

Erstellung von Fichten- und Kiefernanteilskarten auf Basis von Satellitendaten für Bayern

MARKUS IMMITZER^{1,*}, KATHRIN EINZMANN¹, SEBASTIAN BÖCK¹, MATTEO MATTIUZZI¹, WAI-TIM NG¹, ADELHEID WALLNER², NICOLE PINNEL³, ANNE REICHMUTH³, MATTHIAS FROST⁴, ANDREAS MÜLLER³, RUDOLF SEITZ² UND CLEMENT ATZBERGER¹

Zusammenfassung

Der fortschreitende Klimawandel erhöht das Gefährdungspotenzial von Wäldern zunehmend. Eine forstwirtschaftliche Nutzung von Wäldern ist oftmals nur durch den Wechsel auf Baumarten mit größerer Flexibilität gegenüber den Auswirkungen des Klimawandels langfristig sichergestellt. Für derartige Waldumbauvorhaben auf großer Fläche benötigen die Akteure aktuelle Verbreitungskarten einzelner Baumarten in einer räumlichen Auflösung von z. B. 1 Hektar. Um eine regelmäßige Überwachung und Aktualisierung zu ermöglichen, sind die Kosten für derartige Kartenprodukte ein weiterer wichtiger Faktor, der zu berücksichtigen ist. Der vorliegende Beitrag beschreibt und validiert einen fernerkundungsbasierten Ansatz für die Kartierung von einzelnen Baumarten (Fichte und Kiefer) mit einer innovativen Kombination von kommerziellen, sehr hoch aufgelösten Satellitendaten und frei verfügbar Landsat-Zeitreihen. Die Methodik beinhaltet dabei die überwachte Klassifikation von WorldView-2 Daten ausgewählter Testgebiete, gefolgt von einem „Upscaling“ dieser Referenzinformationen auf große Flächen mit multispektralen und multitemporalen Landsat-Daten. Für die Modellierung wurde mit Random Forest (RF) ein auf Entscheidungsbäumen basierender Ansatz gewählt. Mit der entwickelten Methode konnten für ganz Bayern konsistente und genaue Karten der Fichten- und Kiefernverbreitung (kontinuierliche Anteile) mit einer Auflösung von 1 ha erstellt werden. Eine Validierung mit 3798 unabhängigen Referenzzellen ergab für Fichte bzw. Kiefer einen Root-Mean-Square Error (RMSE) von 11 und 14 %, und ein Bestimmtheitsmaß (R^2) von 0.74 bis 0.79. Zwischen 76 und 85 % der Validierungspunkte wurden besser modelliert als die angenommene Unsicherheitsmarge von ± 15 % der Referenzinformation (aus manueller Bildinterpretation von Orthophotos).

Schlüsselwörter

Upscaling; Random Forest Regression; WorldView-2, Landsat; prozentuelle Bedeckung; Baumartenkartierung

Abstract

Intensifying weather events and climate change have the potential to put healthy forests at risk. This calls for an adaption in forest management. To manage large scale forest conversion, stakeholders need up-to-date, accurate and consistent forest distribution maps at hecta-metric spatial resolution covering large areas. To enable such map products, costs are a major factor to consider. This paper describes and validates a remote sensing based approach for mapping individual tree species (here spruce and pine) using an innovative combination of commercial very high resolution (WorldView-2) images and freely available Landsat time series. The mapping involves a supervised classification of very high resolution images over a relatively small selection of test sites. Where after wall-to-wall upscaling of the derived reference information is used on

¹ Universität für Bodenkultur (BOKU) Wien, Institut für Vermessung, Fernerkundung und Landinformation (IVFL), Peter-Jordan-Straße 82, A-1190 Wien, E-Mail: vorname.nachname@boku.ac.at

² Bayerische Landesanstalt für Wald und Forstwirtschaft (LWF), Abteilung Informationstechnologie, Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising, E-Mail: vorname.nachname@lwf.bayern.de

³ Deutsches Zentrum für Luft- und Raumfahrt (DLR): Deutsches Fernerkundungsdatenzentrum (DFD) – Abteilung Landoberfläche, Oberpfaffenhofen, D-82234 Wessling, E-Mail: vorname.nachname@dlr.de

⁴ Bayerische Staatsforsten AöR (BaySF), Bereich Informations- und Kommunikationstechnik, Sonnenstraße 24-26, D-80331 München, E-Mail: vorname.nachname@baysf.de

* Korrespondierender Autor; E-Mail: markus.immitzer@boku.ac.at; Tel.: +43-1-47654-5132; Fax: +43-1-47654-5142.

multi-spectral and multi-temporal Landsat data. Random Forests (RF) regression trees are used for modeling. The method is tested on Bavaria (Germany), its approach yields consistent and accurate maps of fractional tree cover within 1-ha cells. The validation using 3798 independent reference samples provided a root-mean-square error (RMSE) between 11 and 14 % with almost no bias, and R^2 of 0.74-0.79 for spruce and pine, respectively. Between 76 and 85 % of the samples were modeled within the assumed uncertainty of ± 15 % of the reference sample, the latter being derived through manual photo-interpretation of orthoimagery.

Keywords

Upscaling; Random Forest regression; WorldView-2; Landsat; fractional coverage; tree species mapping

1 Einleitung

Durch den fortschreitenden Klimawandel werden sich die Wachstumsbedingungen für die verschiedenen Baumarten weiter verändern. Für Bayern wird dabei prognostiziert, dass Fichte und Kiefer zu den am stärksten betroffenen Baumarten gehören werden (KÖLLING et al. 2009). Um dem Klimawandel zu begegnen fehlen bis jetzt jedoch verlässliche, flächendeckende Karten über die Verbreitung der einzelnen Baumarten in mittlerer Auflösung (Flächengröße zirka 1 ha). Solche Grundlagendaten würden sowohl der forstlichen Beratung, als auch langfristigen Monitoringvorhaben dienen.

Im Gegensatz zu terrestrisch punktweise aufgenommenen Waldinventuren bietet die Auswertung von Fernerkundungsdaten die Möglichkeit flächendeckende Produkte zu liefern. Für eine bayernweite Umsetzung müssen neben der Berücksichtigung der Datenkosten auch der Aufwand für Interpretation und Analyse der Bilddaten beachtet werden. Eine überregionale Auswertung durch visuelle Stereointerpretation von Farbinfrarot-Luftbildern ist beispielsweise auf Grund des hohen Arbeitsaufwandes nicht praktikabel. Als brauchbare Alternative bietet sich die Verwendung von Satellitendaten an. Satelliten erfassen größere Gebiete und stellen homogenere Bilddaten bereit als beispielsweise Orthophotos. Dabei stehen Sensoren mit sehr unterschiedlichen spektralen und geometrischen Auflösungen zur Verfügung. Für großflächige Klassifikationen von groben Landbedeckungsklassen bieten sich Sensoren mit mittlerer (z. B. MODIS) bis hoher Auflösung (z. B. Landsat) an (CARRÃO et al. 2008; CIHLAR 2000; FRANKLIN & WULDER 2002; KNORN et al. 2009). Sehr hoch auflösende Daten können hingegen auch gut für die Klassifikation von einzelnen Baumarten herangezogen werden (CARLEER & WOLFF 2004; IMMITZER et al. 2012b; KIM et al. 2011; WASER et al. 2014).

Durch die Landsat-Satelliten stehen kostenfreie Daten mit hoher Flächenabdeckung und 16-tägiger Wiederholungsrate zur Verfügung. Der Landsat-Sensor hat darüber hinaus Spektralkanäle im nahen und mittleren Infrarot mit hoher Relevanz für waldbezogene Fragestellungen (SCHLERF et al. 2003). Landsat-Daten können neuerdings bereits in hoher Qualität atmosphärenkorrigiert bezogen werden (VUOLO et al. 2015). Die für forstliche Fragestellungen wichtige hohe räumliche und spektrale Auflösung des WorldView-2-Sensors (WV-2) muss durch relativ hohe Beschaffungskosten und geringe Flächenabdeckung ‚erkauft‘ werden. Daher ist eine großflächige Anwendung meist finanziell nicht tragbar. Die Eckdaten beider Systeme finden sich in Tab. 1.

