

Untersuchung der Eignung logistischer Regressionsmodelle
zur Kartierung forstlicher Merkmale mit
Satelliten-Fernerkundungsdaten

DISSERTATION

zur Erlangung des Doktorgrades

des Departments Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

vorgelegt von

ANDRÉ IOST

aus Pirna

Hamburg, 2006

Genehmigt vom Department Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften
an der Universität Hamburg
auf Antrag von Professor Dr. M. Köhl
Weiterer Gutachter der Dissertation:
Herr Professor Dr. Dr. h.c. D. R. Pelz
Tag der Disputation: 31. Oktober 2006

Hamburg, den 8. Oktober 2006



A handwritten signature in black ink, consisting of a large 'R' followed by a stylized 'L' and a long horizontal stroke.

Professor Dr. Reinhard Lieberei
Leiter des Departments Biologie

Meinen Eltern

Danksagung

Der freundlichen Hilfe zahlreicher Mitmenschen verdankt die vorliegende Arbeit ihre Existenz. Ich danke Prof. Dr. Michael Köhl für seine Betreuung, Erfahrung und konstruktive Kritik. Seine steten Bemühungen um eine gesicherte Anstellung an der Universität und die Gewährung eines individuellen Forschungsstiles ermöglichten mir ein angenehmes Arbeiten.

Das Zweitgutachten wurde freundlicherweise von Prof. Dr. Dr. h. c. Dieter R. Pelz übernommen, dem ich dafür herzlich danke.

Das Projekt „Bundeswaldinventur: Weiterentwicklung des Inventurverfahrens durch Integration von Erdbeobachtungsdaten“, in dessen Kontext diese Arbeit entstand, geht u. a. auf das Engagement von Herrn F. Schmitz im BMELV zurück. Die Finanzierung erfolgte über das BMBF unter Koordination durch Dr. H. P. Lüttenberg vom DLR. Ihnen gebührt ein herzlicher Dank!

Für viele Dienste und die Lösung zahlreicher Probleme danke ich den Sekretärinnen der TU Dresden, der Universität Hamburg und der Bundesforschungsanstalt für Forst- und Holzwirtschaft in Hamburg. Darüber hinaus waren Frau Kasten und Frau Böhm von der Verwaltung der TU Dresden sehr engagiert, um nach der Verlängerung des Projektes für meine lückenlose Weiterbeschäftigung zu sorgen. Wertvolle Unterstützung erfuhr diese Arbeit durch die unermüdlichen Bemühungen von Frau Stilke und ihren Kolleginnen der BFH-Bibliothek bei der Beschaffung von Artikeln und Büchern per Fernleihe.

Meinen Kollegen aus Dresden, Hamburg und Eberswalde verdanke ich zahlreiche Anregungen und Problemlösungen. Namentlich erwähnen möchte ich Herrn T. Riedel, Frau K. Oehmichen und Herrn F. Kroihner, mit denen ich gemeinsam an diesem Projekt arbeitete. Die Herren R. Köhler und H. Ziese waren da, um Herausforderungen auf den Gebieten Geographische Informationssysteme, Rechentechnik im Allgemeinen und \LaTeX im Speziellen zu meistern. Den Volleyball-Enthusiasten gebührt Dank für viele belebende Spiele, die den notwendigen Ausgleich zu stundenlangem „Sitzsport“ schufen.

Herr M. Gellrich von der WSL in Birmensdorf (Schweiz) war es, der mich auf wertvolle, weiterführende Literatur aufmerksam machte.

Auch möchte ich der Open-Source-Gemeinschaft einen herzlichen Dank aussprechen, denn sie stellte mit einigen der in der Einleitung aufgeführten Titel hervorragende Software zur Verfügung, die kommerziellen Produkten in Nichts nachsteht.

Ohne die Unterstützung meiner Frau Susanne und meiner Familie wäre die Anfertigung dieser Arbeit ungleich schwieriger geworden. Vielen Dank für die Liebe, Kraft und große Geduld!

Kurzfassung

Zur Kartierung der forstlichen Merkmale Waldstatus, Mischung und Totholzvorkommen wurde die Eignung logistischer Regressionsmodelle untersucht. In fünf Testgebieten, die zusammen ~ 11 % der Waldfläche Deutschlands abdecken, wurden Modelle für Landsat-TM- und -ETM⁺-Daten angepasst, für ein Testgebiet zusätzlich eine multispektrale QuickBird-Szene. Als Informationsquelle diente ein Auszug der Bundeswaldinventur-Datenbank. Mit den Modellen als Basis wurden Waldmasken erzeugt. Die Modelle wurden mit der k -Block-Kreuzvalidierungsmethode auf numerische Stabilität geprüft, für die Genauigkeitsanalyse der Waldmasken wurden Orthophotos verwendet.

