Flächennutzungsmonitoring V
Methodik – Analyseergebnisse – Flächenmanagement

IÖR Schriften Band 61 · 2013

ISBN: 978-3-944101-18-7

Vom Satellitenbild zu Stadtstrukturtypen: Wie Graphen die Flächennutzung charakterisieren

Irene Walde, Sören Hese, Christian Berger, Christiane Schmullius

Walde, I.; Hese, S.; Berger, C.; Schmullius, C. (2013): Vom Satellitenbild zu Stadtstrukturtypen: Wie Graphen die Flächennutzung charakterisieren. In: Meinel, G.; Schumacher, U.; Behnisch, M. (Hrsg.): Flächennutzungsmonitoring V. Methodik – Analyseergebnisse – Flächenmanagement. Berlin: Rhombos, IÖR Schriften 61, S. 179-187.

RHOMBOS Rhombos Verlag – www.rhombos.de

Leibniz-Institut für ökologische Raumentwicklung IÖR – www.ioer.de

Vom Satellitenbild zu Stadtstrukturtypen: Wie Graphen die Flächennutzung charakterisieren

Irene Walde, Sören Hese, Christian Berger, Christiane Schmullius

Zusammenfassung

Dieser Beitrag beschreibt eine Methode zur Klassifikation urbaner Landnutzung aus sehr hoch aufgelösten Satellitenbildern. Landbedeckungsobjekte, abgeleitet aus einer QuickBird-Szene und einem normalisierten Oberflächenmodell der Hansestadt Rostock, dienen der Untersuchung von topologischen Eigenschaften, die durch Nachbarschaftsgraphen modelliert werden. Es werden Graphen-Indikatoren zur Ableitung und Trennung der urbanen Strukturtypen verwendet. Der Einsatz eines Random Forest Klassifikators, basierend auf maschinellem Lernen, ermöglicht die Untersuchung der Bedeutung der Graphen-Indikatoren für das Klassifikationsergebnis. Der Einfluss verschiedener Einstellungsparameter, die Auswirkungen durch Reduktion der Trainingsdaten und Graphen-Indikatoren wird bezüglich der Klassifikationsgüte analysiert. Anschließend werden für einen unabhängigen Testdatensatz die Stadtstrukturtypen zugeordnet und das Ergebnis validiert. Eine Gesamtgenauigkeit von 87 % wurde erreicht. Dabei ist die Ableitung von städtischen Kleingartenanlagen zu 95 % genau. Durch eine hohe Variabilität innerhalb der urbanen Landnutzungsklasse Industrie- und Gewerbegebiet, konnte lediglich eine Genauigkeit von 66 % erreicht werden. Den größten Einfluss auf das Klassifikationsergebnis hat die Höhe des Gebäudes mit dem höchsten Knotengrad innerhalb eines urbanen Blocks.

1 Einführung

Vielfältige Landbedeckungsklassen (z. B. Gebäude, Straßen, Grünflächen) charakterisieren den urbanen Raum (Herold et al. 2002). Anhand ihrer Komposition sind urbane Strukturen mit homogenen räumlichen, topologischen oder morphologischen Eigenschaften erkennbar (Banzhaf, Höfer 2008). Urbane Strukturtypen haben unterschiedliche Anforderungen an die Stadtplanung, z. B. Müllentsorgung, Verkehrsaufkommen, Wasser- und Energieversorgung (Pauleit, Duhme 2000). Die fortwährende Urbanisierung erfordert eine nachhaltige Planung und Entwicklung urbaner Räume auf nationalem (Bock et al. 2001), europäischem (European Commission and Directorate – General for Regional Policy 2009) und globalem Level (UN Habitat 2008). Fernerkundung kann dabei als Basis-Layer das Kartieren, Überwachen und Modellieren urbaner Bereiche unterstützen (Blaschke et al. 2011). Im Rahmen dieses Beitrages wird ein Graphen-basierter Ansatz zur Extraktion urbaner Strukturtypen aus Landbedeckungsobjekten, die aus einem sehr hoch aufgelösten Satellitenbild und einem LiDAR-basiertem normalisiertem Objekthöhenmodell abgeleitet wurden, vorgestellt.