	Landsat-Familie	WorldView-2
Geometrische Auflösung	30 m	2 m
Anzahl Spektralkanäle	7	8
Spektralkanäle	Blau-Grün, Grün, Rot, nahes Infrarot, mittleres Infrarot 1, mittleres Infrarot 2	Küstenblau, Blau, Grün, Gelb, Rot, Red-Edge, nahes Infrarot 1, nahes Infrarot 2
Verfügbarkeit	Flächendeckend, alle 16 Tage	Auf Bestellung
Kosten	Kostenfrei beziehbar	Relativ hohe Kosten (ca. 30 USD/km ²)

Tab. 1: Eckdaten zweier Satellitensystemen (Landsat und WorldView-2)

Tab. 1: Main features of the two satellite sensors (Landsat and WorldView-2)

Das Ziel der vorliegenden Arbeit war es die jeweiligen Vorteile beider Sensoren durch die Kombination beider Datensätze zu nutzen:

- Flächendeckende multi-temporale Aufnahmen (Zeitreihen) hoch aufgelöster Daten (Landsat)
- Kleinflächige mono-temporale Aufnahmen sehr hoch aufgelöster Daten (WV2)

Die damit erzielten Synergieeffekte ermöglichen großflächige Kartierungen bei vertretbaren Datenkosten. Solch kombinierte Ansätze (*Upscaling*), wurden bereits für Baumartengruppen erfolgreich angewendet (IVERSON et al. 1989; ZHU & EVANS 1994; METZLER & SADER 2005). Eine Umsetzung zur Schätzung einzelner Baumarten innerhalb einer Baumartengruppe ist bislang nicht bekannt.

2 Methoden

Die entwickelte Methode für die Erstellung der bayernweiten Fichten- und Kiefernanteilskarten lässt sich in drei Abschnitte unterteilen (Abb. 1).

Der erste Schritt beinhaltet die überwachte, objektbasierte Klassifikation von insgesamt 43 WV2-Satellitenszenen. Dadurch werden relativ kleinflächige, über ganz Bayern verteilte, Baumartenkarten erstellt. Diese Baumartenkarten dienen im zweiten Schritt als Eingangs- bzw. Referenzdatensatz für ein *Upscaling*. Das *Upscaling* dient zur Überführung dieser Informationen auf flächendeckend zur Verfügung stehende Landsat-Daten. Die kalibrierten Modelle lassen sich dann auf die Gesamtfläche übertragen, also auch auf die Bereiche außerhalb der (relativ kleinen) WV2-Daten. Damit werden Karten in einer räumlichen Auflösung von $100\text{ m} \times 100\text{ m}$ erstellt, mit Prozentangaben von *Fichte* (*Picea sp.*), *Kiefer* (*Pinus sp.*) und *Sonstiges* pro ha-Zelle. Im dritten Arbeitsschritt werden die Teilergebnisse dann homogenisiert und zusammengeführt. Anschließend findet eine unabhängige Validierung der Gesamtergebnisse statt. Entsprechend der Aufnahmezeitpunkte der WV2-Szenen entspricht der Statuszeitpunkt zur Kartierung der Fichten- und Kiefernbestände in Bayern den Jahren 2013/2014.

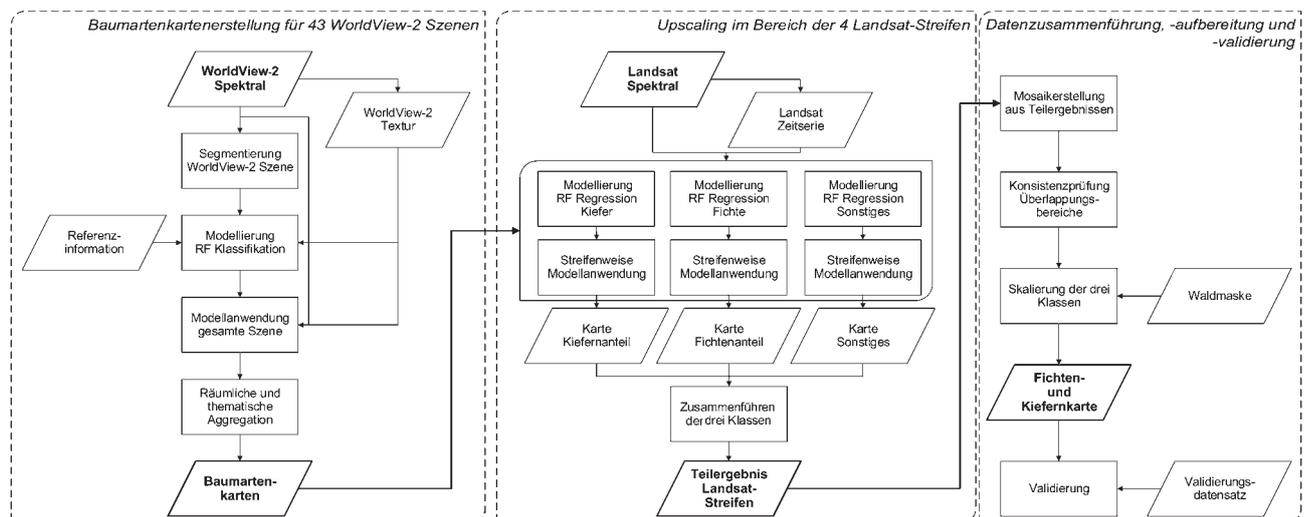


Abb. 1: Workflow der dreistufigen Methodik zur Erstellung von prozentualen Fichten- und Kiefernanteilskarten für Bayern: Erstellung der Baumartenkarten (links), flächendeckendes Upscaling der Baumartenkarten (Mitte) und Aufbereitung der Ergebnisse zu den finalen Karten und Validierung (rechts).

Fig. 1: Workflow of the three-step method for designing the Bavarian spruce and pine probability maps: Creation of tree species maps (left), area-wide upscaling of tree species maps (middle) and producing the final maps and validation (right).

2.1 Erstellung von räumlich hoch aufgelösten Baumartenkarten für kleine Teilgebiete

Für die Erstellung der räumlich sehr hoch aufgelösten Baumartenkarten („Referenz“) wurde auf WV2-Daten (Tab. 1) zurückgegriffen. Ihre Eignung für die Klassifikation von mitteleuropäischen Wäldern ist gut belegt (IMMITZER et al. 2012b; IMMITZER et al. 2012a; WASER et al. 2014). Die bearbeiteten Gebiete wurden entsprechend der forstlichen Wuchsgebiete Bayerns (WALENTOWSKI et al. 2001) sowie der Verteilung der Staatswaldflächen ausgewählt. Zusätzlich stehen Inventurpunkte als Referenzdaten zur Verfügung. Für die ausgewählten Gebiete wurden WV2-Daten aus dem Archiv bestellt bzw. explizit für das Projekt in Auftrag gegeben (Abb. 3 links).

Alle WV2-Szenen wurden zwei Vorprozessierungsschritten mit der Prozesskette CATENA (KRAUSS et al. 2013; REINARTZ 2010) des DLR unterzogen. Diese Prozesskette besteht aus mehreren Komponenten, unter anderem der parametrischen Orthorektifizierung basierend auf ORTHO (MÜLLER et al. 2005), der automatischen Atmosphärenkorrektur basierend auf ATCOR (RICHTER 1996; RICHTER et al. 2006) sowie einer Qualitätskontrolle.

Um die Klassifizierung zu optimieren wurden nicht nur spektrale Merkmale verwendet, sondern auch Texturinformation, wie beispielsweise Wavelets. Diese können bei ähnlichem oder höherem Informationsgehalt deutlich schneller berechnet werden als z. B. die weitverbreiteten Texturmaße nach Haralick (IMMITZER et al. 2014; TOSCANI et al. 2013). Zur Bestimmung der objekt-spezifischen Texturmerkmale wurde neben den Kanälen Rot und nahes Infrarot auch der Vegetationsindex NDVI (ROUSE et al. 1974) einer diskreten stationären Wavelet-Transformation (DSWT) (TOSCANI et al. 2013) unterzogen (Abb. 2c). Der NDVI basiert auf den Spektralkanälen nahes Infrarot und Rot.

