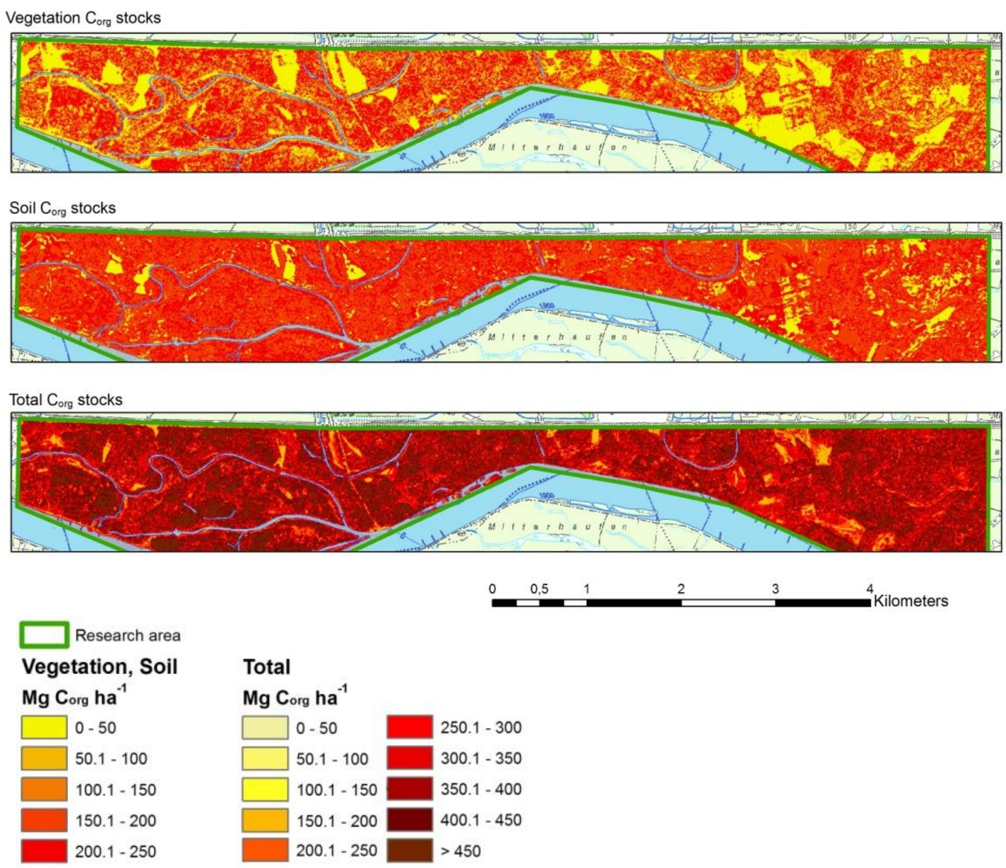

Fig.IV-2a: C_{org} stocks in Vegetation, Soil and total, calculated by SOM method based on RapidEye

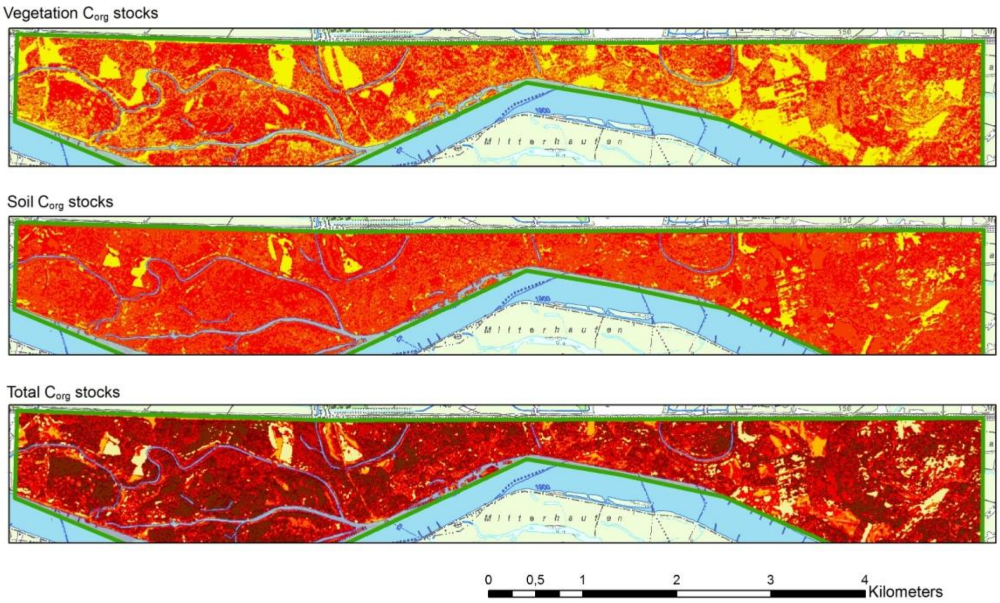

Fig. IV-2b: C_{org} stocks in Vegetation, Soil and total, calculated by SOM method based on RapidEye and additional data (legend: see Fig. IV-2a)

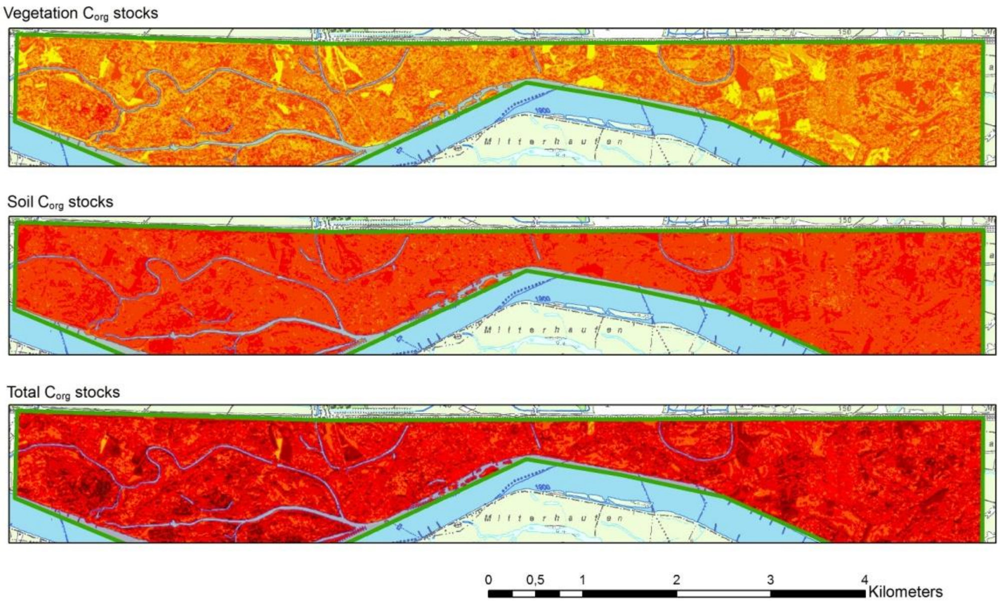

Fig. IV-1a: C_{org} stocks in Vegetation, Soil and total, calculated by kNN method based on RapidEye (legend: see Fig. IV-2a)

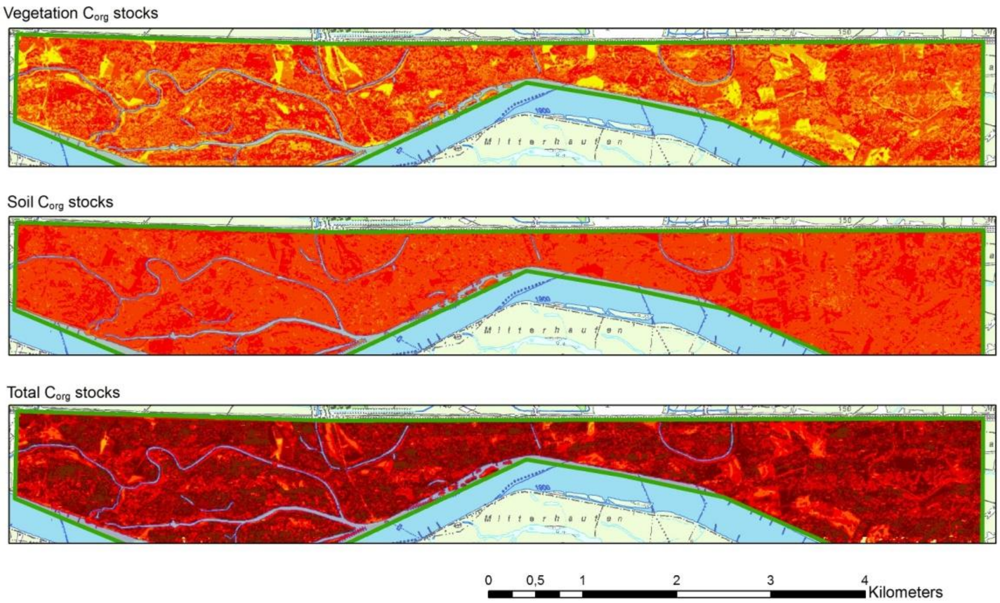

Fig. IV-3b: C_{org} stocks in Vegetation, Soil and total, calculated by kNN method based on RapidEye and additional data (legend: see Fig. IV-2a)

The review of the SOM- and kNN approach was complemented by a comparison of statistical estimates for the test area. Table 2 shows the results of C_{org} stock provided by the SOM- and kNN- approaches. The results presented are based on the entire set of point estimates used for producing the test area maps. The differences between SOM- and kNN-based estimates range between 3.87 Mg ha^{-1} (for soil and total C_{org} stocks) and 46.14 Mg ha^{-1} (for total C_{org} stocks).