2 Datenbeschreibung

Vom Untersuchungsgebiet Rostock liegen eine wolkenfreie, sehr hoch aufgelöste Quickbird-Satellitenbildszene (September 2009, 0,6 m pangeschärfte Bodenpixelauflösung) und ein normalisiertes LiDAR-Objekthöhenmodell (Oktober 2006, 2 m Bodenpixelauflösung) vor. Eine objektbasierte Segmentierung und Klassifikation der Rohdaten ergab die folgenden sechs Landbedeckungsklassen und deren Nutzer-/Produzentengenauigkeit (Congalton 1991) in Prozent: Gebäude (90/90), Boden (96/86), Grasland (84/89), versiegelte Fläche (82/89), Baum (96/98) und Wasser (100/96). Eine detaillierte Beschreibung der Landbedeckungsklassifikation ist in Berger et al. (2013) nachzulesen. Zur Abgrenzung der urbanen Strukturen wurden die Baublockgrenzen aus dem ATKIS Basis-DLM verwendet. Ein Referenzdatensatz der urbanen Strukturen wurde unter Verwendung der DLM Nutzungsinformation semi-automatisch erzeugt. Dabei konnten die Industriegebiete und Gartenanlagen direkt übernommen werden, während die Zuweisung von Einfamilienhaus Wohngebiet, Blockbebauung und Stadtzentrum manuell erfolgte.

3 Methodik

Die angewandte Methode ist in Abbildung 1 schematisch dargestellt. Für die Landbedeckungspolygone werden die Zentroide (Knoten des Graphen) und deren Nachbarschaft (Kanten des Graphen) berechnet. Die DLM-Baublöcke begrenzen die Graphen. Mithilfe eines Random Forest Modells werden Graphen-Indikatoren ausgewertet und die Stadtstrukturtypen abgeleitet. Eine Validierung der Ergebnisse gibt Aufschluss über die erreichte Genauigkeit.

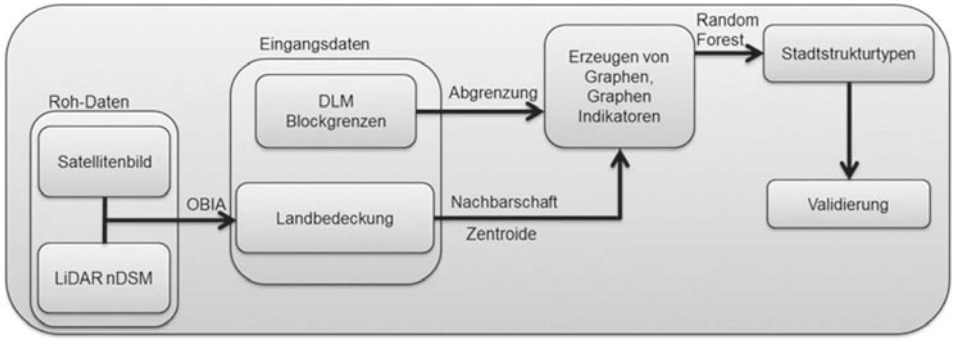

Abb. 1: Arbeitsschritte (Quelle: eigene Bearbeitung)

3.1 Urbane Strukturtypen

Die fünf definierten Stadtstrukturtypen sind in Abbildung 2 dargestellt. Das Stadtzentrum ist gekennzeichnet durch unmittelbar angrenzende mehrstöckige Häuser

in direkter Straßennachbarschaft mit versiegelten oder begrünten Innenhöfen. Die Blockbebauung wird durch drei- und mehrstöckige, rechteckige Gebäude in meist regelmäßiger Anordnung charakterisiert. Freistehende ein- bis zweistöckige Gebäude mit gleichmäßiger bis zu individueller Form und einem hohen Grünflächenanteil beschreiben ein Einfamilienhaus-Wohngebiet. In Gartenanlagen stehen Gartenhäuser mit geringer Höhe und Fläche in meist kompakter Form und sind umgeben von Grünflächen. Industriegebiete bestehen aus großflächigen einfachen Gebäuden mit benachbarten versiegelten Flächen (Parkplätze) und wenig Grünflächenanteilen. Eine detaillierte Beschreibung ist in Walde et al. (2013) nachzulesen.