Für die überwachte, objekt-basierte Klassifizierung wurden die WV2-Szenen mittels der Software eCognition automatisiert segmentiert (Abb. 2a). Ziel war, möglichst klassenreine Segmente zu erhalten. Die Segmentierungseinstellungen wurden dazu manuell an die Gegebenheiten jeder einzelnen Szene angepasst. Die weitere Prozessierung wurde mit der Statistiksoftware R 3.0.1 (R CORE TEAM 2014) umgesetzt. Für jedes Segment wurden unterschiedliche statistische Kennwerte berechnet, und zwar getrennt für die acht Spektralkanäle, den NDVI-Layer und die 12 Texturlayer. Die statistischen Kennwerte zur überwachten Klassifizierung der Segmente umfassen den Mittelwert, die Standardabweichung, Minimum, Maximum sowie zusätzlich noch die 5., 10., 25., 50., 75., 90. und 95. Perzentile. Damit standen insgesamt 231 Eingangsvariablen zur objekt-basierten Klassifikation zur Verfügung (21 Layer \times 11 Kennwerte).

Wo vorhanden, wurden den Segmenten Referenzinformationen aus Inventurdaten zugewiesen. Wenn nötig wurden zusätzliche Referenzdaten durch die manuelle Interpretation von Orthophotos und Stereoluftbildern gewonnen. Bei der Identifikation geeigneter Referenzflächen wurden insbesondere die beiden Zielklassen Fichte und Kiefer beachtet. Die Daten wurden ergänzt um weitere Nadelhölzer sowie eine Laubholz-Klasse. Um eine flächendeckende Klassifikation zu erhalten wurden zusätzlich auch Nichtwaldklassen (Wiese, Acker, Siedlung, Gewässer,...) ausgewiesen.

Als Klassifikator wurde der von BREIMAN (2001) entwickelte ‚Random Forest (RF)‘ Algorithmus verwendet. RF liefert zuverlässige Klassifikationsergebnisse bei geringer Feinabstimmung (LIAW & WIENER 2002; HASTIE et al. 2009). Vorteilhaft sind auch die Out-of-Bag-Schätzungen (OOB), die unabhängige Schätzgenauigkeiten liefern (BREIMAN 2001). Das Verfahren wird zur Gruppe der ‚ensemble learning‘-Klassifikatoren gezählt, welche die Erkennungsleistung durch die Kombination mehrerer Klassifikatoren steigern (WASKE & BRAUN 2009; KROSCHER et al. 2011). Im konkreten Fall wurden für jede Klassifikation 1000 Entscheidungsbäume erstellt. Die Reduktion der großen Anzahl an Eingangsdaten auf die wesentlichen Variablen erfolgte mit einer schrittweisen ‚recursive feature elimination‘ (GUYON et al. 2002), basierend auf den im RF-Modell ermittelten Wichtigkeitsmaßen. Dazu wurden die Eingangsdaten nach Wichtigkeit gereiht und durch Streichung der jeweils unwichtigsten (letztgereihten) Variable reduziert. Dieser Vorgang wurde bis zur Erstellung eines Modells mit nur einer Eingangsvariablen wiederholt. Von allen Modellen wurde die Gesamtgenauigkeit auf Basis der OOB Daten errechnet und das Modell identifiziert, welches mit wenigen Eingangsvariablen eine nahezu optimale Schätzgenauigkeit liefert. Dieses ‚beste‘ Modell wurde anschließend auf die gesamte WV2-Szene angewendet (Abb. 2d).

Das Klassifikationsergebnis wurde einer visuellen Überprüfung unterzogen. Dies erfolgte durch Vergleich und Interpretation von Orthophotos. Erkannte Fehlklassifikationen wurden anschließend manuell korrigiert. Unsichere Bereiche wurden auf ‚nicht eindeutig klassifizierbar‘ gesetzt und von der weiteren Analyse ausgenommen.

Um die resultierenden Baumartenkarten als Eingangsdaten für die *Upscaling*-Modellierung verwenden zu können, erfolgte eine thematische Aggregation zu den Zielklassen *Fichte*, *Kiefer* und *Sonstiges* (Abb. 2e). Diese Informationen wurden in ein für ganz Bayern eindeutig festgelegtes Raster mit $100\text{ m} \times 100\text{ m}$ großen Zellen aggregiert (Abb. 2f). Dazu wurde innerhalb jeder 1 ha-Zelle der Flächenanteil berechnet, der von den drei Klassen bedeckt wird. Diese so ermittelten Flächenanteile stellen im nächsten Schritt die zu modellierenden Zielgrößen dar.

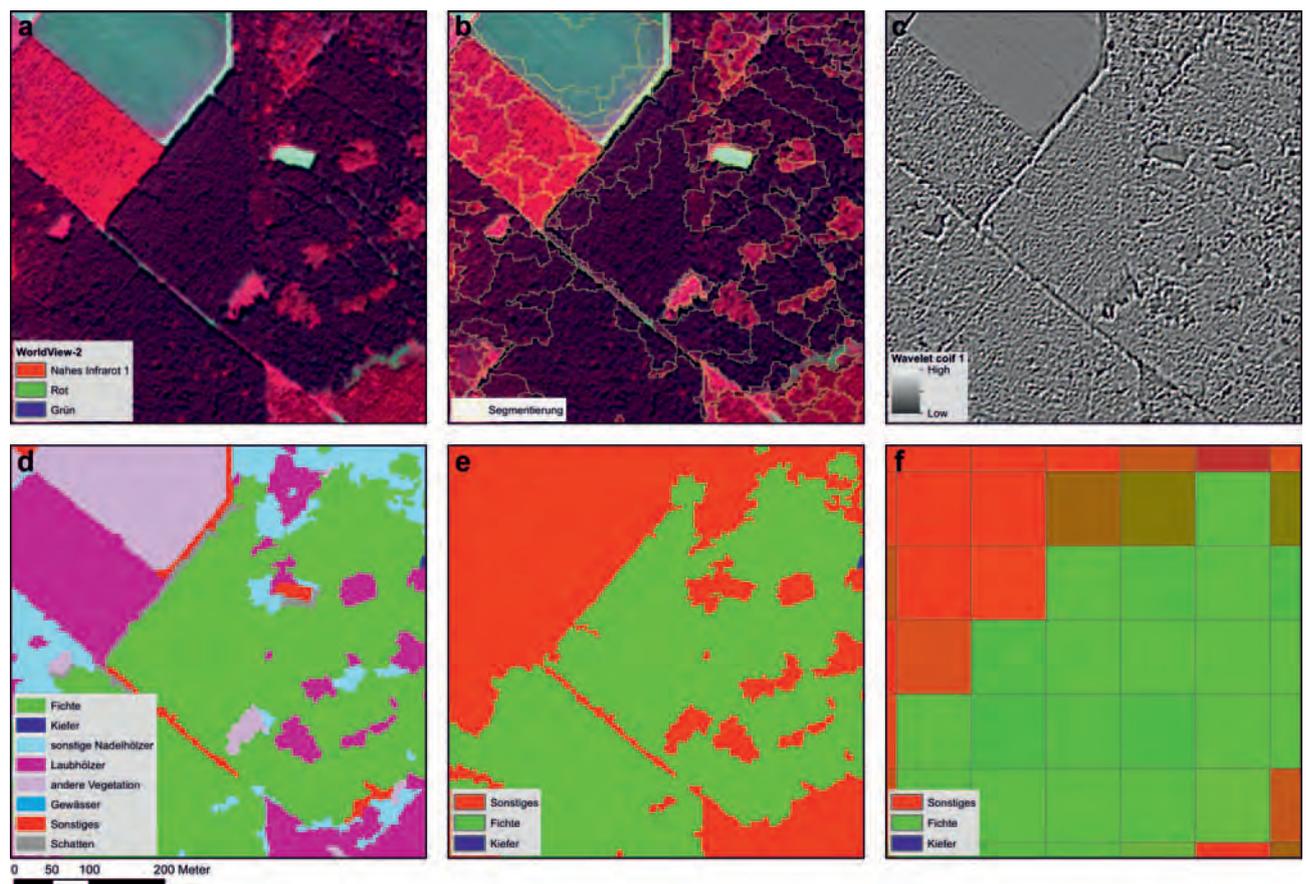


Abb. 2: Ablauf der Baumartenkartenerstellung („Referenz“) mit Datenmaterial und Zwischenergebnissen anhand eines Detailausschnittes: a) WorldView-2-Szene, b) Segmentierung, c) aus den WorldView-2-Daten generierter Textur-layer, d) detailliertes Klassifikationsergebnis, e) thematisch aggregiertes Klassifikationsergebnis, f) räumlich zu den 1 ha-Zellen aggregiertes Klassifikationsergebnis als „Referenz“ für das nachfolgende Upscaling.