The kNN approach with the RapidEye dataset provides generally higher values in comparison to the SOM approach. The differences between the two approaches including additional data do not indicate a one-sided bias structure. Estimations of vegetation, soil and total C_{org} stocks are independent from each other, so vegetation and soil C_{org} stocks do not necessarily add up to total C_{org} stocks. In order to analyze the accuracy in comparison with the field data, we have to consider the error estimates. In general, values are slightly lower than the results of Cierjacks et al. (2010), and Suchenwirth et al. (2012).

Table IV-2. SOM- and kNN-based estimates for vegetation, soil and total C_{org} stocks in the Danube Floodplain

Dataset	Approach	Vegetation C_{org}: Mg C_{org} in total study area (Mg C ha^{-1})	Soil C_{org}: Mg C_{org} in total study area (Mg C ha^{-1})	Total C_{org}: Mg C_{org} in total study area (Mg C ha^{-1})
RapidEye	SOM	144043.49 (127.47)	198390.17 (175.57)	393735.41 (348.44)
	kNN	158791.28 (140.52)	238362.66 (210.94)	398114.52 (352.31)
RapidEye +altitude +MGW + distance	SOM	168056.05 (148.72)	198635.46 (175.78)	389228.63 (344.45)
	kNN	122856.37 (108.72)	203001.62 (179.65)	337092.95 (298.31)

3.2 Error estimates

In order to evaluate their performance, SOM and kNN point estimates that coincided with terrestrial survey plots were each used to carry out an error analysis for the estimation of the average growing stock per unit area (Table 3). The values assessed on the field plots served as control values. In the research area, in total 104 terrestrial plots were available for calculating the bias and RMSE, with normalized values for bias, RMSE and % RMSE.

For vegetation C_{org} measurements, the approaches had positive and negative biases (SOM: -4.26; 11.41; kNN: 39.52; -0.94).

In soil C_{org} assessments, SOM approaches yielded positive biases (3.01; 0.28), while kNN yielded positive bias for RapidEye- classification only (18.22; -4.28).

Both approaches had a positive and a negative bias for classifications of total C_{org} stocks (SOM: -19.90; 3.15), kNN was positive and negative (73.92; -8.23). The positive biases are higher than the negative biases. In most cases, apart from vegetation C_{org} with additional data, the kNN approach is more biased than the SOM results.

Discerning between SOM and kNN, the RMSE does not show a clear tendency. In some classifications, SOM has lower RMSE, in other cases kNN classification is more accurate. The %RMSE of the SOM-approach ranged between 56.42% and 146.99% and had a smaller range than for the kNN approach (40.79-158.32%). Biases are smaller for SOM classifications.

Table IV-3. Error estimates from SOM and kNN for vegetation, soil and total C_{org} stocks in the Danube Floodplain (SOM: start distance δ_{start} of 100000 and an end distance δ_{end} of 100, and have five iterations (t_{max}).; kNN k: 5; Euclidean distance $d_{(x1,x2)}$:2; distance weight $w_{(i),p}$: 2)

Dataset	Approach	Vegetation C_{org} stocks (average 149.65 Mg C ha^{-1})			Soil C_{org} stocks (average 192.1 Mg C ha^{-1})			Total C_{org} stocks (average 361.52 Mg C ha^{-1})		
		Bias	RMSE	% RMSE	Bias	RMSE	% RMSE	Bias	RMSE	% RMSE
RapidEye	SOM	-4.26	229.12	146.99	3.01	113.26	58.99	-19.90	267.33	69.61
	kNN	39.52	177.45	158.32	18.22	85.34	48.27	73.92	210.45	72.52
RapidEye +altitude +MGW +distance	SOM	11.41	198.85	143.29	0.28	108.22	56.42	3.15	226.18	63.11
	kNN	-0.94	182.15	118.46	-4.28	81.26	40.79	-8.23	196.66	52.67

Regarding the use of additional geodata, there is a lower RMSE for the classifications based on additional geodata, than for classifications based on pure RapidEye datasets. Especially for kNN classifications, the error is notably lower (8- 40%), whereas for SOM classifications, errors are only slightly lower (2-6%). Apart from the SOM approach on vegetation C_{org}, the bias is smaller for classifications using additional geodata.

4 Discussion and Conclusion

SOM and kNN have been applied for spatially explicit estimates of C_{org} stocks above and below ground in riparian forest zones. Terrestrial measurements and satellite data as well as additional geodata served as input data to carry out the learning and training process of a neural network. Results show that both methods, SOM and kNN, are able to mimic spatial patterns of vegetation, soil and total C_{org} stocks. Both provide spatially detailed estimates, only limited by the spatial resolution of the used imagery. However, the SOM approach supplies a far more distinct spatial pattern of the C_{org} distribution, while the kNN method results in rather averaged, homogeneous patterns.

In comparison with existing estimations, values for total C_{org} stocks are comparable to the results of Cierjacks et al. (2010), but are considerably lower than results classified by Suchenwirth et al. (2012). This would support the assumption that the SOM and kNN methods can substitute a field-based calculation (such as Cierjacks et al. 2010) better than a mere classification of vegetation types to estimate C_{org} stocks (such as Suchenwirth et al. 2012).

In detail, classifications of vegetation C_{org} stocks (both kNN and SOM, based on satellite sensors and additional data) have an apparently higher RMSE than classifications of soil C_{org} stocks and total C_{org} stocks. This is not the case for bias, which is in several cases higher for the classification of total C_{org} stocks. The RMSE of our estimations of soil (ranging from 40.79 to 58.99%) and total C_{org} (ranging between 52.67 and 72.52%) are in line with results of other studies using SOM (Fuchs et al. 2009; Stümer et al. 2010; Tuominen and Pekkarinen 2005) or kNN (McInerney and Nieuwenhuis 2009) to classify C_{org}, where values range between 44.85 and 70.49%. RMSE for vegetation C_{org} is higher (118.46 to 158.32%) in our classification.

The reason for the higher RMSE within the vegetation classification can be explained by the more complex natural structure and the resulting diversity inside riparian forest vegetation and the national park area in comparison to the structure of conventional working forests and timberland monocultures.

Comparing the results of kNN and SOM-based classifications, we can find that both provide similar results. kNN has smaller RSME estimates for soil C_{org} and for classifications of vegetation and total C_{org} stocks based on RapidEye and additional data, yet has higher RMSE estimates for classifications of vegetation and total C_{org} based solely on RapidEye. In general, we can state that kNN have a better performance regarding RMSE than SOM estimates, which is also coincident with Stümer et al. (2010). Contrarily, the kNN results are much more biased than the results of the SOM. Moreover, the visual impressions of the SOM-generated maps are more distinct; this differs our results from Stümer's results in 2010 who found a smaller bias and a higher level of detail of structures such as roads, planting rows and stand boundaries for kNN results. In conclusion, we can state that in our study kNN provides on average better estimates for C_{org}, but just within the restricted range of values within the test area. For a possible transfer of the method to other regions the less biased SOM approach might be the preferable algorithm.

Both presented approaches provide greater spatial detail than comparable classifications based on object-based image analysis (OBIA) of the area (Suchenwirth et al. 2013; Wagner-Lücker et al. 2013). Even though OBIAs have the advantage of working with distinct image objects, the process of segmentation can be challenging and even misleading for continuous objects, such as natural vegetation or ecosystems in an intricate floodplain area, and may thus be a source of error, as stated by Rocchini et al (2013). In comparison to other remote sensing techniques such as Principal Component Analysis (PCA), SOM has shown demonstrably better performance (Astel et al. 2007; Klobucar and Subasic 2012).

However, some issues may yet occur when using SOMs, as they are not self-explanatory and are generally treated as a “black box” due to unknown weights and the non-linearity of the activation functions. While Hsu and Halgamuge (2003) mention the obliqueness of rectangular lattices as major sources of topographic errors, Klobucar and Subasic (2012) count among the problems of SOM the repeatability of the method. The time needed to calibrate and validate neural networks should not be underrated and the decision about the termination of the learning process may be difficult.

The application of kNN did not impose greater issues, and their applicability to forest and biomass/C_{org} inventions has often been proven, even though the majority of studies have worked with Landsat data which have a lower spatial resolution and thus provide coarser imagery; the application of VHSR data is not so common yet, while the combination with auxiliary geodata has barely been used. Among the commonly mentioned disadvantages of kNN are excessive validations of each distance, the sensitiveness towards irrelevant or noisy attributes as well as towards unbalanced datasets. For our application however, it has served as a valuable alternative to the application of SOMs.

The use of additional geodata improved the performance of both algorithms, in that all RMSE values improved, as well as the bias (with the exception of SOM classification on vegetation C_{org}). Especially for kNN, the notable improvement of RMSE underlines the importance of combined data approaches. This also confirms previous findings of Suchenwirth et al. (2012).

Comparing the study's method with previous methods to quantify C_{org} in the Danube floodplain, our study uses a 'direct remote sensing' approach including machine learning (Goetz et al. 2009), while Suchenwirth et al. (2012) used a 'stratify and multiply' approach, and Suchenwirth et al. (accepted) used a 'combine and assign' approach (Goetz et al. 2009).

In general, we see our study as a contribution to high-detailed C_{org} analyses and large scale maps of intricate ecosystems such as riparian forests or similar wetland areas with interfering aquatic and terrestrial environments, as they impede ground survey measurements through their restricted accessibility and require advanced methods to estimate biomass and organic carbon, such as remote sensing or machine learning.

For prospective applications, we envisage comparable studies with extensions of start distances and numbers of iterations, as the focus of this study lays on the comparative estimations of C_{org} stocks in vegetation, soil, and total, with varying parameters and with two methods, and not with different settings of SOM and kNN.