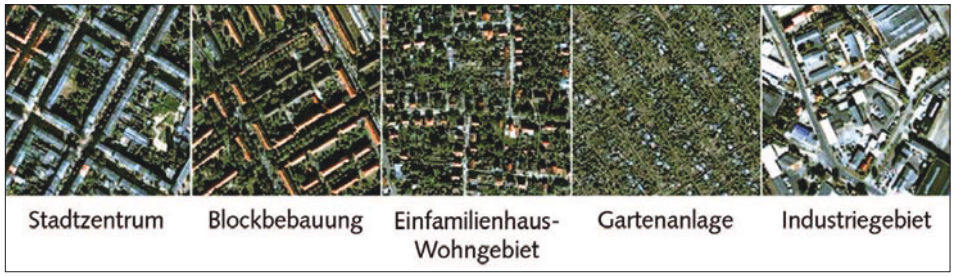

Abb. 2: Quickbird-Ausschnitte der urbanen Strukturtypen (Quelle: ©DigitalGlobe, Inc., 2011)

3.2 Polygon Nachbarschaft und Graphen

Die direkten Nachbarschaften aller Landbedeckungspolygone wurden pro Baublock berechnet. Dabei sind auch Polygone mit nur einem gemeinsamen Eckpunkt benachbart. Besteht eine Nachbarschaft werden die Zentroide der Polygone (Knoten des Graphen) mit einer Kante verbunden und ein Graph entsteht. In Abbildung 3 sind zwei Graphen-Beispiele jeweils mit und ohne geographische Referenz abgebildet.

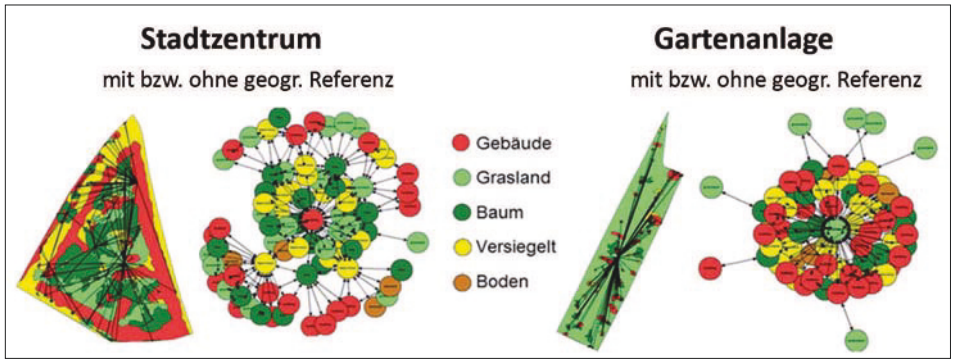

Abb. 3: Graphen aus Landbedeckungsklassen mit und ohne geographische Referenz (Quelle: eigene Bearbeitung)

3.3 Graphen-Indikatoren

Insgesamt wurden 22 Graphen-Indikatoren, gruppiert nach vier Kategorien, berechnet. Drei Zentralitätsmaße wurden verwendet, um die dominante Landbedeckungsklasse im Graph zu bestimmen. Dafür wurde zum einen für jeden Graphen die Landbedeckungsklasse mit dem höchsten Knotengrad (Knoten mit der höchsten Kantenanzahl) ermittelt (Borgatti, Halgin 2011). Zum anderen wurde die Landbedeckung mit der höchsten „Betweenness"-Zentralität, die durch die Anzahl der kürzesten Wege eines Graphen durch einen bestimmten Knoten definiert ist, festgestellt (Borgatti, Halgin 2011). Als drittes Zentralitätsmaß wurde der mittlere Knotengrad von Gebäuden berechnet. Die zweite Kategorie sind die Adjazenz-Event-Maße, die 13 Graphen-Indikatoren enthält. Sie berücksichtigen die Häufigkeit und räumliche Anordnung der Landbedeckungsklassen (Barnsley, Barr 1996). Eine Adjazenz-Event-Matrix speichert dabei die Anzahl der Kanten aller Landbedeckungskombinationen. Ausgehend von dieser Matrix wurden die folgenden Indikatoren je Graph abgeleitet: normalisierte Kantenanzahl zu versiegelter Fläche (respektive Baum und Gras), normalisierte Kantenanzahl von Gebäuden zu versiegelter Fläche (respektive Baum und Gras), Landbedeckungsdiversität um Gebäude, Prozent der Gebäude mit mindestens einer Kante zu versiegelter Fläche (respektive Baum und Gras), Prozent der Gebäude die von versiegelter Fläche (respektive Baum und Gras) umschlossen sind. Vier zusätzliche Graphenmaße, bezogen auf das Gebäude eines Graphen mit dem höchsten Knotengrad, wurden ermittelt. Dazu gehören die folgenden Gebäudeeigenschaften: Gebäudehöhe, Gebäudefläche, Gebäudevolumen und Gebäudeform (Shape Index). Abschließend wurden zwei Konnektivitätsmaße berechnet, die nicht auf Basis der Nachbarschaftsgraphen, sondern auf Basis von Distanzgraphen der Gebäude erzeugt wurden. Dabei handelt es sich um den Beta-Index (Rodrigue et al. 2009), der das Verhältnis von Kanten zu Knoten in einem Graph beschreibt. Die detaillierte Vorgehensweise zur Erzeugung der Graphen und Berechnung der Konnektivitätsmaße ist in Walde et al. (2013) nachzulesen.