Fig. 2: Workflow for the creation of the tree species map (‘reference’) with data and intermediate results by means of a subset: a) WorldView-2 scene, b) segmentation, c) from WorldView-2 data derived texture layer, d) detailed classification result, e) thematically aggregated classification result, f) aggregated classification result spatially linked to 1 ha-cells as ‘reference’ for subsequent upscaling.

2.2 Upscaling der Baumartenkarten zu flächendeckenden Fichten- und Kiefernarten

Im zweiten Schritt der Methodik erfolgte die Überführung (*Upscaling*) der teilflächig vorliegenden Informationen auf die gesamte Landesfläche. Dazu wurden in den Bereichen, die durch Baumartenkarten abgedeckt waren, *Upscaling*-Modelle erstellt und diese in weiterer Folge auf die Gebiete außerhalb der Baumartenkarten angewendet. Als Referenzinformation für die Modelle wurden die Flächenanteile der drei

Klassen in den 1 ha-Zellen aus den Baumartenkarten herangezogen (Abb. 2f). Flächendeckend zur Verfügung stehende Landsat-Satellitenszenen mehrerer Aufnahmezeitpunkte bzw. daraus abgeleitete Zeitserien wurden als erklärende Variablen verwendet. Die Abdeckung Bayerns mit Landsat-Szenen (siehe Tab. 1) wird durch vier überlappende Streifen (192 bis 195) mit jeweils maximal drei Szenen sichergestellt (Abb. 3 rechts). Durch die Überlappung und relativ hohe zeitliche Wiederholrate (alle 16 Tage je Sensor) sind für ganz Bayern wolkenfreie Landsat-Daten vorhanden.

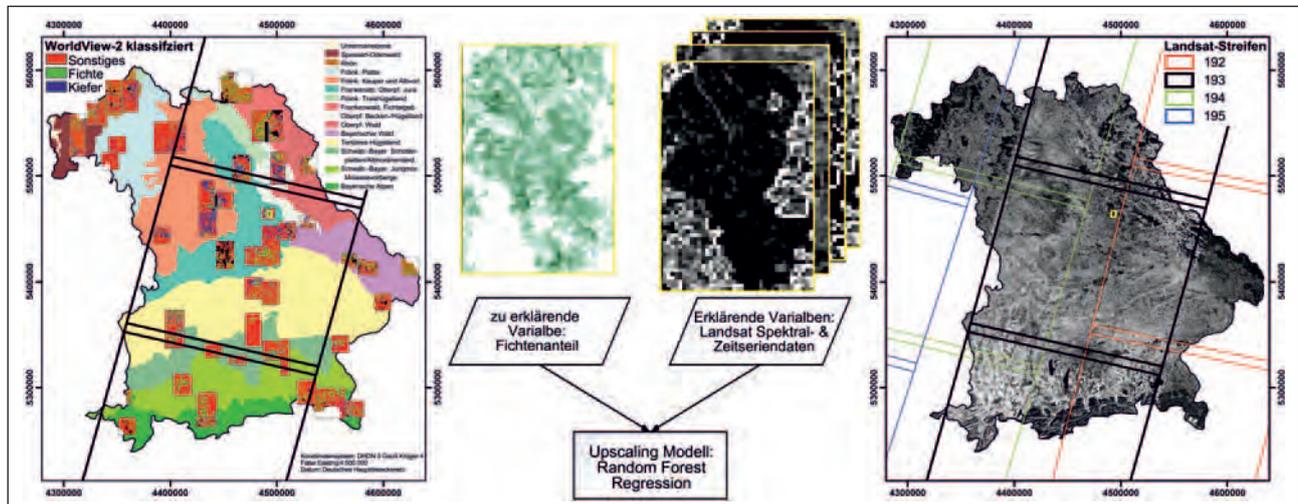


Abb. 3: Eingangsdaten für die Erstellung der Upscaling-Modelle: über Bayern verteilte Klassifikationsergebnisse (Baumartenkarten) der WorldView-2 Szenen hinterlegt mit den forstlichen Wuchsgebieten (links) und flächendeckend zur Verfügung stehende Landsat-Daten in stark überlappenden Streifen (rechts). In der Mitte ist dargestellt, wie die erklärenden Variablen zur Schätzung der zu erklärenden Variable (Flächenanteile) herangezogen werden.

Fig. 3: Input data for creation of upscaling-models: classification results (tree species maps) of the WorldView-2 scene distributed over Bavaria with forest growth zones in the background (left) area-wide available Landsat-data with high overlapping stripes (right). Middle: Usage of the explanatory variable for estimating the explanatory variable (surface area).

Für die Modellierung wurden die spektralen Informationen einzelner Szenen sowie Zeitreihendaten verwendet. Für die Extraktion der spektralen Information wurden flächendeckend jeweils zwei bis drei Szenen verteilt über die Vegetationsperioden der letzten fünf Jahre verwendet, wobei aktuellere Daten bevorzugt wurden.

Zusätzlich wurde der NDVI aller verfügbaren Szenen für die Zeitreihenanalyse berechnet und dieser innerhalb eines Jahresverlaufs analysiert. Daraus wurden pro 1 ha-Zelle verschiedene Kennwerte ermittelt, wie z. B. die Amplitude oder der kumulierte NDVI der Vegetationsperiode.

Zur Veranschaulichung der unterschiedlichen Abdeckungen sind in Abb. 3 die beiden wesentlichen Eingangsdaten gegenüber gestellt:

- die über Bayern verteilten sehr hoch aufgelösten WV2-Daten bzw. die bereits daraus abgeleiteten Baumartenkarten,
- die flächendeckend zur Verfügung stehenden Landsat-Daten; hier als Mosaik aus ursprünglich vier getrennten, jedoch stark überlappenden Streifen.

Die Upscaling-Modelle wurden getrennt für jeden Landsat-Streifen und für jede der drei Klassen (Fichte, Kiefer und Sonstiges) erstellt. Für die Modellierung wurde das RF Regressionsverfahren verwendet (BREIMAN 2001). Ähnlich zum bereits erwähnten RF Klassifikationsverfahren basiert es auf Entscheidungsbäumen. Die erstellten Modelle wurden anschließend auf die jeweiligen Landsat-Streifen angewendet (Abb. 4b und 4c).

2.3 Datenaufbereitung zu bayernweiten Kartenprodukten

Für die finale Kartenerstellung (Abb. 4, untere Zeile) und Validierung wurden die Teilergebnisse der einzelnen Landsat-Streifen im letzten Schritt zu einer bayernweiten Karte zusammengeführt (Mosaikerstellung). Dabei erfolgte eine abstandsabhängige Interpolation der sich überlappenden Ergebnisse. Bei allen Arbeitsschritten wurde darauf geachtet, dass jede 1 ha-Zelle in Summe genau 100 % Deckung aufweist. Gegebenenfalls wurde neu skaliert. Unter Berücksichtigung der Waldmaske wurden alle Bereiche außerhalb des Waldes auf 100 % Sonstiges gesetzt.

2.4 Konsistenzprüfung und Validierung

Da die Schätzung der einzelnen stark überlappenden Landsat-Streifen unabhängig voneinander erfolgte, konnten die Überlappungsbereiche zur Überprüfung der Modellstabilität (Konsistenz) herangezogen werden. Zur Validierung wurde ein regelmäßiges Raster mit 2.5 km x 2.5 km Netzweite erstellt. An allen Punkten innerhalb der Waldmaske wurden in der jeweiligen 1 ha-Zelle die drei Klassen in 5 %-Schritten durch visuelle Orthophotointerpretation geschätzt.