Another improvement for future research on the estimation of C_{org} with remote sensing data may be to include imagery with an even higher spatial resolution, as e.g. provided by the commercial sensors Ikonos (1 m), QuickBird 2 (0.64 m), or Worldview (0.5 m). The inclusion of further datasets such as surface models including tree height, e.g. based on Lidar, and other auxiliary data are able to additionally improve the performance.

Acknowledgements

This study was funded by the German Research Foundation (DFG; project number KL 2215/2-1 and KL 2215/2-2). We acknowledge the DLR for RapidEye images as part of the RapidEye Science Archive (proposal 454). We would like to thank the staff of the Danube Floodplain National Park as well as Friederike Lang, Arne Cierjacks and Isaak Rieger for the provision of data. We would also like to thank Kyle Pipkins for checking our English.

Chapter V: Synthesis

The synthesis contains citations from the already published papers or submitted manuscripts in chapter II (Suchenwirth et al. 2012), chapter III (Suchenwirth et al. 2013)**, and chapter IV (Suchenwirth et al. submitted)**. Respective sections are marked in grey with subsequent asterisks.*

The overarching objective of this thesis was the assessment of the capability of very high spatial resolution imagery and auxiliary geodata with the aim to model organic carbon (C_{org}) stocks in floodplains. Based on the research questions posed at the introduction of the thesis, this chapter discusses research results and gives recommendations for future investigations.

1 Main Conclusions

The first overall research objective of this thesis was to develop and evaluate methods to integrate existing very high resolution satellite imagery and additional geodata to either a knowledge-based or machine-learning based classification process with the aim to model and map the spatial distribution of carbon stocks in vegetation and soil. The applicability of datasets could be reached, however some limitations of the approaches became visible.

This objective was approached by applying different techniques to estimate the carbon stocks ranging from indirect to more direct methods and to estimate the quality of the resulting classifications.

Research question 1: *Which methods can be sufficiently applied to model C_{org} stocks in soil and vegetation in floodplains by remote sensing and additional geodata?*

The question aims at the feasibility of the applied methods. In a nutshell, all of the used approaches have their benefits and disadvantages when being applied for large-scale maps of floodplain C_{org}. Based on the methods used, the approaches applied in the chapters can be defined and compared to the ones described by Goetz et al. 2009 including their specific features.

The estimation of C_{org} stocks based on vegetation types classified with object-based image analysis (OBIA) as presented in chapter II can be regarded as a stratify and multiply approach, as a certain amount of C_{org} is assigned to every vegetation type class gained from the classification. In order to mitigate the described hitch of wide range of biomass/ C_{org} stock, the results were calculated by a Monte Carlo simulation with a 1000fold repetition. It is one of the disadvantages that the accuracy of the C_{org} map depends on the accuracy of the classification of the specific vegetation types. Inherent problems of vegetation classification have already been described by Rocchini et al (2013). Yet, this approach is easy to understand and can be transferred to other research areas, as long as there are data on C_{org} stocks in specific vegetation types available and an accuracy assessment on vegetation types is feasible. Accuracy assessment is perceived by the often applied measures of overall accuracy, kappa value and producer's / user's accuracy.

The estimation of C_{org} stocks based on quantile classes generated by a CART approach using OBIA, as presented in chapter III is comparable to the combine and assign approach, but also shows signs of a direct remote sensing approach (Goetz et al. 2009). It uses a wider variety of data including all sensor channels, indices, texture parameters and information derived from the additional datasets, along with finer spatial aggregation units. It is one of the main achievements of this approach to use this wide variety of data along with the open generation of rule sets by CART; a point of critique may be that this approach applied quantile classes instead of clustered classes; the representativeness of these quantile classes remains a subject to debates. Yet, the application of quantiles enables an easily reproducible data schema. The provision of spatially consistent datasets was no problem for the research area but may become a challenge for larger research areas, not only for remote sensing data, but aso for auxiliary datasets. Accuracy assessment is perceived by overall accuracy, location and quality disagreement, as suggested by Pontius and Millones (2011).

Chapter IV can be described as a direct remote sensing approach, as it applies machine learning, in this case SOM and kNN, with a direct extension of satellite measurements (and additional geodata) to the map. The field measurements are used directly through iterative repeated data analysis and accordingly develops an optimal rule set for the classification. As soon as rules are optimized for training data, they are applied for the entire dataset (satellite imagery and additional geodata). Results are issued as direct values, there are no class boundaries, which helps to give a very clear result for each spatial unit. This is one of the main benefits of this method over the previous ones. Even the area size does not provide difficulties. Yet, the application of SOM has the disadvantage of being a black box with unknown weights and the non-linearity of activation functions. Classification accuracy is slightly trickier to measure and to define in comparison to the previous approaches that used overall accuracy measures besides Kappa which are based on classes, such as vegetation types

(*) or quantile classes (**). RMSE and bias values however provide a good alternative to calculate the quality of results.

Table V-1: Strengths and weaknesses of applied techniques

Applied technique	Strengths	Weaknesses
Estimation of C_{org} vegetation type classified with OBIA (Methods: OBIA + expansion of results using a Monte Carlo simulation) *	-simple applicability -simple to understand -Rules are modifiable	-Dependency on vegetation type classification -Preliminary data on C_{org} in vegetation types needed
Classification of quantile classes of C_{org} in vegetation, soil and total (Methods: OBIA + CART)**	-Enables the use of various datasets -Searches the datasets for viable rules -Provides rule sets (decision trees) -higher accuracy for vegetation C_{org}	-Difficult repeatability -Single CART models may be unstable -CART interpretation may become confusing with increasing number of branches -the use of quantile classes does not automatically represent clusters -lower accuracy for soil C_{org}
Estimation of C_{org} in vegetation, soil and total (Method: SOM)***	-high repeatability -clear, unambiguous results for each spatial unit	-SOM is a black-box algorithm, rules are not presented -High calculation times -Accuracy assessment based on RMSE errors differs from error matrix used for classifications
Estimation of C_{org} in vegetation, soil and total (Method: kNN)***	-simple applicability -frequent usage -higher accuracy for soil C_{org}	-lazy learning algorithm -reduction on spatial relations

The findings of this thesis suggest that there is a potential for remote sensing based estimation of C_{org} stocks in floodplains. The capabilities of the applied methods can generally be regarded as sufficient for the purpose of mapping and modeling C_{org} stocks in riparian vegetation and soils. Based on the presented methods, an extension of C_{org} monitoring concepts based on remote sensing and additional geodata to other floodplain areas seems practicable. The applied approaches present a methodological step towards the remote sensing based monitoring of C_{org} stocks and contributes to establishment of conceptual frameworks. Yet, the floodplain environment still causes difficulties for remote sensing applications, due to their heterogeneous nature, resulting in decreased classification accuracy. Further research needs to evaluate the operational relevance of differential strategies in greater detail. In general, the advantages and disadvantages are compared in table V-1.

Research question 2: *How can additional geodata be included and their significance for the model be measured?*

The implementation of multiple data with different background for the models of C_{org} is one of the core ideas of this thesis. It is a challenge to combine geographic datasets (with different spatial reference systems, acquisition dates, contents, formats and data source reliability). The integration of data requires a high volume of preliminary work, among the different work steps are the definition and unification of a spatial reference system including geographic datum and projection, resulting in the reprojection of datasets with different spatial reference.

The majority of the used software packages are able to process both raster and vector formats. For those cases when a software package in use is not able to work on raster or vector data, conversion tools provided by ESRI ArcGIS are used. The chapters show that the inclusion of multiple data in the classification process is possible and provide enhanced results.

In chapter II, Ikonos satellite data formed the basic layer for the OBIA. Additional data were added as thematic layers, and were weighted equally in the applied eCognition software. It shows a comparison of different classification sets; on the one hand the application of pure Ikonos data, on the other hand the use of Ikonos data combined with knowledge-based classification rules, derived from the digital elevation model, distance to river and specific rules regarding area size. The classification set including all knowledge-based rules resulted in the highest overall accuracy and Kappa value, albeit values for producer's and user's accuracy and kappa per class are partly higher for differing classification sets. The classification based on modeling the total C_{org} stocks without a differentiation for soil and vegetation.

In chapter III, the same eCognition software was used, however including a second satellite dataset (RapidEye) and more additional data**. Like in chapter II, an equal weighting of layers was applied. The chapter shows a different approach to chapter II, as it compares classifications based on different class numbers, while using the same set of parameters for all classifications; yet, the usage of parameters is analyzed in detail. It turns out that for the classifications the importance of parameters is volatile, as parameters have been used for all classifications. Differences become visible for the classifications of C_{org} stocks in soil, vegetation and total C_{org} stocks. While for the classification of C_{org} in vegetation remote sensing parameters were most important.

In chapter IV, the comparison of raster-based SOM- and kNN approaches, all data were resampled to a same cell size of 5 meters, in order to provide comparable results between the approaches. The paper compares approaches of SOM and kNN, for datasets based on pure remote sensing parameters, and combined datasets of remote sensing and additional geodata (DEM, medium ground water level and distance to river). Like in chapter II and III, the use of additional data improved the model, resulting on lower RMSEs and biases for both SOM and kNN.

In order to examine the impact of individual datasets and geofactors for the estimation of C_{org} stocks, the importance of the additional datasets was to be assessed. Chapter II and IV compared results generated by classification sets with pure remote sensing data and classification sets with combined datasets of remote sensing and additional geodata; however, the comparison did not give quantitative results on the specific significance of included individual parameters from each individual applied dataset.