3.4 Random Forest

Durch die vielfältigen Graphen-Indikatoren wurde ein maschinell lernender Klassifikator angewandt, um die urbanen Strukturen zu bestimmen. Die nicht-parametrische Methode Random Forest (Breiman 2001) wurde bereits erfolgreich für Fernerkundungsfragestellungen eingesetzt (Hüttich et al. 2011; Loosvelt et al. 2012). Es wird ein Ensemble von Entscheidungsbäumen generiert, wobei jeweils ein unterschiedlicher Auszug aus den Trainingsdaten genutzt wird, um die einzelnen Bäume zu erzeugen. Ein Drittel der Trainingsdaten für jeden Baum wird für die Validierung des Baumes verwendet. Die Klassenzuweisung basiert auf der Mehrheit der Stimmenvorhersagen der Bäume. Da die berechneten Variablen (Graphen-Indikatoren) aus kategorischen und numerischen Variablen bestehen und einige Variablen miteinander korrelieren, wurde der von Hothorn et al. (2006) entwickelte Ansatz cforest (party package) innerhalb der freien Software R verwendet.

4 Ergebnisse

Die Methode wurde auf 1 588 Baublöcke in Rostock angewandt, die halb-automatisch in die fünf definierten urbanen Strukturtypen vorklassifiziert wurden. Basierend auf der Klasse mit dem geringstem Vorkommen (Stadtzentrum: 123 Baublöcke) wurden 60 Trainingsgebiete pro Stadtstrukturtyp zufällig gewählt. Mit einer Gesamtzahl von 300 Trainingsgebieten wurde das Random Forest Model erzeugt, während die übrigen 1 288 Baublöcke als unabhängige Testdatensätze verwendet wurden. Eine Anzahl von 500 Bäumen erwies sich als stabil bezüglich der Klassifikationsgenauigkeit der Testdatensätze. Im nächsten Schritt wurde die Variablen-Wichtigkeit untersucht, die in Abbildung 4 als Ranking visualisiert ist. Die Klassifikationsgenauigkeit wurde maßgeblich durch die Höhe des Gebäudes mit dem höchsten Knotengrad bestimmt, gefolgt von der normalisierten Kantenanzahl zu Grasflächen. Die Graphen-Indikatoren, die die Prozentzahl an Gebäuden wiedergeben, die von einer anderen Landbedeckungsklasse umschlossen sind, hatten den geringsten Einfluss auf die Klassifikationsgüte.