3 Ergebnisse

3.1 Räumlich hoch aufgelöste Baumartenkarten für kleine Teilbereiche

Die einzelnen WV-Szenen wurden in 8 bis 15 Klassen klassifiziert (Median: 11 Klassen). Die Klassenanzahl variiert, da nicht immer alle Landbedeckungsklassen vorkommen. Für die Erstellung der Klassifikationsmodelle wurden zwischen 500 und 1500 (Median: 850) Referenzpolygone pro Szene berücksichtigt. Die auf den OOB-Daten basierenden und mittels Konfusionsmatrizen ermittelten Gesamtgenauigkeiten der Klassifikationsmodelle lagen zwischen 85 und 95 % (Median: 92 %). Ein beispielhaftes Klassifikationsergebnis ist in Abb. 2d dargestellt. Vor der Verwendung als Eingangsdaten für das Upscaling wurden die Klassifikationsergebnisse thematisch und räumlich aggregiert (Abb. 2d-f). Dadurch wurden für ca. 20 % der Landesfläche flächendeckende Informationen bezüglich der prozentualen Bedeckung mit Fichte, Kiefer und Sonstiges generiert (Abb. 4 links). Von der Analyse ausgenommen wurden Flächen, die nicht eindeutig einer dieser Klassen zugewiesen werden konnten.

3.2 Upscaling in den Landsat-Streifen

Die erklärten Varianzanteile der Upscaling-Modelle je Klasse und Landsat-Streifen sind in Tab. 2 aufgeführt. Tendenziell wurden in den Streifen, für welche eine größere Anzahl an WV2-Szenen vorhanden war, bessere Ergebnisse erzielt. Ebenso zeigten die Landsat-Daten einen Einfluss auf die Modellgüte: die Verfügbarkeit von aktuelleren Szenen sowie Szenen in einer größeren Anzahl wirkten sich positiv auf die Modellgüte aus.

	192	193	194	195
Fichte	83.7	84.1	87.4	78.7
Kiefer	78.2	83.2	86.6	74.9
Sonstiges	89.2	91.4	92.0	89.0
n (WV2-Szenen)	20	33	22	6
n (Landsat-Szenen)	2	2	3	2

Tab. 2: Erklärter Varianzanteil (R^2) aus den OOB Ergebnissen für die drei Klassen pro Landsat-Streifen sowie die für die Modellerstellung verfügbare Anzahl an WV-2- bzw. Landsat-Szenen pro Streifen (n).

Tab. 2: Explanatory variance percentage of the OOB results for the three classes per Landsat-stripe and the available number of WV-2 respectively Landsat scenes per stripe (n).

3.3. Fichten- und Kiefernkarte

Die bayernübergreifenden Verbreitungskarten wurden durch die Zusammenführung der drei Klassen in den Streifen sowie durch eine anschließende entfernungsabhängige gewichtete Interpolation (Mosaikerstellung) der Überlappungsbereiche der streifenweisen Teilergebnisse erstellt (Abb. 4, untere Zeile). Die beiden Karten in Abb. 5 und Abb. 6 zeigen den kontinuierlichen Flächenanteil von Fichte bzw. Kiefer pro 1 ha-Zelle.

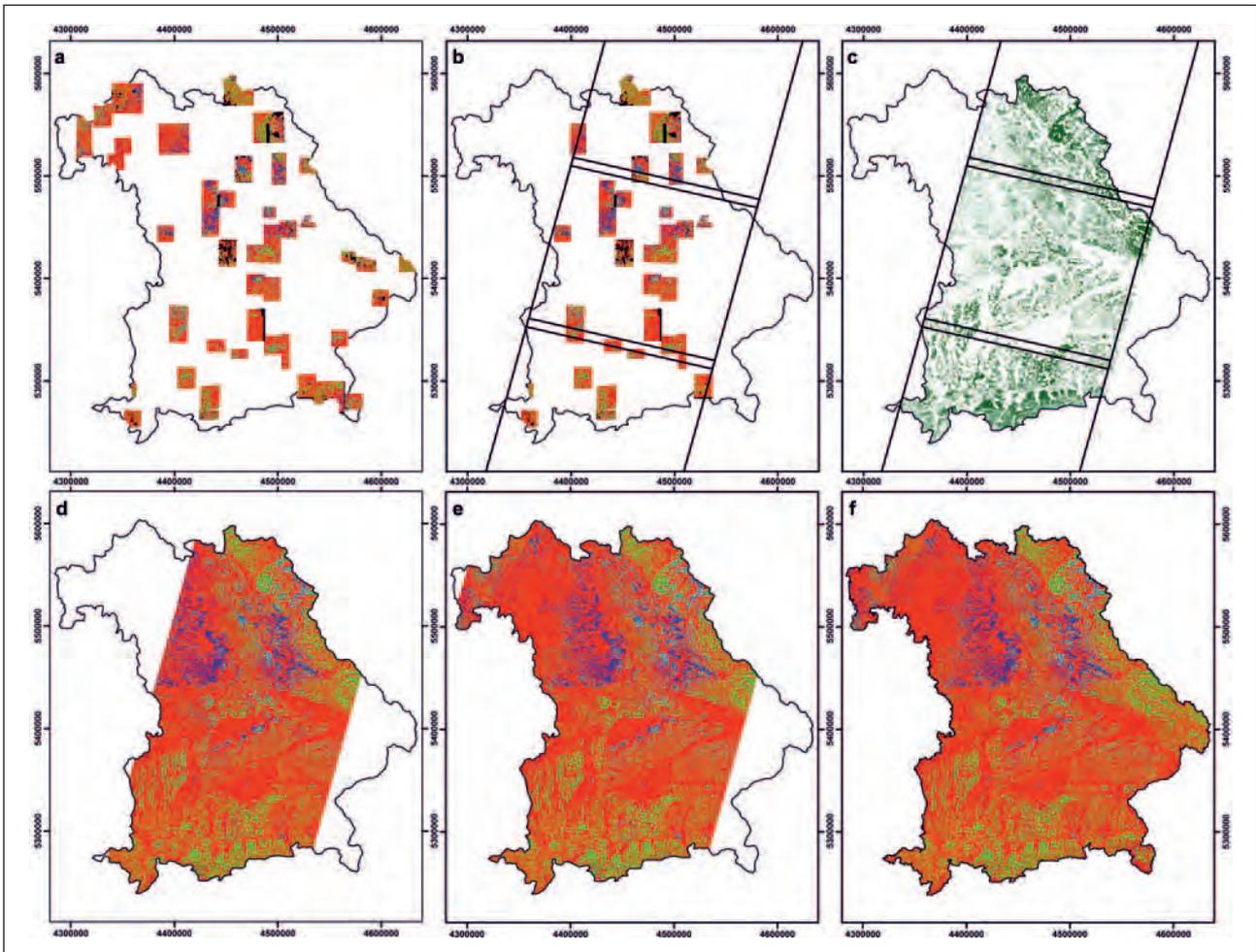


Abb. 4: Vom streifenweisen Upscaling der WorldView-2 Ergebnisse bis zur finalen Karte: a) klassifizierte WorldView-2 Szenen, b) Auswahl der Szenen für die Erstellung der Schätzmodelle für den Landsat-Streifen 193, c) Ergebnis der Upscaling-Modellanwendung für die Klasse Fichte, d) Kombination der drei Klassen im Landsat-Streifen 193, e) Mosaikerstellung aus den benachbarten Streifenergebnissen, f) fertige Karte mit allen drei Klassen.

Fig. 4: From stripe wise upscaling of the WorldView-2 results to the final map: a) classified WorldView-2 scene, b) scene selection for designing the estimation model for Landsat-stripe 193, c) result of applying the upscaling model for class 'spruce', d) combination of the three classes in Landsat-stripe 193, e) mosaicking of neighboring stripes, f) final map with the three classes.