The importance of the use of additional datasets can therefore only be derived indirectly, by the comparison of classification results and their accuracies and error values. The overall accuracy of vegetation type classification* improved by 0.1 from 0.6 to 0.7 when using additional geodata, while based on individual classes the producer accuracy was able to improve by 0.46, and the kappa per class value by 0.36 respectively.

In comparison, for the applications on SOM and kNN***, the RMSE values decreased when using additional geodata (SOM: 6%; kNN: 40%). It indicates an improved classification accuracy, and confirms the assumption that the additional datasets support the modeling of C_{org}.

Chapter III** shows an application of data mining; while all classifications were based on the same datasets. The input of every dataset was numerically specified. The method to exposit parameter relevance has hardly been described in literature, except by Erasmi et al. (2013). Differences become visible for the classifications of C_{org} stocks in soil, vegetation and total C_{org} stocks. While for the classification of C_{org} in vegetation, remote sensing parameters were most important along with groundwater features, the use of historical, groundwater and digital elevation model were specific for the classification of C_{org} in soil. Even though parameter importance is varying, all additional parameters contribute to the classification success of the applied methods. The analysis of individual parameter and/or dataset importance or relevance remains difficult. An increased repetition of analyses might give an enhanced insight into the parameter relevance.

Research question 3: *What are the specific advantages of automated C_{org} mapping on local scale for operational monitoring purposes?*

Approaches of automated C_{org} mapping on local scale has benefits, compared to estimations at small scale, for instance on national base, which tend to be coarse and dissatisfying for local needs. C_{org} stocks modeled on a local base feature not only a higher spatial resolution and specific details, but offer also more information on the connections between attributes of soil, vegetation, morphology, hydrology, historical background, and the specific C_{org} content.

But there are also advantages towards terrestrial measurements. Data measured directly in a field-based campaign may be more accurate and provide more detailed information. Yet, the acquisition efforts, regarding precious resources such as time, equipment, human labor and consequently financial means, are certainly higher and more demanding, compared to an automated model. The use spectral information from satellite sensors as well as auxiliary geodata is able to show up new up existing interrelations that may not be visible in the field.

Furthermore, once the model is well-trained, it provides a more objective way to generate information C_{org} stocks.

Conclusively, the applied methods can be seen as a auspicious novelty to bridge the gap between small-scale approaches on national or international base, such as coarse national or continental estimations, and exceedingly detailed terrestrial studies. While local terrestrial studies provide a micro-level, national or international models provide a macro-level. This thesis shows up solutions at a medium level, i.e. for planning and monitoring purposes of protected areas such as national parks, FH areas, within national and international legislations.

2 Future research

In this thesis the application of various methods to model carbon stocks in floodplains was demonstrated. Several issues interesting for follow-up research beyond the scope of this research evolved during the course of this thesis. Amongst them, the integration of further datasets, the transferability of methods to other areas, and the link to ecosystem services including biodiversity, and the question of floodplain/wetland protection shall be discussed.

2.1 Technical issues and integration of further datasets

In order to transfer the method, several issues shall be kept in mind. One central issue is the availability of data, hardware and software:

Data:

-Besides the provision of satellite imagery with appropriate spatial resolution, the availability and spatial contiguity of additional data is crucial as it can e.g. give important hints for the existence of historic riverbeds and resulting differences in C_{org} stocks.

-A good geodatabase requires specific conditions from the area/ state in which the carbon model is to be established. For countries with good public geodata, especially regarding the availability of historical maps, the pursuit for additional geodata is probably easier to handle and organize than for countries or regions with a restricted availability and public data policy. The establishment of C_{org} modeling systems developing countries or countries in transition may impose problems of data availability.

-A fundamental base for our study were the terrestrial survey data collected in 2008 and the extension of data in 2010. As they have been carefully sampled and collected over the whole area covering all habitat types, their quality is invaluable. They provided the ground truth for the calibration and validation of classifications and models. For similar studies a comparable database is necessary.

A further question of transferability is the provision of hardware and software infrastructure. As the demand of disk space by datasets can increase with spatial resolution of satellite imagery, the size of research area, but also the general amount of datasets can be very demanding. The computer system at use has the following specifications: 12 GB RAM, Intel Xeon CPU 2.53 GHz, 64 Bit Operating system, and a 931 GB hard disk. As operating system, a windows 7 professional system was used. Software components consist of ESRI ArcGIS 10.0, ERDAS Imagine 2011, eCognition developer 64 8.7.1, and various additional software packages.

For additional datasets, it might be arguable to include the following datasets into a carbon model of floodplains:

-Digital surface model including tree heights: There are plenty of studies showing the successful modeling of carbon stocks using detailed surface models. They are used to derive information on tree height (Omasa et al. 2003; Balzter et al. 2007). As these data were not available for this study, the estimation of tree age and thus resulting more specific information on C_{org} stocks was impeded. According to Patenaude et al. 2005 (2005), radar data are most appropriate to estimate forest C_{org} stocks.

-Satellite imagery of flooding: The available satellite imagery was taken during non-flooding conditions, i.e. while river had a mean water (MW) level . It could be interesting to analyze imagery taken during floods, such as the floods in Central Europe in 2002 or 2013, and to investigate flooding effects and possible correlations with C_{org} distribution.

-Use of hyperspectral data: Sensor systems being able to receive a wide spectrum of many, closely attached wavelengths are used to calculate vegetation indices in order to derive information on plant health; they may as well contain information on biomass and C_{org}; there are studies using hyperspectral data for C_{org} monitoring in soils (Gomez et al. 2008; Jaber et al. 2011; Yang and Li 2013), while Adam et al (2010) show the use of hyperspectral data for wetlands in general.

2.2 Transferability to other floodplain/ wetland areas

As floodplain and wetland areas and riparian forests are to be found not only in Central Europe and other temperate climate zones, but all over the world, and have high relevance due to their C_{org} stocks, the creation of large-scale floodplain maps in different environments is worthwhile from a scientific point of view. Even for a general public, detailed maps of floodplains and their specific hotspots may be of interest e.g. for touristic purposes.

Regarding the environment, our research area reflects a floodplain area in comparatively pristine state, within a Central European environment, i.e. a moderate climate zone, temperate forests. Therefore, the specific and singular classification rules developed by knowledge -base (chapter II) or the use of CART (chapter III) may be challenging to transfer to areas with different habitats; the SOM and kNN approaches develop rules, yet they are a black box, and the user has no insight. Even though specific rules cannot be transferred to other datasets research areas or habitats, the methods in general can be transferred, if accordant or comparable datasets are available.

Along the Danube river and its feeder rivers, there is an entire network of protected areas such as national parks (NP) and other reserved and protected areas, in the various adjacent countries along the river and within the Danube's catchment basin. The network starts with the re-established riparian forest close to the German cities of Neuburg and Ingolstadt, and continues with the Danube Floodplain NP in Austria. It is followed in Slovakia by the Záhorie as well as the Dunajské Luhy Protected Landscape Area, while in Hungary, the Duna-Ipoly NP, the Fertö-Hansag NP and the Duna-Dráva NP are part of the river system's natural resources. Further downstream, the Lonjsko Polje Nature Park, Kopački rit Nature Park in Croatia, the Gorne Poduavlje Nature Reserve and Đerdap National Park (both Serbia), the nature parks of Persina and Rusenski Lom as well as the Kalimok Brushlen Site in Bulgaria, and at the estuary into the Black Sea, the Danube Delta Biosphere Reserve in Romania completes the system of the Danube's protected areas.

In Germany, besides the Danube river's basin and floodplain systems, the riparian areas of the Elbe, Weser, Rhine, and Oder river are of major importance. However most floodplain systems have been subject to major losses, according to BfN (2009). For instance, the straightening of the Upper Rhine was initiated by Gottfried Tulla in the 19th century; other floodplain ecosystems were changed in the 20th century. Few natural floodplains have remained intact. Among the protected wetland areas in Central Europe, the biosphere reserves of Spreewald, and along the Elbe river, or the Müritz NP, the Lower Oder Valley NP in Germany and the Thayatal NP in Austria, to name a few, may be worth for comparative research, also for their carbon sequestration potential.

2.3 Integration of further ecosystem services

As mentioned in the chapter I, floodplains have a high importance not only for C_{org} sequestration, but also for other ecosystem services, the concepts can be expanded and other ecosystem services can be integrated. The concept of ecosystem services has been frequently discussed in recent years. The Millennium Ecosystem Assessment (2005) differentiates between 4 service categories (-provisioning, e.g. food, water, fiber, fuel; -regulating, e.g. climate regulation, water, disease; -cultural, e.g. spiritual, aesthetic, recreation, education; -and supporting, e.g. primary production, soil formation), while Constanza et al. (1997) itemize 17 ecosystem services and goods (gas regulation, climate regulation, disturbance regulation, water regulation, water supply, erosion control and sediment retention, soil formation, nutrient cycling, waste treatment, pollination, biological control, refugia, food production, raw materials, genetic resources, recreation, cultural). The associated monetization, i.e. conversion of ecosystems services and goods into monetary categories is a constant subject of research but also internal and external critics.

According to the data of Constanza et al. (1997), floodplains rank second among ecosystem services only after estuaries; gas regulation by floodplains (apparently including carbon sequestration) is however financially estimated much lower than the importance of riparian wetlands for habitat/refugia, disturbance regulation, water supply, or waste treatment or cultural purposes. The effect on climate regulation is not even taken into account. Yet, one might assume that during the study the crucial significance of floodplains for carbon sequestration and thus climate regulation was not honored sufficiently in the given time 1997. Even so the ecosystem service habitat/refugia is mentioned in the text, a link towards biodiversity is missing.