Das Merkmal *Waldvorkommen* war mit einer maximalen Gesamtgenauigkeit von 95,4 % (Modell) bzw. 94,6 % (Karte) mit Landsat-Daten klassifizierbar. Das für die QuickBird-Szene angepasste Modell lieferte eine vergleichsweise niedrige Gesamtgenauigkeit von 83 %, weshalb keine Waldmaske erzeugt wurde. Ein 3×3 km großer Ausschnitt innerhalb des QuickBird-Bildes wurde getrennt untersucht; dort konnte durch die Delinierung von Trainingsgebieten unabhängig von der BWI-Datenbank mit umfangreicheren Parametrisierungsdatensätzen gearbeitet werden. Die Modellanpassung gestattete dann eine Modellklassifizierungsgenauigkeit von ~ 99 %; die resultierende Waldmaske wies 94%ige Übereinstimmung mit der Referenz auf.

Weniger genau ließen sich unterschiedliche Waldmischungen klassifizieren. Bei drei Klassen—*Nadel*-, *Misch*-, *Laubwald*—waren Mischwaldpixel per Model nur marginal besser als durch Zufallswahl zu klassifizieren. Die ungemischten Klassen waren aus Sicht des Erstellers mit ca. 75–83 % Übereinstimmung zwischen Trainingsdaten und BWI-Referenz trennbar.

Das Vorkommen des in der BWI-Datenbank aufgenommenen Totholzes konnte ebenfalls erfolgreich mit logistischen Regressionsmodellen in den Landsat-ETM⁺-Szenen vorhergesagt werden. Die Streuung der Gesamtgenauigkeiten (ca. 70 bis fast 90 %) war größer als bei der Klassifizierung von Waldvorkommen.

Inhaltsverzeichnis

Abbildungsverzeichnis	xi
Tabellenverzeichnis	xiii
Verwendete Abkürzungen	xv
1 Einleitung	1
1.1 Problemstellung	1
1.2 Ziel der Arbeit	2
1.3 Gliederung der Arbeit	2
1.4 Verwendete Hard- und Software	3
2 Satellitengestützte Fernerkundungssysteme	5
2.1 Grundlagen der Fernerkundung	5
2.2 Das Sensorensystem Landsat	6
2.3 Das Sensorensystem QuickBird	7
2.4 Datenverfügbarkeit	8
2.4.1 Landsat	9
2.4.2 QuickBird	10
2.4.3 Visueller Vergleich beider Systeme	11
3 Die logistische Regression	13
3.1 Einführung	13
3.2 Theoretische Grundlagen	15
3.2.1 Aufbau von Generalized Linear Models	15
3.2.2 Schätzverfahren für Generalized Linear Models	18
4 Literaturüberblick	21
4.1 Anwendungsbeispiele für fernerkundungsgestützte Waldkartierung	21
4.1.1 Landsat	21
4.1.2 QuickBird und andere hoch auflösende Systeme	24
4.2 Logistische Regression zur Klassifizierung von Fernerkundungsdaten	26
5 Testgebiete und Datengrundlagen	29
5.1 Die Testgebiete im Überblick	29
5.1.1 Testgebiet 1: Westliche Mittelgebirgsschwelle	30
5.1.2 Testgebiet 2: Nordostdeutsches Tiefland	30
5.1.3 Testgebiet 3: Östliche Mittelgebirgsschwelle	32
5.1.4 Testgebiet 4: Alpen	32