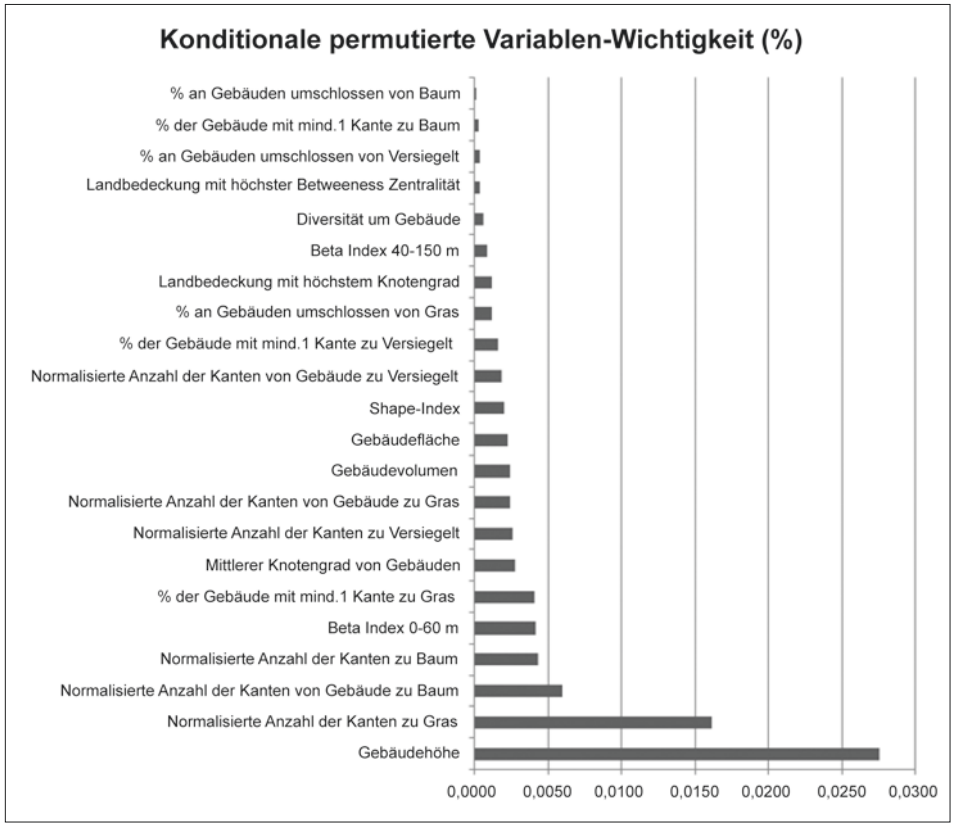

Abb. 4: Konditionale permutierte Variablen-Wichtigkeit (Quelle: eigene Bearbeitung)

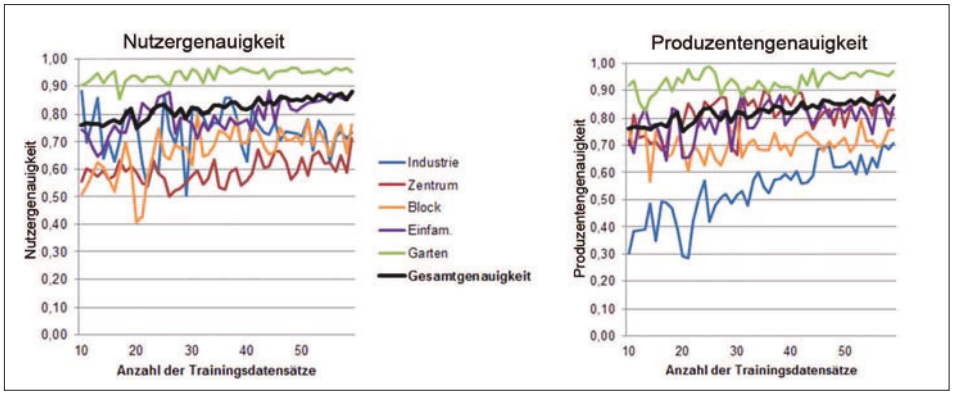

Abb. 5: Auswirkung der Reduktion der Trainingsgebiete auf Nutzer- und Produzentengenauigkeit (Quelle: eigene Bearbeitung)