Biodiversity as a concept can be defined as the degree of variation of life forms inside a certain species, ecosystem, biome or even planet. There are traditionally three sources of biodiversity, which are species biodiversity, ecosystem biodiversity and genetic biodiversity; molecular biodiversity as a fourth level is being discussed (Campbell 2003). Biodiversity can be defined by statistical values, such as alpha, beta, gamma etc. biodiversity (Whittaker 1972; Diamond 1988); Stoms and Estes (1993) transferred this knowledge already to remote sensing applications for biodiversity.
There is a linkage between high biodiversity and high C_{org} stocks in many ecosystems and efforts to maintain areas with high biodiversity and at the same time high C_{org} stocks may be combined (Huston and Marland 2003; Strassburg et al. 2010) there are even map compilations atlases on global and international scale by (Groombridge and Jenkins 2002; UNEP-WCMC 2008).

Biodiversity in floodplains has been described in detail and in its importance in several studies (Junk et al. 2006). Local studies have shown the tight connection of biodiversity in floodplains and increased carbon stocks, e.g. for Australia (Shiel et al. 1998; Horner et al. 2010), Brazil (Agostinho et al. 2005; Ferretti and de Britez 2006), or globally (Schindler et al. 2013; Mitra et al. 2005); for the Danube floodplain in Austria, the studies of Tockner et al. (1998) and Ward et al. (1999) have already shown the high biodiversity.

As the database for this project, especially the field survey data, contains information on plant biodiversity, especially tree species, it can provide information of spatial and statistical models for the distribution C_{org} stocks and biodiversity.

2.4 Extension to wetland protection and floodplain restoration projects

While the extension of research questions towards biodiversity seems clear and feasible, there will be new research areas and material for the upcoming future. Planned extensions and or restorations of floodplains providing a protection function as inundation zones can be researched and integrated into carbon modeling processes.

A number of flood events, some of them in recent years (e.g. August 2002, April 2006, June 2013) in this area demonstrated again the imperative necessities and the need for sustainable flood protection concepts in Central Europe (Petrow et al. 2006) and thus to maintain and restore floodplain systems, that have been straightened during the 19^{th} and 20^{th} century, particularly in Germany, but also other European countries (BfN 2009).

In the recent years there have been various studies describing tasks and targets for the restoration of floodplains (Pedroli et al. 2002; Hale and Adams 2007; Schindler et al. 2013) Recently, a promising project has been launched by the Bavarian Water Authority in 2010 in the Upper Danube between the Bavarian cities of Neuburg and Ingolstadt in a former riparian forest area that had been drained by a fluvial power plant several decades ago. In order to re-establish a floodplain habitat and to bring back water and sediment dynamics into the area, a new river side channel was dug up, partly along existing old oxbow lakes, partly through erosion. Additional controlled ecological floodings help tom establish a new floodplain environment. The area is subject to intensive monitoring such as geomorphology (Stammel et al. 2011) and biodiversity. It might be a promising approach to combine the research on C_{org} sequestration potential in the re-established floodplain system with the monitoring of floodplain restoration projects, and to assess the naturalness of the river floodplain systems.

In general, the need for an integrated floodplain management has been recognized, as they form one of the most diverse, dynamic and productive, but also one of the most threatened global ecosystems. This thesis contribute to a better understanding of floodplains and their C_{org} storage capacities and properties.

References

Achard F, DeFries R, Eva H, Hansen M, Mayaux P, Stibig HJ (2007) Pan-tropical monitoring of deforestation. Environmental Research Letters 2 (4): 045022

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. doi: 10.1007/s11273-009-9169-z

Adis J, Junk WJ (2002) Terrestrial invertebrates inhabiting lowland river floodplains of Central Amazonia and Central Europe: a review. Freshwater Biology 47 (4): 711-731. doi: 10.1046/j.1365-2427.2002.00892.x

Adamczyk JJ, Kurzac M, Park YS, Kruk A (2013) Application of a Kohonen's self-organizing map for evaluation of long-term changes in forest vegetation. J Veg Sci 24 (2): 405-414. doi: 10.1111/j.1654-1103.2012.01468.x

Agostinho AA, Thomaz SM, Gomes LC (2005) Conservation of the Biodiversity of Brazil's Inland Waters. Conservación de la Biodiversidad de las Aguas Interiores de Brasil. Conservation Biology 19 (3): 646-652. doi: 10.1111/j.1523-1739.2005.00701.x

Arribas-Bel D, Nijkamp P, Scholten H (2011) Multidimensional urban sprawl in Europe: A self-organizing map approach. Computers, Environment and Urban Systems 35 (4): 263-275. doi: j.compenvurbsys.2010.10.002

Asner GP, Powell GVN, Mascaro J, Knapp DE, Clark JK, Jacobson J, Kennedy-Bowdoin T, Balaji A, Paez-Acosta G, Victoria E, Secada L, Valqui M, Hughes RF (2010) High-resolution forest carbon stocks and emissions in the Amazon. Proceedings of the National Academy of Sciences 107 (38): 16738-16742. doi: 10.1073/pnas.1004875107

Astel A, Tsakovski S, Barbieri P, Simeonov V (2007) Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Research 41 (19): 4566-4578. doi: 10.1016/j.watres.2007.06.030

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. In: Shiyomi M, Kawahata H, Koizumi H, Tsuda A, Awaya Y (eds) Global Environmental Change in the Ocean and on Land. Terrapub, Tokyo, pp 425-440

Baatz M, Schäpe A (2000) Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. In: Strobl J (ed) Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, 2000. Herbert Wichmann Verlag, Karlsruhe, Karlsruhe, pp 12-23

Baccini A, Goetz SJ, Walker WS, Laporte NT, Sun M, Sulla-Menashe D, Hackler J, Beck PSA, Dubayah R, Friedl MA, Samanta S, Houghton RA (2012) Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Clim Change 2 (3): 182-185. doi: 10.1038/nclimate1354

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. doi: 10.1016/j.foreco.2005.05.059

Balzter H, Rowland CS, Saich P (2007) Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. Remote Sensing of Environment 108 (3): 224-239. doi: http://dx.doi.org/10.1016/j.rse.2006.11.014

Baptist MJ, Penning WE, Duel H, Smits AJM, Geerling GW, Van der Lee GEM, Van Alphen JSL (2004) Assessment of the effects of cyclic floodplain rejuvenation on flood levels and biodiversity along the Rhine River. River Research and Applications 20 (3): 285-297. doi: 10.1002/rra.778

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. doi: 10.1016/j.foreco.2010.03.025

Behrens T, Scholten T (2006) Digital soil mapping in Germany—a review. Journal of Plant Nutrition and Soil Science 169 (3): 434-443. doi: 10.1002/jpln.200521962

BfN (2009) Auenzustandsbericht. Bonn

Bonn A, Hagen K, Reiche DW-v (2002) The significance of flood regimes for carabid beetle and spider communities in riparian habitats—a comparison of three major rivers in Germany. River Research and Applications 18 (1): 43-64. doi: 10.1002/rra.632

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. doi: 10.1186/1750-0680-4-7

Breijo EG, Pinatti CO, Peris RM, Fillol MA, Martinez-Manez R, Camino JS (2013) TNT detection using a voltammetric electronic tongue based on neural networks. Sens Actuator A-Phys 192: 1-8. doi: 10.1016/j.sna.2012.11.038

Breiman L (2001) Random Forests. Machine Learning 45 (1): 5-32. doi: 10.1023/a:1010933404324

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA

Busse LB, 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. doi: 10.1016/S0075-9511(01)80033-5

BVEL (Bundesministerium für Verbraucherschutz EuL (2005) Gesunder Boden – gesunder Wald. Die zweite Bodenzustandserhebung im Wald 2006 - 2008. Bonn

Campbell AK (2003) Save those molecules! Molecular biodiversity and life*. Journal of Applied Ecology 40 (2): 193-203. doi: 10.1046/j.1365-2664.2003.00803.x

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. doi: 10.1016/j.foreco.2010.07.053

Chubey MS, Franklin SE, Wulder MA (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. doi: 10.1002/jpln.200900209

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. doi: 10.1002/rra.1409

Conchedda G, Durieux L, Mayaux P (2008) An object-based method for mapping and change analysis in mangrove ecosystems. ISPRS Journal of Photogrammetry and Remote Sensing 63 (5): 578-589. doi: 10.1016/j.isprsjprs.2008.04.002

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37 (1): 35-46. doi: 10.1016/0034-4257(91)90048-B

Costanza R, d'Arge R, Groot Rd, Farber S, Grasso M, Hannon B, Limburg K, Naeem S, O'Neill RV, Paruelo J, Raskin RG, Sutton P, Belt Mvd (1997) The value of the world's ecosystem services and natural capital. Nature 387: 253-261

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. doi: 10.1016/S0176-1617(99)80314-9

De'ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81 (11): 3178-3192. doi: 10.1890/0012-9658(2000)081[3178:cartap]2.0.co;2

Deering DW, Rouse JW, Haas RH, Schell JA (1975) Measuring "forage production" of grazing units from Landsat MSS data. Proceedings of the 10th International Symposium Remote Sensing of Environment II: 1169-1178

Dillabaugh KA, King DJ (2008) Riparian marshland composition and biomass mapping using Ikonos imagery. Canadian Journal of Remote Sensing 34 (2): 143-158. doi: 10.5589/m08-011

Diamond J (1988) Factors Controlling Species Diversity: Overview and Synthesis. Annals of the Missouri Botanical Garden 75 (1): 117-129

Dissanska M, Bernier M, Payette S (2009) Object-based classification of very high resolution panchromatic images for evaluating recent change in the structure of patterned peatlands. Canadian Journal of Remote Sensing 35 (2): 189-215. doi: doi:10.5589/m09-002