5.1.5	Testgebiet 5: Pfälzerwald	32
5.2	Datengrundlagen	33
5.2.1	BWI-Daten	33
5.2.2	Landsat-TM- und -ETM ⁺ -Szenen	33
5.2.3	QuickBird-Szenen	35
5.2.4	Orthophotos	36
5.2.5	Topographische Karten	37
6	Methodik	39
6.1	Datenvorbereitung für TM- und ETM ⁺ -Szenen	39
6.1.1	NDVI-Berechnung, Grauwertextraktion, Datensynthese	39
6.1.2	Ausschluss von Gewässerflächen	40
6.2	Datenvorbereitung für QuickBird-Szene	41
6.2.1	NDVI-Berechnung, Grauwertextraktion, Datensynthese	41
6.2.2	Ausschluss von Gewässerflächen	41
6.3	Logistische Regression zur Kartierung von Waldvorkommen	41
6.3.1	Zusammenstellung der Datensätze	41
6.3.2	Vorbereitende Untersuchungen	42
6.3.3	Uni- und multivariate Analysen	42
6.3.4	Beurteilung und Interpretation der Regressionsmodelle	42
6.3.5	Kartenerstellung	47
6.3.6	Validierung	49
6.3.7	Verfahrenstest ohne terrestrische Daten	50
6.3.8	Zum Einfluss der Testgebietswahl	51
6.4	Logistische Regression zur Kartierung von Waldmischungen	53
6.5	Logistische Regression zur Kartierung von Totholzvorkommen	54
6.5.1	Zusammenstellung der Datensätze	55
6.5.2	Uni- und multivariate Analysen, Validierung	55
7	Ergebnisse	57
7.1	Merkmal <i>Waldstatus</i>	57
7.1.1	Landsat TM	58
7.1.2	Landsat ETM ⁺	59
7.1.3	Klassifizierung und Kartierung ohne terrestrische Daten	66
7.1.4	QuickBird	68
7.1.5	Einfluss der Testgebietswahl	70
7.1.6	Diskussion	72
7.2	Merkmal <i>Mischung</i>	73
7.3	Merkmal <i>Totholzvorkommen</i>	79
7.3.1	Regressionsmodelle	79
7.3.2	Validierung der Regressionsmodelle	79
8	Zusammenfassung	83
8.1	Allgemeine Eigenschaften des Verfahrens	83
8.2	Merkmal <i>Waldstatus</i>	84
8.3	Merkmal <i>Mischung</i>	84
8.4	Merkmal <i>Totholzvorkommen</i>	85
8.5	Wahl der Testgebiete	85
	Literaturverzeichnis	87

Abbildungsverzeichnis

2.1	Das elektromagnetische Spektrum	5
2.2	Landsat 7 schematisch	7
2.3	Die Systeme Landsat und QuickBird	9
2.4	Vergleich räumliche Auflösung ETM ⁺ vs. QuickBird	11
3.1	Logistisches Modell am Beispiel <i>Waldstatus</i>	15
5.1	Altersstruktur und Baumarten der Testgebiete im Vergleich	30
5.2	Lage der Testgebiete	31
5.3	Lage der QuickBird-Kacheln in TG5	36
6.1	Spatial Modeler zur Kartenerstellung	47
6.2	Entstehung einer Waldkarte	48
6.3	Stichprobennetz zur Validierung der Waldmaske	51
6.4	Kleinsttestgebiet mit Validierungsnetz	52
7.1	Modellvergleich für Landsat TM, Merkmal <i>Waldstatus</i>	58
7.2	Unterschiedliche Überlappung im ETM ⁺ -Band 2	62
7.3	Probleme bei der Waldkartierung	63
7.4	Ergebnisse der Kreuzvalidierung der ETM ⁺ -basierten Modelle	65
7.5	Auswirkung des Medianfilters	67
7.6	3D-Streudiagramm des Merkmals <i>Mischung</i> (TG5 Süd)	74
7.7	Unterschiedliche Trennbarkeit der Mischungsklassen	74
7.8	Kreuzvalidierung des Totholzgesamtmodells	81

Tabellenverzeichnis

2.1	Geometrische und spektrale Auflösung von TM und ETM ⁺	7
2.2	Geometrische und spektrale Auflösung von QuickBird	8
5.1	Übersicht der BWI-Traktecken pro Testgebiet	34
5.2	Im WEI-BWI-Projekt verwendete Landsat-Daten	35
6.1	SPSS-Ausgabe: Beispiel einer Kontingenztafel	44
7.1	Koeffizienten; Modelle für Merkmal <i>Waldstatus</i> (TM)	59
7.2	Klassifikationsergebnisse; Modelle für Merkmal <i>Waldstatus</i> (TM)	60
7.3	Koeffizienten; Modelle für Merkmal <i>Waldstatus</i> (ETM ⁺)	60
7.4	Klassifikationsergebnisse; Modelle für Merkmal <i>Waldstatus</i> (ETM ⁺)	61
7.5	Vergleich Modellklassifizierungsergebnisse für Merkmal <i>Waldstatus</i> (QB)	70
7.6	Validierungsergebnisse; Kleingebiets-Waldkartierung (QB)	71
7.7	Koeffizienten; Modelle für Merkmal <i>Mischung</i> (ETM ⁺)	76
7.8	Klassifikationsergebnis; 1-phasiges Modell für <i>Mischung</i> (ETM ⁺)	76
7.9	Klassifikationsergebnis; 2-phasiges Modell für <i>Mischung</i> (ETM ⁺)	77
7.10	Klassifikationsergebnis; 2-phasiges, realistisches Modell für <i>Mischung</i> (ETM ⁺)	78
7.11	Klassifikationsergebnisse; Modelle für Merkmal <i>Totholz</i> (ETM ⁺)	80