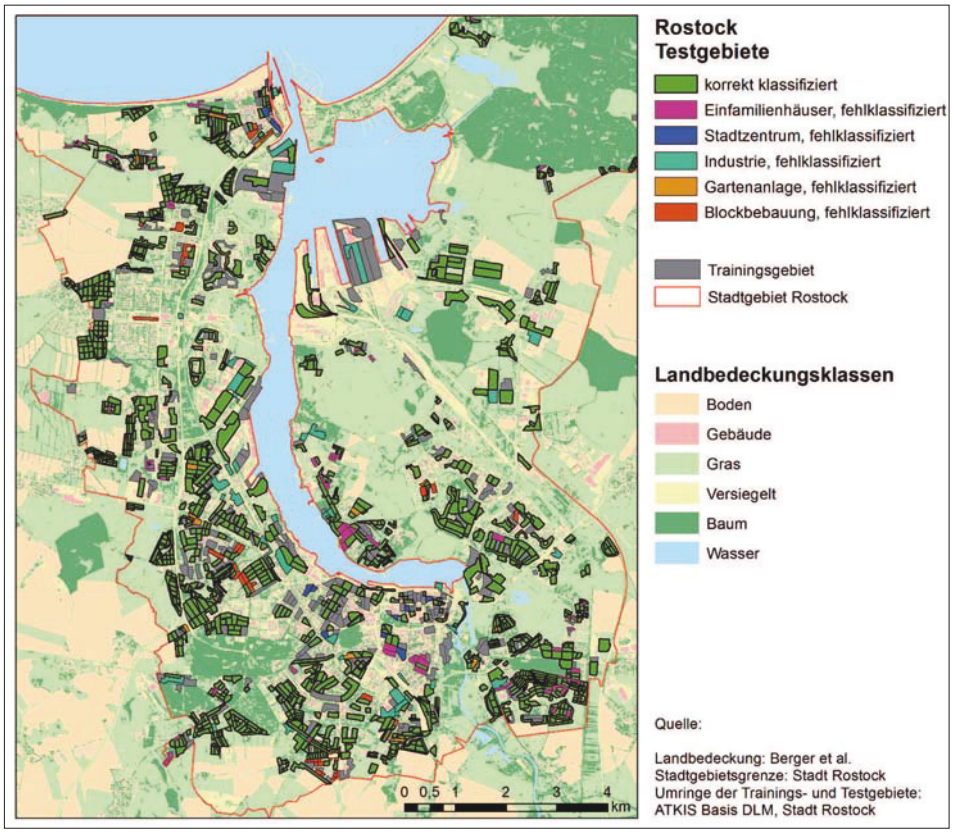

Abb. 6: Vergleich korrekt und falsch klassifizierter Testdaten

Anschließend wurde der Einfluss der Reduktion der Trainingsdatensätze auf die Nutzer- und Produzentengenauigkeit der Testdatensätze untersucht (Abb. 5). Dafür wurden die Trainingsdaten schrittweise um eins reduziert bis eine Anzahl von zehn Trainingsgebieten pro Stadtstrukturtyp erreicht wurde. Wie erwartet, nimmt die Klassifikationsgenauigkeit mit abnehmender Anzahl an Trainingsgebieten ab. Eine zufriedenstellende und stabile Gesamtgenauigkeit von 80 % wurde ab ca. 35 Trainingsdatensätzen pro Stadtstrukturtyp erreicht. Abbildung 5 zeigt nicht nur die Robustheit bezüglich der Trainingsdatenreduktion, sondern auch die Intra-Klassen Variabilität. Je größer die Amplitude, umso größer ist die Intra-Klassen Variabilität.

Erkennbar ist die relativ stabile Kurve der Gartenanlagen im Vergleich zu den anderen Klassen. Das bedeutet, dass Gartenanlagen in Rostock sehr homogen ausgeprägt sind, während z. B. Block- oder Industriegebiete sehr heterogen in ihrer Landbedeckungszusammensetzung sind. Für die finale Klassifikation wurden die 17 besten Graphen-Indikatoren und 60 Trainingsdatensätze pro Klasse verwendet. Eine Gesamtgenauigkeit von 87 % wurde erreicht, wobei 1 118 Baublöcke korrekt und 170 falsch klassifiziert wurden (Abb. 6). Folgende Nutzer- und Produzentengenauigkeiten konnten erreicht werden: Industriegebiete (66 %/71 %), Stadtzentrum (81 %/64 %), Blockbebauung (73 %/73 %), Einfamilienhaus Wohngebiete (81 %/87 %) und Gartenanlagen (97 %/95 %).

5 Fazit

Aus sechs Landbedeckungsklassen wurden Nachbarschaftsgraphen generiert, die durch Baublockgrenzen voneinander abgetrennt wurden. Fünf Stadtstrukturtypen wurden mithilfe von 17 Graphen-Indikatoren und einer überwachten Random Forest Klassifikation abgeleitet. Eine Gesamtgenauigkeit von 87 % konnte erreicht werden. Die Höhe des Gebäudes mit dem höchsten Knotengrad stellte sich als einflussreichste Variable heraus, neben der Gruppe der Adjazenz-Event-Maße. Eine hohe Intra-Klassen-Variabilität wurde bei Industrie und Blockbebauung festgestellt, während Gartenanlagen sehr homogen ausgeprägt sind. Eine Verbesserung des Ansatzes kann durch eine Reduzierung der Intra-Klassen-Variabilität erreicht werden, indem eine Einteilung in entsprechende Unterklassen vorgenommen wird. Auch für den praktikablen Einsatz müssen die vorhandenen Strukturtypen in homogenere unterteilt werden, um eine ausreichende Genauigkeit bei gleichzeitiger Reduktion der Trainingsdaten beizubehalten. Durch die detailliertere Anpassung der Klassen an das Untersuchungsgebiet nimmt die Robustheit des Ansatzes bezüglich der Übertragbarkeit auf andere Städte ab. Ein Übertragbarkeitstest auf andere Städte mit ähnlichen Stadtstrukturtypen ist geplant und setzt eine Landbedeckungsklassifikation mit vergleichbarer Genauigkeit voraus. Der Vergleich der Variablen-Wichtigkeit mehrerer Städte gibt Aufschluss über die Relevanz der Graphen-Indikatoren. Eine Untersuchung, ob das für Rostock erzeugte Random