Donato DC, Kauffman JB, Murdiyarso D, Kurnianto S, Stidham M, Kanninen M (2011) Mangroves among the most carbon-rich forests in the tropics. Nature Geoscience advance online publication. doi: 10.1038/ngeo1123

eCognition (2012) eCognition Developer Reference Book 8.8. Trimble Germany GmbH, Munich

Ellenberg H (1986) Vegetation Mitteleuropas mit den Alpen in ökologischer Sicht. 4th edn. Ulmer, Stuttgart, Germany

Erasmi S, Riembauer G, Westphal C (2013) Mapping habitat diversity from multi-temporal RapidEye and RADARSAT-2 data in Brandenburg, Germany. In: Borg E, Daedelow H, Johnson R (eds) 5th RESA Workshop, Neustrelitz, 20th -21st March 2013 2013. GITO Berlin, pp 75-89

Evans TL, Costa M, Telmer K, Silva TSF (2010) Using ALOS/PALSAR and RADARSAT-2 to Map Land Cover and Seasonal Inundation in the Brazilian Pantanal. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 3 (4): 560-575. doi: 10.1109/JSTARS.2010.2089042

Farid A, Goodrich DC, 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. doi: 10.1016/j.jaridenv.2007.04.010

Farley K, Kelly E, Hofstede RM (2004) Soil Organic Carbon and Water Retention after Conversion of Grasslands to Pine Plantations in the Ecuadorian Andes. Ecosystems 7 (7): 729-739. doi: 10.1007/s10021-004-0047-5

Federal Ministry of Agriculture F, Environment and Water Management (2008) Sustainable Forest Management in Austria. Austrian Forest Report 2008. Republic of Austria, Federal Ministry of Agriculture, Forestry, Environment and Water Management, Vienna

Ferretti AR, de Britez RM (2006) Ecological restoration, carbon sequestration and biodiversity conservation: The experience of the Society for Wildlife Research and

Environmental Education (SPVS) in the Atlantic Rain Forest of Southern Brazil. Journal for Nature Conservation 14 (3â€“4): 249-259. doi: http://dx.doi.org/10.1016/j.jnc.2006.04.006

Fierke MK, Kauffman JB (2005) Structural dynamics of riparian forests along a black cottonwood successional gradient. Forest Ecology and Management 215 (1-3): 149-162. doi: 10.1016/j.foreco.2005.06.014

Fitzgerald RW, Lees BG (1994) Assessing the classification accuracy of multisource remote sensing data. Remote Sensing of Environment 47 (3): 362-368. doi: 10.1016/0034-4257(94)90103-1

Förster M, Frick A, Walentowski H, Kleinschmit B (2008) Approaches to utilising QuickBird data for the monitoring of NATURA 2000 habitats. Community Ecology 9 (2): 155-168. doi: 10.1556/ComEc.9.2008.2.4

Förster M, Kleinschmit B (2008) Object-based classification of QuickBird data using ancillary information for the detection of forest types and NATURA 2000 habitats. In: Blaschke T, Lang S, Hay G (eds) Object-Based Image Analysis. Springer, Heidelberg, Berlin, New York, pp 275-290

Fuchs H, Magdon P, Kleinn C, Flessa H (2009) Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory. Remote Sensing of Environment 113 (3): 518-531. doi: 10.1016/j.rse.2008.07.017

Gallaun H, Zanchi G, Nabuurs G-J, Hengeveld G, Schardt M, Verkerk PJ (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. doi: 10.1016/j.foreco.2009.10.011

GeoEye (2009) GeoEye Product Guide.57 pages

Gibbs HK, Brown S, Niles JO, Foley JA (2007) Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environmental Research Letters 2 (4): 045023. doi: 10.1088/1748-9326/2/4/045023

Gibbs HK, Ruesch AS, Achard F, Clayton MK, Holmgren P, Ramankutty N, Foley JA (2010) Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. Proceedings of the National Academy of Sciences 107 (38): 16732-16737. doi: 10.1073/pnas.0910275107

Giese LA, Aust WM, Trettin CC, Kolka RK (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. doi: 10.1016/S0925-8574(99)00081-6

Giraudel JL, Lek S (2001) A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. Ecological Modelling 146 (1-3): 329-339. doi: 10.1016/S0304-3800(01)00324-6

Gitelson A, Merzlyak MN (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. doi: 10.1016/S0176-1617(11)81633-0

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. doi: 10.1186/1750-0680-4-2

Gomez C, Rossel RAV, McBratney AB (2008) Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma 146 (3-4): 403-411. doi: 10.1016/j.geoderma.2008.06.011

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. doi: 10.1016/j.geoderma.2008.05.008

Groombridge B, Jenkins MD (2002) World atlas of biodiversity: earth's living resources in the 21st century: Earth's living Resources in the 21st century. Prepared by UNEP World Monitoring Centre University of California Press, Berkeley, USA

Haaren C v (2004) Landschaftsplanung. Ulmer, Stuttgart

Hale BW, Adams MS (2007) Ecosystem management and the conservation of river-floodplain systems. Landscape and Urban Planning 80 (1â€“2): 23-33. doi: 10.1016/j.landurbplan.2006.05.002

Hall P, Park BU, Samworth RJ (2008) Choice of neighbor order in nearest-neighbor classification. Annals of Statistics 20 (3): 1236-1265. doi: 10.1214/aos/1176348768

Haralick RM, Shanmugam K, Dinstein Ih (1973) Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics SMC-3 (6): 610-621

Harrison AF, Howard PJA, Howard DM, Howard DC, Hornung M (1995) Carbon storage in forest soils. Forestry 68 (4): 335-348. doi: 10.1093/forestry/68.4.335

Hazlett PW, Gordon AM, Sibley PK, Buttle JM (2005) Stand carbon stocks and soil carbon and nitrogen storage for riparian and upland forests of boreal lakes in northeastern Ontario. Forest Ecology and Management 219 (1): 56-68. doi: 10.1016/j.foreco.2005.08.044

Hilker T, Coops NC, Wulder MA, Black TA, Guy RD (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. doi: 10.1016/j.scitotenv.2007.11.007

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. doi: 10.1016/j.geomorph.2007.11.015

Hofmann G, Anders S (1996) Waldökosysteme als Quellen und Senken für Kohlenstoff - Fallstudie ostdeutsche Länder. Beiträge Forstwirtschaft und Landschaftsökologie 30 (1): 9-16

Horner GJ, Baker PJ, Nally RM, Cunningham SC, Thomson JR, Hamilton F (2010) Forest structure, habitat and carbon benefits from thinning floodplain forests: Managing early stand density makes a difference. Forest Ecology and Management 259 (3): 286-293. doi: http://dx.doi.org/10.1016/j.foreco.2009.10.015

Houghton RA, Butman D, Bunn AG, Krankina ON, Schlesinger P, Stone TA (2007) Mapping Russian forest biomass with data from satellites and forest inventories. Environmental Research Letters 2 (4): 045032. doi: 10.1088/1748-9326/2/4/045032

Hsu AL, Halgamuge SK (2003) Enhancement of topology preservation and hierarchical dynamic self-organising maps for data visualisation. International Journal of Approximate Reasoning 32 (2-3): 259-279. doi: 10.1016/S0888-613X(02)00086-5

Huston MA, Marland G (2003) Carbon management and biodiversity. Journal of Environmental Management 67 (1): 77-86. doi: http://dx.doi.org/10.1016/S0301-4797(02)00190-1

IPCC (2000) Special report on land use, land-use change and forestry. Cambridge, UK

IPCC (2003) Good practice Guidance for Land Use, Land-Use Change and Forestry. Institute for Global Environmental Strategies, Japan

Jaber SM, Lant CL, Al-Qinna MI (2011) Estimating spatial variations in soil organic carbon using satellite hyperspectral data and map algebra. International Journal of Remote Sensing 32 (18): 5077-5103. doi: 10.1080/01431161.2010.494637

Jaramillo VJ, Kauffman JB, Rentería-Rodríguez L, Cummings DL, Ellingson LJ (2003) Biomass, Carbon, and Nitrogen Pools in Mexican Tropical Dry Forest Landscapes. Ecosystems 6 (7): 609-629. doi: 10.1007/s10021-002-0195-4

Kanevski M, Maignan M (2004) Analysis and Modelling of Spatial Environmental Data. EFPL Press, Lausanne

Kanevski M, Timonin V, Pozdnukhov A (2009) Machine learning algorithms for spatial data analysis and modelling. EFPL Press, Lausanne

Klobucar D, Subasic M (2012) Using self-organizing maps in the visualization and analysis of forest inventory. iForest - Biogeosciences and Forestry 5 (5): 216-223. doi: 10.3832ifor0629-005

Köhl M, Stümer W, Kenter B, Riedel T (2008) Effect of the estimation of forest management and decay of dead woody material on the reliability of carbon stock and carbon stock changes--A simulation study. Forest Ecology and Management 256 (3): 229-236. doi: 10.1016/j.foreco.2008.04.004

Kohonen T (1982) Self-organized formation of topologically correct feature maps. Biological Cybernetics 43 (1): 59-69. doi: 10.1007/bf00337288

Kollár S, Vekerdy Z, Márkus B (2011) Forest Habitat Change Dynamics in a Riparian Wetland. Procedia Environmental Sciences 7 (0): 371-376. doi: 10.1016/j.proenv.2011.07.064

Kooch Y, Hosseini SM, Zaccone C, Jalilvand H, Hojjati SM (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. doi: 10.1039/C2EM30410D

Kooistra L, Wamelink W, Schaepman-Strub G, Schaepman M, van Dobben H, Aduaka U, Batelaan O (2008) Assessing and predicting biodiversity in a floodplain ecosystem: Assimilation of net primary production derived from imaging spectrometer data into a dynamic vegetation model. Remote Sensing of Environment 112 (5): 2118-2130. doi: 10.1016/j.rse.2007.10.010