Verwendete Abkürzungen

ASCII	American Standard Code for Information Interchange
AVHRR	Advanced Very High Resolution Radiometer
BFH	Bundesforschungsanstalt für Forst- und Holzwirtschaft
BFI	Blattflächenindex
BMBF	Bundesministerium für Bildung und Forschung
BMELV	Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz
BRD	Bundesrepublik Deutschland
BWI	Bundeswaldinventur
CLC	CORINE Land Cover
CORINE	Coordination of Information on the Environment
DEM	Digital Elevation Model
DHM	Digitales Höhenmodell
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DTK25-V	Digitale Topographische Karte 1:25.000—Vorläufige Version
ETM⁺	Enhanced Thematic Mapper Plus
ESRI	Environmental Systems Research Institute
EU	Europäische Union
FE	Fernerkundung
FFH	Flora, Fauna, Habitat
GIMP	GNU Image Manipulation Program
GIS	Geographisches Informationssystem
GLM	Generalized Linear Model
GW	Grauwert
GNU	GNU's Not Unix (Rekursives Akronym)
HRVIR	High Resolution Visible/Infrared
ICP Forests	International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests

Verwendete Abkürzungen

IRS	Indian Remote Sensing Satellite
JACIE	Joint Agency Commercial Imagery Evaluation
JRC	Joint Research Centre
<i>k</i>NN	<i>k</i> nächste Nachbarn bzw. Nearest Neighbours
LAI	Leaf Area Index
LogReg	Logistische Regression
MCPFE	Ministerial Conference on the Protection of Forests in Europe
MIR	Mittleres Infrarot
ML	Maximum Likelihood
MEZ	Mitteleuropäische (Winter-) Zeit
MW	Mischwald
n/v	nicht verfügbar
NASA	National Aeronautics and Space Administration
NIR	Nahes Infrarot
NDVI	Normalized Difference Vegetation Index
PC	Personal Computer
Pixel	Picture Element
QB	QuickBird
SAS[®]	ursprünglich: Statistical Analysis System, heute als Markenname verwendet
SLC	Scanline Corrector
SPOT	Systeme Probatoire d'Observation de la Terre
SPSS[®]	ursprünglich: Statistical Package for the Social Sciences, heute als Markenname verwendet
TG	Testgebiet
TIR	Thermisches Infrarot
TM	Thematic Mapper
TREES	Tropical Ecosystem Environment Observations by Satellite
USA	United States of America
USGS	United States Geological Survey
VD	Validierungsdatensatz
VHR	Very High Resolution
VI	Vegetationsindex/-indizes
WEI-BWI	Bundeswaldinventur: Weiterentwicklung des Inventurverfahrens durch Integration von Erdbeobachtungsdaten

WG	Wuchsgebiet
WRS	Worldwide Reference System
WWSK	Wald-Wahrscheinlichkeit

Kapitel 1

Einleitung

1.1 Problemstellung

Die Waldbewirtschaftung erfolgt in Deutschland gemäß den Grundsätzen der forstlichen Nachhaltigkeit. Gesetzliche Grundlagen dafür liefern sowohl das Bundeswaldgesetz (BUNDESMINISTERIUM FÜR ERNÄHRUNG, LANDWIRTSCHAFT UND FORSTEN, 1998a) als auch die jeweiligen Landeswaldgesetze, beispielsweise das Waldgesetz für den Freistaat Sachsen (SÄCHSISCHES STAATSMINISTERIUM FÜR LANDWIRTSCHAFT, ERNÄHRUNG UND FORSTEN, 1992). Ein wichtiges Instrument zur Sicherstellung dieser Nachhaltigkeit stellt die Bundeswaldinventur (BWI) dar. Sie soll umfassende Informationen zu zahlreichen Kenngrößen der Wälder¹ liefern (BUNDESMINISTERIUM FÜR ERNÄHRUNG, LANDWIRTSCHAFT UND FORSTEN, 1998b). Dabei werden weitestgehend „klassische“ Informationen gesammelt, um die nationale Forstwirtschaft und Politik mit Entscheidungsgrundlagen zu versorgen.