Forest Model direkt für ein anderes Untersuchungsgebiet verwendet werden kann (unüberwacht) oder ob neue Trainingsdaten aus dem jeweiligen Untersuchungsgebiet für eine Neuberechnung des Random Forest Models benötigt werden (überwacht), ist vorgesehen. Darüber hinaus soll eine Methode zur automatischen Partitionierung der Graphen in die urbanen Strukturtypen entwickelt werden, um unabhängig von Baublockgrenzen eine Ableitung der Stadtstrukturtypen zu gewährleisten.

6 Literatur

Banzhaf, E.; Höfer, R. (2008): Monitoring Urban Structure Types as Spatial Indicators With CIR Aerial Photographs for a More Effective Urban Environmental Management. IEEE Journal of selected topics in Applied Earth Observations & Remote Sensing, 1(2)/2008, 129-138.

Barnsley, M. J.; Barr, S. L. (1996): Inferring Urban Land Use from Satellite Sensor Images Using Kernel-Based Spatial Reclassification. Photogrammetric Engineering & Remote Sensing, 62(8)/1996, 949-958.

Berger, C. et al. (2013): Robust Extraction of Urban Land Cover Information from HSR Multi-Spectral and LiDAR Data. IEEE Journal of selected topics in Applied Earth Observations & Remote Sensing, 6(6)/2013, 1-16.

Blaschke, T. et al. (2011): Collective Sensing: Integrating Geospatial Technologies to Understand Urban Systems – An Overview. Remote Sensing, 3(8)/2011, 1743-1776.

Bock, S.; Hinzen, A.; Libbe, J. (Hrsg.) (2011): Nachhaltiges Flächenmanagement – Ein Handbuch für die Praxis: Ergebnisse aus der REFINA-Forschung. Berlin: Spree Druck Berlin GmbH.

Borgatti, S. P.; Halgin, D. S. (2011), (2011c): Analyzing Affiliation Networks. In: Scott, J.; Carrington, P. J. (Hrsg.): The SAGE handbook of social network analysis. SAGE.

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

Congalton, R. G. (1991): A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment, 37(1)/1991, 35-46.

European Commission and Directorate-General for Regional Policy (Hrsg.) (2009): Promoting sustainable urban development in Europe: Achievements and Opportunities. Belgium: European Communities.

Herold, M.; Scepan, J.; Clarke, K. C. (2002): The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. Environment and Planning, 34/2002, 1443-1458.

Hothorn, T.; Hornik, K.; Zeileis, A. (2006): Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics, 15(3)/2006, 651-674. (R-Paket: party package).

Hüttich, C. et al. (2011): Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. Remote Sensing of Environment, 115(10)/2011, 2445-2459.

Loosvelt, L. et al. (2012): Impact of Reducing Polarimetric SAR Input on the Uncertainty of Crop Classifications Based on the Random Forests Algorithm. IEEE Transactions on Geoscience and Remote Sensing, 50(10)/2010, 4185-4200.

Pauleit, S.; Duhme, F. (2000): Assessing the environmental performance of land cover types for urban planning. Landscape and Urban Planning, 52(1)/2000, 1-20.

Rodrigue, J.-P. (2009): The geography of transport systems. www.people.hofstra.edu/geotrans (Zugriff: 13.06.2013).

UN Habitat (Hrsg.) (2008): State of the Worlds Cities 2010/2011: Bridging the urban divide. London: Earthscan.

Walde, I. et al. (2013): Graph based Mapping of Urban Structure Types from High Resolution Satellite Image Objects – A case study of the German cities Rostock and Erfurt. IEEE Geoscience and Remote Sensing Letters, 10(4)/2013, 932-936.