Koukal T, Suppan F, Schneider W (2007) The impact of relative radiometric calibration on the accuracy of kNN-predictions of forest attributes. Remote Sensing of Environment 110 (4): 431-437. doi: 10.1016/j.rse.2006.08.016

Krause B, Culmsee H, Wesche K, Bergmeier E, Leuschner C (2011) Habitat loss of floodplain meadows in north Germany since the 1950s. Biodiversity and Conservation 20 (11): 2347-2364. doi: 10.1007/s10531-011-9988-0

Kreibich H, Thieken A (2009) Coping with floods in the city of Dresden, Germany. Natural Hazards 51 (3): 423-436. doi: 10.1007/s11069-007-9200-8

Kurz WA, Apps MJ (2006) Developing Canada's National Forest Carbon Monitoring, Accounting and Reporting System to Meet the Reporting Requirements of the Kyoto Protocol. Mitigation and Adaptation Strategies for Global Change 11 (1): 33-43. doi: 10.1007/s11027-006-1006-6

Lair GJ, Zehetner F, Fiebig M, Gerzabek MH, van Gestel CAM, Hein T, Hohensinner S, Hsu P, Jones KC, Jordan G, Koelmans AA, Poot A, Slijkerman DME, Totsche KU, Bondar-Kunze E, Barth JAC (2009) How do long-term development and periodical changes of river-floodplain systems affect the fate of contaminants? Results from European rivers. Environmental Pollution 157 (12): 3336-3346. doi: 10.1016/j.envpol.2009.06.004

Lal R (2005) Forest soils and carbon sequestration. Forest Ecology and Management 220 (2005): 242-258. doi: 10.1016/j.foreco.2005.08.015

Laliberte AS, Rango A, Herrick JE, Fredrickson EL, 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. doi: 10.1016/j.jaridenv.2006.08.016

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33 (1): 159-174

Lek S, Guégan JF (1999) Artificial neural networks as a tool in ecological modelling, an introduction. Ecological Modelling 120 (2-3): 65-73. doi: 10.1016/S0304-3800(99)00092-7

Liao C, Luo Y, Fang C, Li B (2010) Ecosystem Carbon Stock Influenced by Plantation Practice: Implications for Planting Forests as a Measure of Climate Change Mitigation. PLoS ONE 5 (5): e10867

Lillesand M, Kiefer R, Chipman J (2004) Remote Sensing and Image Interpretation. 5th edn., New York

Loh W-Y (2011) Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 1 (1): 14-23. doi: 10.1002/widm.8

Maeda EE, Pellikka P, Clark BJF (2010) Monte Carlo simulation and remote sensing applied to agricultural survey sampling strategy in Taita Hills, Kenya. African Journal of Agricultural Research 5 (13): 1647-1654

Magnussen S, McRoberts RE, Tomppo EO (2009) Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. Remote Sensing of Environment 113 (3): 476-488. doi: 10.1016/j.rse.2008.04.018

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. doi: 10.1007/s11273-009-9162-6

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. doi: 10.1111/j.1365-2699.2004.01073.x

McBratney AB, Mendonça Santos ML, Minasny B (2003) On digital soil mapping. Geoderma 117 (1-2): 3-52. doi: 10.1016/S0016-7061(03)00223-4

McInerney DO, Nieuwenhuis M (2009) A comparative analysis of kNN and decision tree methods for the Irish National Forest Inventory. International Journal of Remote Sensing 30 (19): 4937-4955. doi: 10.1080/01431160903022936

Mertes LAK (2002) Remote sensing of riverine landscapes. Freshwater Biology 47 (4): 799-816. doi: 10.1046/j.1365-2427.2002.00909.x

Midwood J, Chow-Fraser P (2010) Mapping Floating and Emergent Aquatic Vegetation in Coastal Wetlands of Eastern Georgian Bay, Lake Huron, Canada. Wetlands 30 (6): 1141-1152. doi: 10.1007/s13157-010-0105-z

Milllennium Ecosystem Assessment (2005) Ecosystwems and Human Wellbeing. Island Press, Washington, DC
Mitchell JK (2003) European River Floods in a Changing World. Risk Analysis 23 (3): 567-574. doi: 10.1111/1539-6924.00337
Mitra S, Wassmann R, Vlek PLG (2005) An appraisal of global wetland area and its organic carbon stock. Anglais 88: 25-35
Mitsch W, Bernal B, Nahlik A, Mander Ü, Zhang L, Anderson C, Jørgensen SE, Brix H (2012) Wetlands, carbon, and climate change. Landscape Ecology 28 (4): 583-597. doi: 10.1007/s10980-012-9758-8
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. doi: 10.1007/s11273-009-9164-4
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. doi: 10.1080/014311600209742
Neeff T, de Alencastro Graça PM, Dutra LV, 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. doi: 10.1016/j.rse.2004.12.002
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. doi: 10.1186/1750-0680-4-11
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. doi: 10.5194/bg-5-817-2008
Omasa K, Qiu GY, Watanuki K, Yoshimi K, Akiyama Y (2003) Accurate Estimation of Forest Carbon Stocks by 3-D Remote Sensing of Individual Trees. Environmental Science & Technology 37 (6): 1198-1201. doi: 10.1021/es0259887
OpenCV-Wiki (2010) Decision Trees. http://opencv.willowgarage.com/documentation/cpp/ml_decision_trees.html. Accessed January 21 2013
Opperman JJ, Galloway GE, Fargione J, Mount JF, Richter BD, Secchi S (2009) Sustainable Floodplains Through Large-Scale Reconnection to Rivers. Science 326 (5959): 1487-1488. doi: 10.1126/science.1178256
Ozesmi SL, Bauer ME (2002) Satellite remote sensing of wetlands. Wetlands Ecology and Management 10 (5): 381-402. doi: 10.1023/a:1020908432489
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. doi: 10.1186/1750-0680-5-7
Parish F, Sirin A, Charman D, Joosten H, Minaeva T, Silvius M (2008) Assessment on peatlands, biodiversity and climate change. . Kuala Lumpur and Wetlands International Wageningen
Patenaude G, Hill RA, Milne R, Gaveau DLA, Briggs BBJ, Dawson TP (2004) Quantifying forest above ground carbon content using LiDAR remote sensing. Remote Sensing of Environment 93 (3): 368-380. doi: http://dx.doi.org/10.1016/j.rse.2004.07.016

Patenaude G, Milne R, Dawson TP (2005) Synthesis of remote sensing approaches for forest carbon estimation: reporting to the Kyoto Protocol. Environmental Science & Policy 8 (2): 161-178. doi: 10.1016/j.envsci.2004.12.010

Pedroli B, de Blust G, Looy K, Rooij S (2002) Setting targets in strategies for river restoration. Landscape Ecology 17 (1): 5-18. doi: 10.1023/a:1015221425315

Petrow T, Thieken AH, Kreibich H, Merz B, Bahlburg CH (2006) Improvements on Flood Alleviation in Germany: Lessons Learned from the Elbe Flood in August 2002. Environmental Management 38 (5): 717-732. doi: 10.1007/s00267-005-6291-4

Pontius RG, 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. doi: 10.1080/01431161.2011.552923

Quinlan JR (1986) Induction of decision trees. Machine Learning 1 (1): 81-106. doi: 10.1007/bf00116251

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. doi: 10.1016/j.ecolind.2012.03.017

Richter K, Atzberger C, Hank TB, Mauser W (2012) Derivation of biophysical variables from Earth observation data: validation and statistical measures. Journal of Applied Remote Sensing 6 (1): 063557-063551. doi: 10.1117/1.JRS.6.063557

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. doi: 10.1007/s11104-013-1638-8

Rocchini D, Foody GM, Nagendra H, Ricotta C, Anand M, He KS, 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. doi: 10.1016/j.cageo.2012.05.022

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. doi: 10.1007/s11273-011-9213-7

Rotach P (2004) Poplars and biodiversity. In. International Plant Genetic Resources Institute (IPGRI), Rome, pp 79-100

Rothenbücher J, Schaefer M (2005) Conservation of Leafhoppers in Floodplain Grasslands - Trade-off between Diversity and Naturalness in a Northern German National Park. Journal of Insect Conservation 9 (4): 335-349. doi: 10.1007/s10841-005-0514-0

Rouse JW, Haas RH, Schell JA, Deering DW Monitoring vegetation systems in the Great Plains with ERTS. In: Proceedings of the Third ERTS Symposium, Washington DC, 1973. pp 309-317

Saatchi SS, Houghton RA, Dos Santos AlvalÁ RC, Soares JV, Yu Y (2007) Distribution of aboveground live biomass in the Amazon basin. Global Change Biology 13 (4): 816-837. doi: 10.1111/j.1365-2486.2007.01323.x

Sali-Bazze M (1981) Entwicklung und Altersstellung von Böden der Donau- und Marchauen im Gebiet von Orth an der Donau und Marchegg / Baumgarten. Universität für Bodenkultur, Vienna

Samaritani E, Shrestha J, Fournier B, Frossard E, Gillet F, Guenat C, Niklaus PA, Pasquale N, Tockner K, Mitchell EAD, Luster J (2011) Heterogeneity of soil carbon pools and fluxes in a channelized and a restored floodplain section (Thur River, Switzerland). Hydrol Earth Syst Sci 15 (6): 1757-1769. doi: 10.5194/hess-15-1757-2011

Sandau R (2010) Status and trends of small satellite missions for Earth observation. Acta Astronautica 66 (1-2): 1-12. doi: 10.1016/j.actaastro.2009.06.008