Forstwirtschaft im weiteren Sinne unterliegt inzwischen jedoch längst nicht mehr nur nationalen Gesetzen und Verträgen, sondern ist in einen internationalen Kontext eingebettet. Neben dem grenzüberschreitenden Handel mit Holz sei das International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) genannt, welches inzwischen 42 Mitgliedsstaaten umfasst. Weitere Beispiele sind das Kyoto-Protokoll der United Nations Framework Convention on Climate Change (UNITED NATIONS, 1997, 2002) und die FFH-Richtlinien (Flora, Fauna, Habitat) für das EU-weite Naturschutzgebietsnetz „Natura 2000“ (RAT DER EUROPÄISCHEN GEMEINSCHAFTEN, 1992). Beide stellen internationale Verträge dar, zu deren Ratifizierung sich Deutschland verpflichtet hat. Neben der Bodenzustandserhebung liefert die BWI bereits seit Abschluss der ersten Inventur Daten zu Kohlenstoffmengen für das Kyoto-Protokoll. Durch eine mögliche Erweiterung des Aufnahmekatalogs künftiger Bundeswaldinventuren um ökologische Parameter² bestünde die Chance, den aus den Verträgen erwachsenden Informationsbedarf besser decken und den Stellenwert der BWI vergrößern zu können.

Die Durchführung der zweiten Bundeswaldinventur erfolgte mit einem Mittelaufwand von ca.

¹ Betriebsart, Eigentumsart, Bestandesstruktur, Baumarten, Alter, Baumdurchmesser, Baumhöhe an ausgewählten Probestämmen, Walderschließung (Forstwege) in den neuen Bundesländern, Geländeform, Schäden, Waldränder, Totholz

² zum Beispiel Informationen, die eine erweiterte Kohlenstoffbilanzierung oder Einschätzung notwendiger Naturschutzmaßnahmen ermöglichen

18.629.000 € und bedurfte inklusive Vorbereitung und Datenauswertung eines Zeitraumes von über 4 Jahren (POLLEY ET AL., 2006). Die Berichtspflichten auf der einen Seite, der Zeit- und Kostenaufwand auf der anderen Seite lassen eine Kluft entstehen. Zur Einsparung von Kosten wäre eine Streckung des Intervalls zwischen aufeinander folgenden Inventuren angezeigt—allerdings wäre damit der aus den internationalen Verträgen erwachsende Informationsbedarf schwieriger zu decken.

1.2 Ziel der Arbeit

Das Ziel des BMBF-Projektes „Bundeswaldinventur: Weiterentwicklung des Inventurverfahrens durch Integration von Erdbeobachtungsdaten (WEI-BWI)“ war es deshalb, nach geeigneten Möglichkeiten zur Optimierung des jetzigen Inventurdesigns der BWI zu suchen. Darüber hinaus wurden Aktualisierungsverfahren untersucht, die eine kostengünstige Fortschreibung der BWI-Ergebnisse bzw. Pflege der Daten zwischen den Inventurzeitpunkten gestatten. Als Schlüsselkomponenten kamen dabei Fernerkundungsdaten³ und neue Inventurverfahren in Betracht. Hauptauftragnehmer des Projektes war das Institut für Forstökologie und Walderfassung in Eberswalde, welches der Bundesforschungsanstalt für Forst- und Holzwirtschaft (BFH) mit Sitz in Hamburg angehört. Die Technische Universität Dresden bzw. später die Universität Hamburg beteiligten sich als Untervertragnehmer der BFH.

Die vorliegende Arbeit entstand als Teil dieses Projektes. Sie befasst sich speziell mit dem Aspekt der Klassifizierung von multispektralen Satelliten-Szenen unterschiedlicher Auflösung auf Basis eines logistischen Regressionsansatzes. Die Wahl des Verfahrens liegt in der Fähigkeit des logistischen Regressionsmodells begründet, nominale Daten verarbeiten zu können. Eine lineare Regression auf Basis der kleinsten Abweichungsquadrate würde diesen Anforderungen nicht gerecht werden können. Als Klassifizierungsmerkmale kamen dabei der Waldstatus, die Baumartenmischung und das Vorhandensein von Totholz in Betracht.