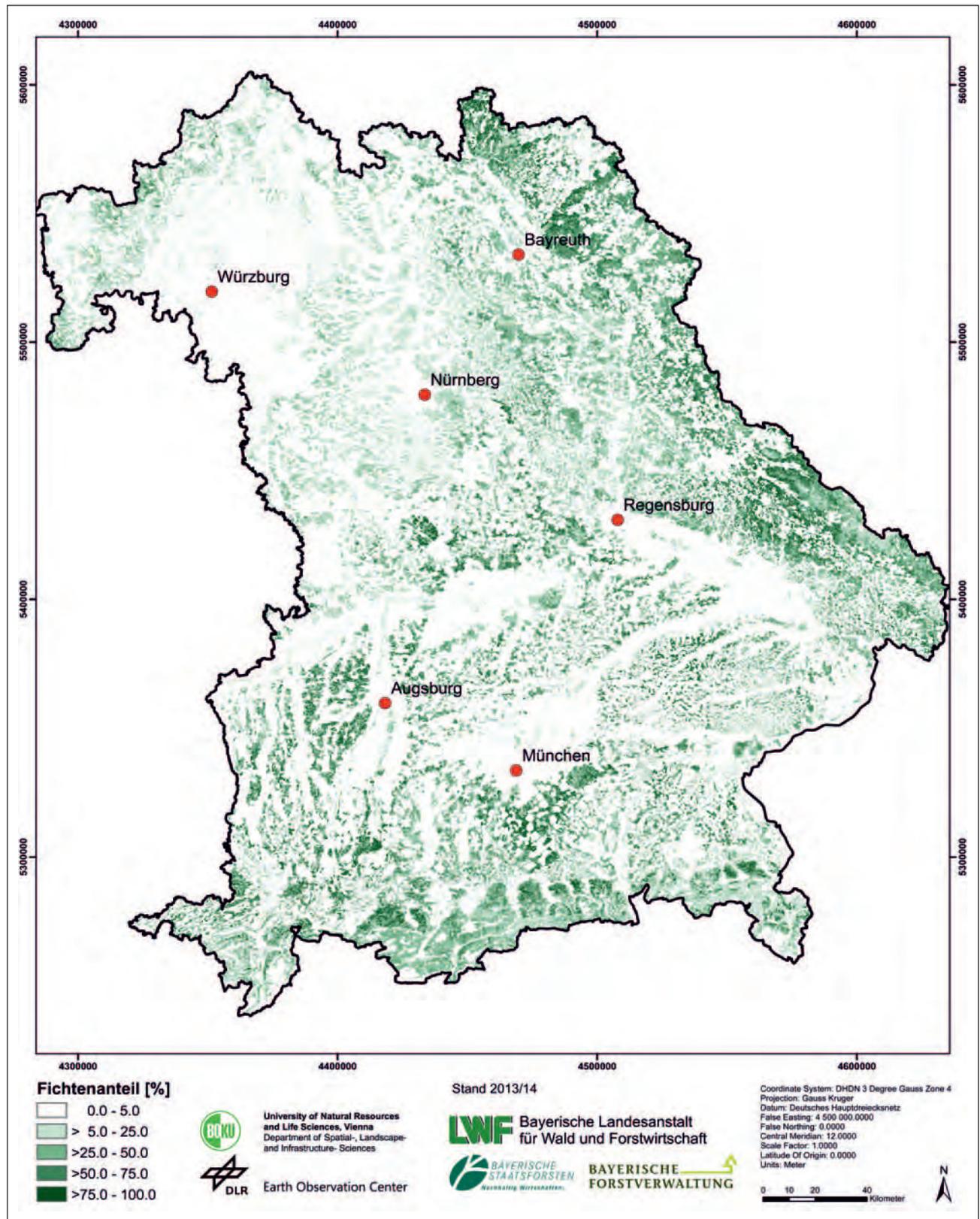


Abb. 5: Darstellung des ermittelten Fichtenanteils in Prozent pro 1 ha-Zelle (Farbsättigung entspricht Anteil der Klasse).

Fig. 5: Map with calculated spruce percentage per 1 ha-cell (color saturation equals class percentage).

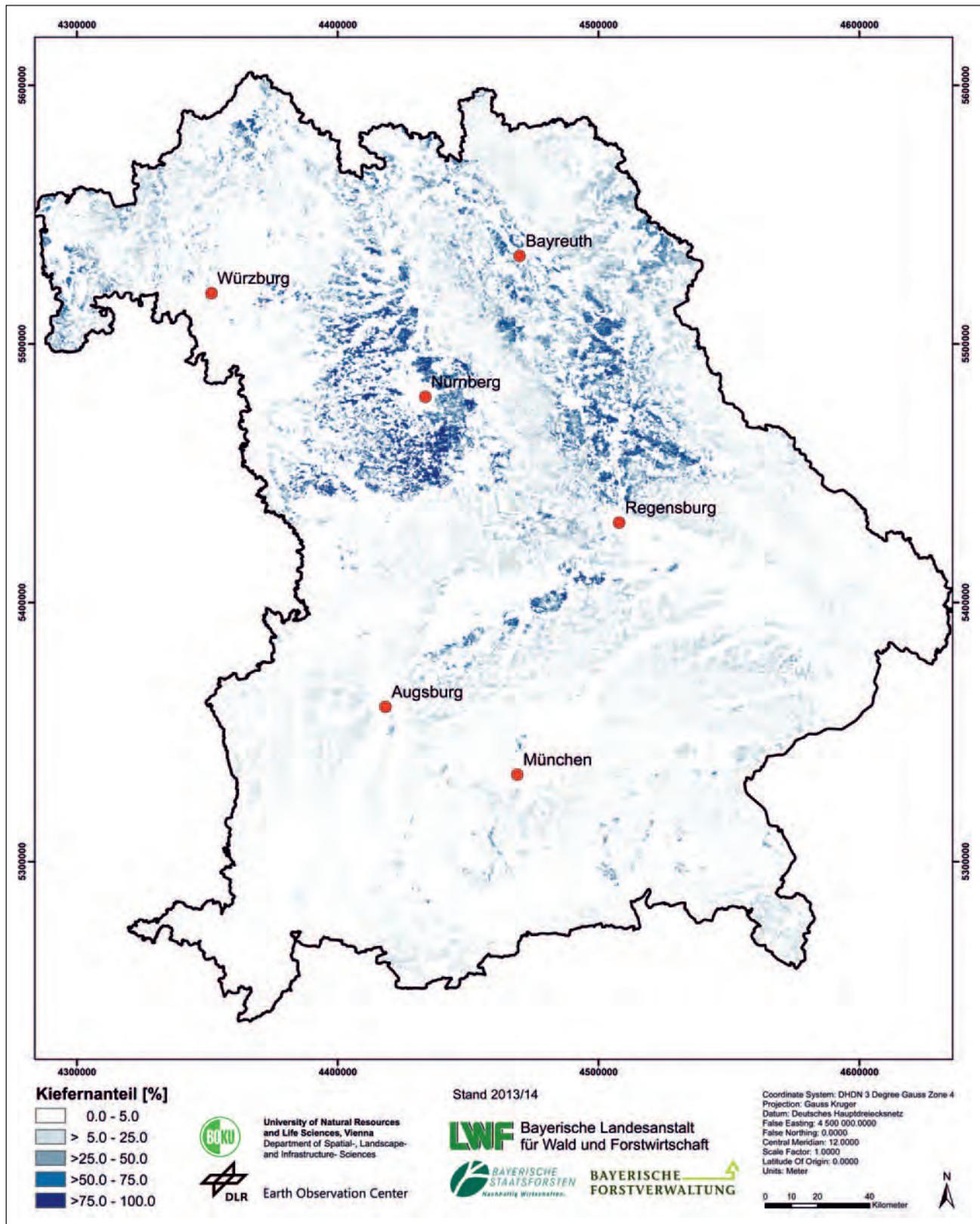


Abb. 6: Darstellung des ermittelten Kiefernanteils in Prozent pro 1 ha-Zelle (Farbsättigung entspricht Anteil der Klasse).

Fig. 6: Map with calculated pine percentage per 1 ha-cell (color saturation equals class percentage).

3.4 Konsistenzprüfung und Validierung

Beim Vergleich der Schätzergebnisse in den Überlappungsbereichen benachbarter Landsat-Streifen wurden sehr hohe Korrelationen zwischen den Schätzungen festgestellt (Tab. 3). Dies indiziert eine hohe Stabilität und räumliche Konsistenz der erstellten *Upscaling*-Modelle.

Streifen 1	Streifen 2	Klasse	R ²	RMSE	BIAS	N
192	193	Fichte	0.92	8.89	-5.41	748693
		Kiefer	0.89	5.13	0.64	
		Sonstiges	0.96	7.90	4.75	
193	194	Fichte	0.86	9.00	-2.26	1021170
		Kiefer	0.84	8.56	0.36	
		Sonstiges	0.94	7.39	1.90	
194	195	Fichte	0.90	5.43	-1.51	182026
		Kiefer	0.88	6.12	0.45	
		Sonstiges	0.94	6.05	1.04	

Tab. 3: Zusammenfassende Vergleiche der Upscaling-Ergebnisse in den Überlappungsbereichen aneinandergrenzender Landsat-Streifen. Mit N wird die Anzahl der Pixel im Überlappungsbereich bezeichnet. Die Schätzwerte der einzelnen Klassen liegen zwischen 0 und 100 Prozent.