Schindler S, Kropik M, Euller K, Bunting S, Schulz-Zunkel C, Hermann A, Hainz-Renetzeder C, Kanka R, Mauerhofer V, Gasso V, Krug A, Lauwaars S, Zulka K, Henle K, Hoffmann M, Biro M, Essl F, Jaquier S, Balazs L, Borics G, Hudin S, Damm C, Pusch M, van der Sluis T, Sebesvari Z, Wrbka T (2013) Floodplain management in temperate regions: is multifunctionality enhancing biodiversity? Environmental Evidence 2 (1): 10

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. doi: 10.1080/01431161.2012.666812

Shanmuganathan S, Sallis P, Buckeridge J (2006) Self-organising map methods in integrated modelling of environmental and economic systems. Environmental Modelling & Software 21 (9): 1247-1256. doi: 10.1016/j.envsoft.2005.04.011

Shiel R, Green J, Nielsen D (1998) Floodplain biodiversity: why are there so many species? Hydrobiologia 387-388 (0): 39-46. doi: 10.1023/a:1017056802001

Silva TSF, Costa MPF, Melack JM (2010) Spatial and temporal variability of macrophyte cover and productivity in the eastern Amazon floodplain: A remote sensing approach. Remote Sensing of Environment 114 (9): 1998-2010. doi: 10.1016/j.rse.2010.04.007

Stammel B, Cyffka B, Geist J, Müller M, Pander J, Blasch G, Fischer P, Gruppe A, Haas F, Kilg M, Lang P, Schopf R, Schwab A, Utschik H, Weißbrod M (2011) Floodplain restoration on the Upper Danube (Germany) by re-establishing water and sediment dynamics: a scientific monitoring as part of the implementation. River Systems 20 (1-2): 55-70

Stockmann U, Adams MA, Crawford JW, Field DJ, Henakaarchchi N, Jenkins M, Minasny B, McBratney AB, Courcelles VDRD, Singh K, Wheeler I, Abbott L, Angers DA, Baldock J, Bird M, Brookes PC, Chenu C, Jastrow JD, Lal R, Lehmann J, O'Donnell AG, Parton WJ, Whitehead D, Zimmermann M (2013) The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agriculture, Ecosystems and Environment 164: 80-99

Stoms DM, Estes JE (1993) A remote sensing research agenda for mapping and monitoring biodiversity. International Journal of Remote Sensing 14 (10): 1839-1860. doi: 10.1080/01431169308954007

Strassburg BBN, Kelly A, Balmford A, Davies RG, Gibbs HK, Lovett A, Miles L, Orme CDL, Price J, Turner RK, Rodrigues ASL (2010) Global congruence of carbon storage and biodiversity in terrestrial ecosystems. Conservation Letters 3 (2): 98-105. doi: 10.1111/j.1755-263X.2009.00092.x

Stümer W, Kenter B, Köhl M (2010) Spatial interpolation of in situ data by self-organizing map algorithms (neural networks) for the assessment of carbon stocks in European forests. Forest Ecology and Management 260 (3): 287-293. doi: 10.1016/j.foreco.2010.04.008

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. doi: 10.1007/s11273-012-9252-8

Suchenwirth L, Forster M, Lang F, Kleinschmit B (2013) Estimation and Mapping of Carbon Stocks in Riparian Forests by using a Machine Learning Approach with Multiple

Geodata. PFG Photogrammetrie, Fernerkundung, Geoinformation 2013 (4): 333-349. doi: 10.1127/1432-8364/2013/0181

Thevs N, Zerbe S, Schnittler M, Abdusalih N, Succow M (2008) Structure, reproduction and flood-induced dynamics of riparian Tugai forests at the Tarim River in Xinjiang, NW China. Forestry 81 (1): 45-57. doi: 10.1093/forestry/cpm043

Tockner K, Schiemer F, Ward JV (1998) Conservation by restoration: the management concept for a river-floodplain system on the Danube River in Austria. Aquatic Conservation: Marine and Freshwater Ecosystems 8 (1): 71-86. doi: 10.1002/(sici)1099-0755(199801/02)8:1<71::aid-aqc265>3.0.co;2-d

Tockner K, Stanford JA (2002) Riverine flood plains: present state and future trends. Environmental Conservation 29 (03): 308-330. doi: 10.1017/S037689290200022X

Tomppo E (1991) Satellite image-based national forest inventory of Finland. International Archive of Photogrammetry and Remote Sensing 28: 419-424

Tomppo E, Goulding C, Katila M (1999) Adapting Finnish Multi-Source Forest Inventory Techniques to the New Zealand Preharvest Inventory. Scandinavian Journal of Forest Research 14 (2): 182-192. doi: 10.1080/02827589950152917

Tomppo E, Halme M (2004) Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. Remote Sensing of Environment 92 (1): 1-20. doi: 10.1016/j.rse.2004.04.003

Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8 (2): 127-150. doi: 10.1016/0034-4257(79)90013-0

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. doi: 10.1016/j.rse.2004.10.001

Turner DP, Ollinger SV, Kimball JS (2004) Integrating Remote Sensing and Ecosystem Process Models for Landscape- to Regional-Scale Analysis of the Carbon Cycle. BioScience 54 (6): 573-584. doi: 10.1641/0006-3568(2004)054[0573:irsaep]2.0.co;2

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

Vogt K, Rasran L, Jensen K (2006) Seed deposition in drift lines during an extreme flooding event - Evidence for hydrochorous dispersal? Basic and Applied Ecology 7 (5): 422-432. doi: 10.1016/j.baae.2006.05.007

Wagner-Lücker I, Lanz E, Förster M, Janauer GA, Reiter K (2013) Knowledge-based framework for delineation and classification of ephemeral plant communities in riverine landscapes to support EC Habitat Directive assessment. Ecological Informatics 14 (0): 44-47. doi: 10.1016/j.ecoinf.2012.11.003

Wagner I (2009) The Danube Floodplain Habitats - application of the Object-based Image Analysis approach. In: Car/Griesebner/Strobl (ed) Geospatial Crossroads @ GI_Forum '09. Proceedings of the Geoinformatics Forum Salzburg. Wichmann, Heidelberg, pp 218-227

Wang ZS, Bian SR, Liu Y, Liu ZH (2013) The load characteristics classification and synthesis of substations in large area power grid. International Journal of Electrical Power & Energy Systems 48: 71-82. doi: 10.1016/j.ijepes.2012.11.032

Ward JV, Tockner K, Arscott DB, Claret C (2002) Riverine landscape diversity. Freshwater Biology 47 (4): 517-539. doi: 10.1046/j.1365-2427.2002.00893.x

Ward JV, Tockner K, Schiemer F (1999) Biodiversity of floodplain river ecosystems: ecotones and connectivity. Regulated Rivers: Research & Management 15 (1-3): 125-139. doi: 10.1002/(sici)1099-1646(199901/06)15:1/3<125::aid-rrr523>3.0.co;2-e

Wendroth O, Reuter HI, Kersebaum KC (2003) Predicting yield of barley across a landscape: a state-space modeling approach. Journal of Hydrology 272 (1â€“4): 250-263. doi: http://dx.doi.org/10.1016/S0022-1694(02)00269-X

Wenger EL, Zinke A, Gutzweiler K-A (1990) Present situation of the European floodplain forests. Forest Ecology and Management 33â€“34 (0): 5-12. doi: 10.1016/0378-1127(90)90180-J

Wharton TN, Kriticos DJ (2004) The fundamental and realized niche of the Monterey Pine aphid, Essigella californica (Essig) (Hemiptera: Aphididae): implications for managing softwood plantations in Australia. Diversity and Distributions 10 (4): 253-262. doi: 10.1111/j.1366-9516.2004.00090.x

Whittaker RH (1972) Evolution and Mesaurement of Species Diversity. Taxon 21 (2/3): 213-251

Williams C, Hanan N, Neff J, Scholes R, Berry J, Denning AS, Baker D (2007) Africa and the global carbon cycle. Carbon Balance and Management 2 (1): 3. doi: 10.1186/1750-0680-2-3

Winiwarter V, Schmid M, Dressel G (2013) Looking at half a millennium of co-existence: the Danube in Vienna as a socio-natural site. Water History 5 (2): 101-119. doi: 10.1007/s12685-013-0079-x

Xuan SY, Wu YB, Chen XF, Liu J, Yan AX (2013) Prediction of bioactivity of HIV-1 integrase ST inhibitors by multilinear regression analysis and support vector machine. Bioorg Med Chem Lett 23 (6): 1648-1655. doi: 10.1016/j.bmcl.2013.01.081

Yang H, Li J (2013) Predictions of soil organic carbon using laboratory-based hyperspectral data in the northern Tianshan mountains, China. Environmental Monitoring and Assessment 185 (5): 3897-3908. doi: 10.1007/s10661-012-2838-z

Zadeh LA (1989) Knowledge Representation in Fuzzy Logic. IEEE Trans on Knowl and Data Eng 1 (1): 89-100. doi: http://dx.doi.org/10.1109/69.43406

ZAMG ZfMuG (2002) Klimadaten von Österreich 1971-2000. Zentralanstalt für Meterorologie und Geodynamik, Vienna, Austria

Zedler JB (2003) Wetlands at your service: reducing impacts of agriculture at the watershed scale. Frontiers in Ecology and the Environment 1 (2): 65-72. doi: 10.1890/1540-9295(2003)001[0065:waysri]2.0.co;2

Zehetner F, Lair GJ, Gerzabek MH (2009) Rapid carbon accretion and organic matter pool stabilization in riverine floodplain soils. Global Biogeochem Cycles 23 (4): GB4004. doi: 10.1029/2009gb003481