Tab. 3 Comparison of the upscaling results in overlapping areas of neighboring Landsat-stripes. N is the number of pixels in the overlapping area. The estimated value of the single classes ranges between 0 and 100 percent.

Der Vergleich der Validierungsdaten aus der Orthophotointerpretation mit den erzielten *Upscaling*-Ergebnissen zeigt sehr gute Übereinstimmungen (Abb. 7).

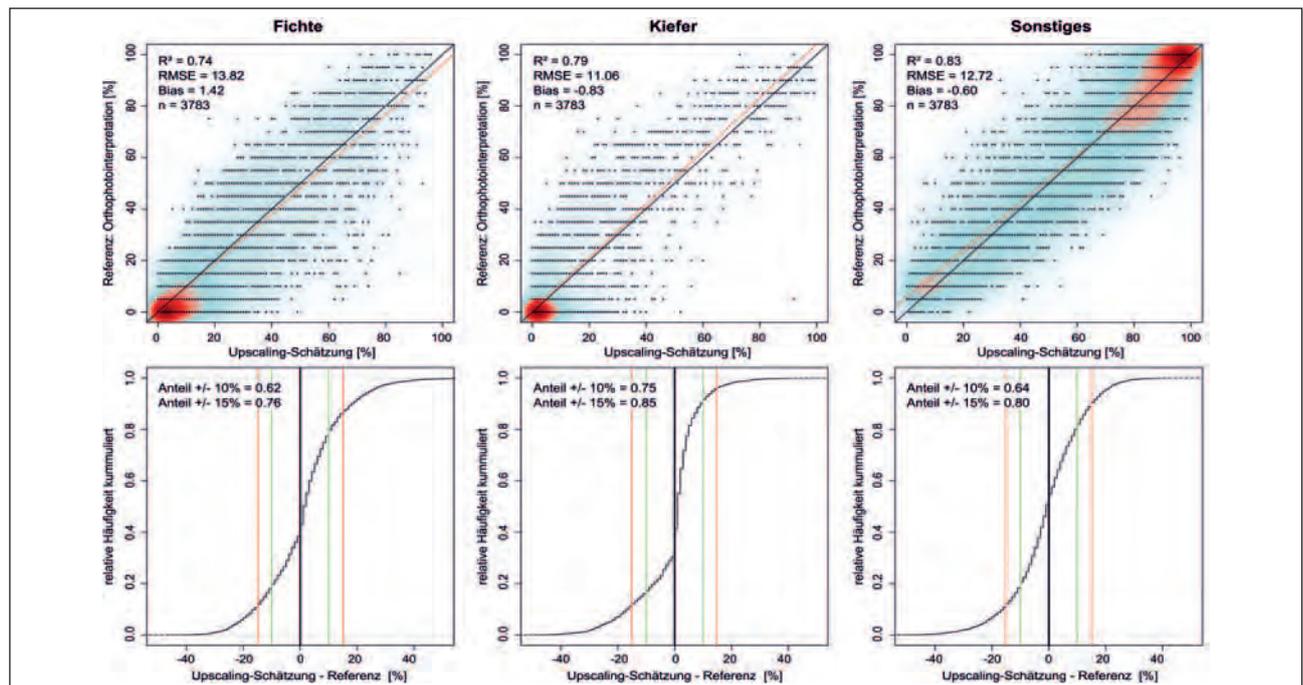


Abb. 7: Unabhängige Validierung der drei Klassen mit 3798 am Orthophoto visuell interpretierten Rasterzellen: Streudiagramme der visuellen Schätzung (in 5% Schritten) versus Modellergebnisse. In schwarz die 1:1 Linie; in rot die Ausgleichsgerade (oben). Dichtefunktionen der Differenzen aus Modellierungsergebnis und visueller Referenz. Die roten senkrechten Linien geben die wahrscheinliche Unsicherheitsmarge der visuellen Referenzschätzungen an ($\pm 15\%$); grün für angenommene $\pm 10\%$ (unten).

Fig. 7: Independent validation of the three classes applied on 3798 on orthophotos visually interpreted raster cells: Scatterplot of the visual assessment (in 5% steps) versus model results. In black 1:1 line, in red best-fit line (above). Probability density functions of the difference between model results and visual reference. The vertical lines display the probable uncertainty margin of the visual reference estimation: red: $\pm 15\%$, green: $\pm 10\%$ (below).

Das Bestimmtheitsmaß für Fichte lag bei 0.74, jenes für Kiefer bei 0.79. Berücksichtigt man die durch die Orthophotointerpretation erzielbare Genauigkeit von $\pm 15\%$ (rote senkrechte Linien in den kumulierten Häufigkeitsverteilungen) wurden an 76 % der Validierungspunkte der Fichtenanteil und an 85 % der Kiefernanteil korrekt geschätzt.

4 Diskussion

Der mittlere absolute Fehler der beiden Zielklassen betrug für Fichte 10 % und für Kiefer 7 %. Die Modellgüte der *Upscaling*-Modelle der beiden Baumarten lag dabei deutlich höher als bei vergleichbaren Studien zur Schätzung von Baumartengruppen (METZLER & SADER 2005). Bezüglich der räumlichen Verteilungsmuster besteht eine hohe Übereinstimmung mit den vom Europäischen Forstinstitute (EFI) erstellten Baumartenkarten (BRUS et al. 2012). Am auffälligsten sind Unterschiede im Detaillierungsgrad (räumliche Auflösung von 1 ha der beschriebenen Karten gegenüber 1 km² beim EFI Produkt). Auch sind bei den EFI-Karten Artefakte zu erkennen, die auf das zugrundeliegende Kriging-Verfahren zurückzuführen sind. Dies betrifft insbesondere die seltener vorkommenden Baumarten wie die Kiefer (tlw. deutliche Überschätzung).

Interessant ist auch der Vergleich der vorliegenden Kartierungen mit den BWI Daten Bayerns. Die Baumartenanteile für Bayern aus den BWI³-Daten werden mit 41.8 % für Fichte und 17.1 % für Kiefer beziffert (KLEMMT et al. 2014). Die Summe der in diesem Projekt erstellten Karten fällt für diese beiden Baumarten hingegen deutlich geringer aus. Bezogen auf die Waldfläche von 2 560 000 ha ergeben sich Anteile von 31.2 % für Fichte und 14.5 % für Kiefer. Die Unterschiede sind wahrscheinlich auf unterschiedliche Definitionen zurückzuführen. In der vorliegenden Arbeit werden die tatsächlichen Überschirmungsflächen ermittelt und dargestellt. Bei den Inventuren werden dagegen beispielsweise auch Bestandslücken innerhalb eines Fichtenforstes als „Fichte“ definiert.

5 Zusammenfassung und Ausblick

Verteilungskarten für einzelne Baumarten sind eine wichtige Informationsgrundlage für die Steuerung der Waldentwicklung, für Monitoringvorhaben oder als Eingangsdaten für diverse Modellierungsansätze. Im vorgestellten Projekt wurden flächenhafte Verbreitungskarten für Fichte und Kiefer produziert, die den Gesamtwald in Bayern darstellen. Die Karten geben dabei die prozentualen Anteile von Fichte und Kiefer in 1 ha-Zellen an. Aus diesen zellenweisen Daten lassen sich ohne Aufwand Flächenabschätzungen der überschirmten Fläche für jede beliebige (administrative) Einheit errechnen.

Zur Kartenerstellung wurde ein effizientes Verfahren entwickelt, das sich auch auf andere Baumarten anwenden lässt. Kartierungen lassen sich gegebenenfalls regelmäßig wiederholen. Weitere Kostenreduzierungen sind potentiell möglich, wenn auf alternative hochaufgelöste Fernerkundungsdaten zurückgegriffen wird. In der vorliegenden Arbeit wurden beispielsweise WV2-Daten für die Detailkartierung verwendet. Die WV2-Satellitendaten wurden gewählt wegen ihrer relativ hohen Qualität und flexiblen Datenbestellung. Nachteilig sind jedoch die hohen Datenkosten. Andere – kostengünstigere – kommerzielle Satellitendaten wurden im Projekt nicht alternativ untersucht.

Um eine wirtschaftliche Auswertung zu erreichen, wurde ein mehrstufiges Verfahren entwickelt. Dabei wurden die WV2-Daten lediglich zur ‚kleinflächigen‘ Kalibrierung kostenfreier Landsat-Satellitendaten mit 30 m Auflösung verwendet. Damit war eine bezüglich der Reproduzierbarkeit robuste Kartierung der bayernweiten Verbreitung von Fichte und Kiefer möglich. Hiermit stehen erstmalig flächendeckende, digitale Karten mit einer räumlichen Auflösung von 1 ha zur Verfügung, welche die prozentualen Anteile der von Fichte und Kiefer überschirmten Fläche darstellen. Eine unabhängige Validierung bescheinigte den Karten einen Fehler (RMSE) zwischen 11 und 14 %. Die geplante Verschneidung mit anderen Geodaten – beispielsweise mit den aus dem Projekt »KLIP 4« erarbeiteten digitalen Standortinformationen wird in nächster Zeit erfolgen. Ebenso ist ein detaillierterer Vergleich mit den BWI³ Ergebnisse geplant.

6 Literaturverzeichnis

- BREIMAN, L. 2001: Random forests. *Machine learning*, **45**(1): 5–32.
- BRUS, D.J.; HENGEVELD, G.M.; WALVOORT, D.J.J.; GOEDHART, P.W.; HEIDEMA, A.H.; NABUURS, G.J. & GUNIA, K. 2012: Statistical mapping of tree species over Europe. *European Journal of Forest Research*, **131** (1): 145–157.
- CARLEER, A. & WOLFF, E. 2004: Exploitation of very high resolution satellite data for tree species identification. *Photogrammetric Engineering and Remote Sensing*, **70** (1): 135–140.
- CARRÃO, H.; GONÇALVES, P. & CAETANO, M. 2008: Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, **112** (3): 986–997.
- CIHLAR, J. 2000: Land cover mapping of large areas from satellites: Status and research priorities. *International Journal of Remote Sensing*, **21** (6–7): 1093–1114.
- FRANKLIN, S.E. & WULDER, M.A. 2002: Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography*, **26** (2): 173–205.
- GUYON, I.; WESTON, J.; BARNHILL, S.; & VAPNIK, V. 2002: Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, **46** (1-3): 389–422.
- HASTIE, T.; TIBSHIRANI, R. & FRIEDMAN, J. 2009: *The elements of statistical learning: Data mining, inference, and prediction*. 2. Auflage. New York: Springer.
- IMMITZER, M.; ATZBERGER, C. & KOUKAL, T. 2012A: Eignung von WorldView-2 Satellitenbildern für die Baumartenklassifizierung unter besonderer Berücksichtigung der vier neuen Spektralkanäle. *Photogrammetrie – Fernerkundung – Geoinformation*, **5**: 573–588.
- IMMITZER, M.; ATZBERGER, C. & KOUKAL, T. 2012B: Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, **4** (9): 2661–2693.
- IMMITZER, M.; TOSCANI, P. & ATZBERGER, C. 2014: The Utility of Wavelet-based Texture Measures to Improve Object-based Classification of Aerial Images. *South-Eastern European Journal of Earth Observation and Geomatics*, **3**: 79–84.
- IVERSON, L.R.; COOK, E.A. & GRAHAM, R.L. 1989: A technique for extrapolating and validating forest cover across large regions. Calibrating AVHRR data with TM data. *International Journal of Remote Sensing*, **10** (11): 1805–1812.
- KIM, S.-R.; LEE, W.-K.; KWAK, D.-A.; BIGING, G.S.; GONG, P.; LEE, J.-H. & CHO, H.-K. 2011: Forest cover classification by optimal segmentation of high resolution satellite imagery. *Sensors*, **11** (2): 1943–1958.
- KLEMMT, H.-J.; NEUBERT, M.; MÖSSNANG, M. & HOPF, C. 2014: *Nachhaltig und naturnah: Wald und Forstwirtschaft in Bayern – Ergebnisse der dritten Bundeswaldinventur*. Freising: Bayerische Landesanstalt für Wald und Forstwirtschaft (LWF).
- KNORN, J.; RABE, A.; RADELOFF, V.C.; KUEMMERLE, T.; KOZAK, J. & HOSTERT, P. 2009: Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment*, **113** (5): 957–964.
- KÖLLING, C.; DIETZ, E.; FALK, W. & MELLERT, K.-H. 2009: Provisorische Klima-Risikokarten als Planungshilfen für den klimagerechten Waldumbau. *LWF-Wissen*, **63**: 31–39.
- KRAUSS, T.; ANGELO, P.D.; SCHNEIDER, M. & GSTAIGER, V. 2013: The Fully Automatic Optical Processing System CATENA at DLR. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-1/W1*: 177–183.
- KROSCHEL, K.; RIGOLL, G. & SCHULLER, B. 2011: *Statistische Informationstechnik : Signal- und Mustererkennung, Parameter- und Signalschätzung*. Berlin: Springer.
- LIAW, A. & WIENER, M. 2002: Classification and regression by randomForest. *R news*, **2** (3): 18–22.
- METZLER, J.W. & SADER, S.A. 2005: Model development and comparison to predict softwood and hardwood per cent cover using high and medium spatial resolution imagery. *International Journal of Remote Sensing*, **26** (17): 3749–3761.
- MÜLLER, R.; HOLZWARTH, S.; HABERMEYER, M. & MÜLLER, A. 2005: Ortho Image Production within an Automatic Processing Chain for hyperspectral Airborne Scanner ARES. In: *Proceedings of the EARSeL Workshop 3D-Remote Sensing*. Porto, 10.–11.6.2005
- R CORE TEAM 2014: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>.
- REINARTZ, P. 2010: The CATENA Processing Chain-Multi-Sensor Pre-processing: Orthorectification, Atmospheric Correction, Future Aspects. In: *Proceedings of Geoland Forum 6*. Toulouse, 24.–25.3.2010.
- RICHTER, R. 1996: A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing*, **17** (6): 1201–1214.

- RICHTER, R.; SCHLÄPFER, D. & MÜLLER, A. 2006: An automatic atmospheric correction algorithm for visible/NIR imagery. *International Journal of Remote Sensing*, **27** (10): 2077–2085.
- ROUSE, J.; HAAS, R.; SCHELL, J. & DEERING, D. 1974: Monitoring vegetation systems in the Great Plains with ERTS. In: *NASA Special Publication*, 309–317.
- SCHLERF, M.; ATZBERGER, C. & HILL, J. 2003: Tree species and age class mapping in a Central European woodland using optical remote sensing imagery and orthophoto derived stem density-performance of multispectral and hyperspectral sensors. In: *Geoinformation for European-wide integration. Proceedings of the 22nd Symposium of the European Association of Remote Sensing Laboratories*, Prag, 4.–6.6.2002, 413–418.
- TOSCANI, P.; IMMITZER, M. & ATZBERGER, C. 2013: Texturanalyse mittels diskreter Wavelet Transformation für die objektbasierte Klassifikation von Orthophotos. *Photogrammetrie – Fernerkundung – Geoinformation*, **2**: 105–121.
- VUOLO, F.; MATTIUZZI, M. & ATZBERGER, C. 2015: Comparison of the Landsat Surface Reflectance Climate Data Record (CDR) and manually atmospherically corrected data in a semi-arid European study area. *International Journal of Applied Earth Observation and Geoinformation* **42**: 1–10.
- WALENTOWSKI, H.; GULDER, H.-J.; KÖLLING, C.; EWALD, J. & TÜRK, W. 2001: *Die regionale natürliche Waldzusammensetzung Bayerns*. Freising: Bayerische Landesanstalt für Wald und Forstwirtschaft.
- WASER, L.T.; KÜCHLER, M.; JÜTTE, K. & STAMPFER, T. 2014: Evaluating the Potential of WorldView-2 Data to Classify Tree Species and Different Levels of Ash Mortality. *Remote Sensing*, **6** (5): 4515–4545.
- WASKE, B. & BRAUN, M. 2009: Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64** (5): 450–457.
- ZHU, Z. & EVANS, D.L. 1994: U.S. forest types and predicted percent forest cover from AVHRR data. *Photogrammetric Engineering and Remote Sensing*, **60** (5): 525–531.