Zur Parametrisierung der Klassifizierungsmodelle wurden BWI-Daten herangezogen. Nach erfolgreicher Aufstellung eines funktionierenden Modells sollten über die Klassifikation Karten erstellt werden, die eine visuelle Repräsentation des kartierten Merkmals ermöglichen würden.

1.3 Gliederung der Arbeit

Diese Arbeit ist in mehrere Themenkomplexe unterteilt, welche die einzelnen für die Bearbeitung relevanten Wissensgebiete behandeln. Im Kapitel 2 wird in aller Kürze das Konzept der Fernerkundung beleuchtet und ein kurzer Überblick der verwendeten Sensorensysteme geboten. Mit der Theorie der logistischen Regression befasst sich Kapitel 3. Im sich anschließenden Kapitel ist eine Zusammenfassung des Kenntnisstandes zum Thema der Arbeit zu finden. Über die Testgebiete und das verwendete Datenmaterial wird im Kapitel 5 informiert, während Kapitel 6 die Methodik zur Bearbeitung der Fragestellung umreißt. Die letzten zwei Abschnitte sind den Ergebnissen und deren Diskussion sowie einer Zusammenfassung gewidmet.

³ zur Einbettung des Projektes in staatliche Bemühungen um die Fernerkundungsnutzung siehe auch BACKHAUS (2004)

1.4 Verwendete Hard- und Software

Die Analysen für das Projekt und diese Arbeit wurden auf einer Reihe von PC-Systemen durchgeführt. Spezielle Hardware war nicht erforderlich, da keinerlei Digitalisierungen und Messungen an Bildmaterial im Papierformat vorgenommen wurden.

Ein breites Spektrum an Software kam zum Einsatz, um die Daten aufzubereiten, zu analysieren und darzustellen. Microsoft[®] Excel und Access 2003 dienten der Bearbeitung der BWI-Datenbankauszüge. Die eigentlichen Untersuchungen zur logistischen Regression fanden unter Zuhilfenahme von SPSS[®] 12.0.2 statt. Weitere statistische Berechnungen erfolgten mit SAS[®] V8. Erdas Imagine[®] 8.7 sowie ESRI[®] ArcView[™] 3.3 und ArcGIS[™] 9.0 waren notwendig, um einerseits die Fernerkundungsdaten bearbeiten zu können und andererseits eine Visualisierung und weitere Bearbeitung der gefundenen Zusammenhänge zu ermöglichen. Für den \TeX tsatz der vorliegenden Arbeit wurde $\text{\LaTeX} 2_{\epsilon}$ verwendet. The GIMP erwies sich als nützliches Werkzeug zur Bearbeitung und Konvertierung von Abbildungen.

Kapitel 2

Satellitengestützte Fernerkundungssysteme

2.1 Grundlagen der Fernerkundung

Das weite Betätigungsfeld der Fernerkundung (FE) wird seit mehr als 100 Jahren bearbeitet (ALBERTZ, 2001). Insofern verwundert es nicht, dass es bereits unzählige Texte zur Einführung und Vertiefung in dieses Thema gibt. Aus diesem Grund sollen hier zur Einstimmung nur in aller Kürze einige Grundlagen der FE ausgeführt werden. So facettenreich wie die FE selbst ist, fallen auch die Feinheiten der möglichen Definitionen aus. Nach ALBERTZ (2001) ist „Fernerkundung“ ein indirektes Beobachtungsverfahren, welches berührungsfrei Informationen von Objekten auf der Erde sammelt, in dem es von den Objekten emittierte/reflektierte elektromagnetische Strahlung bzw. Wellen registriert. Als Trägersysteme kommen in den meisten Fällen Flugzeuge oder Satelliten in Frage. Einen umfassenden Überblick bereits im Einsatz befindlicher bzw. geplanter Satelliten wird in STÜMER (2004) gegeben.

Innerhalb des elektromagnetischen Spektrums (Abbildung 2.1) gibt es mehrere Bereiche, die durch spezielle Sensoren erfasst werden. Besonderes Interesse gilt den Bereichen des sichtbaren „Lichtes“, der Infrarot- und Wärmestrahlung sowie den Radar-/Mikrowellen. Die Messgeräte kön-

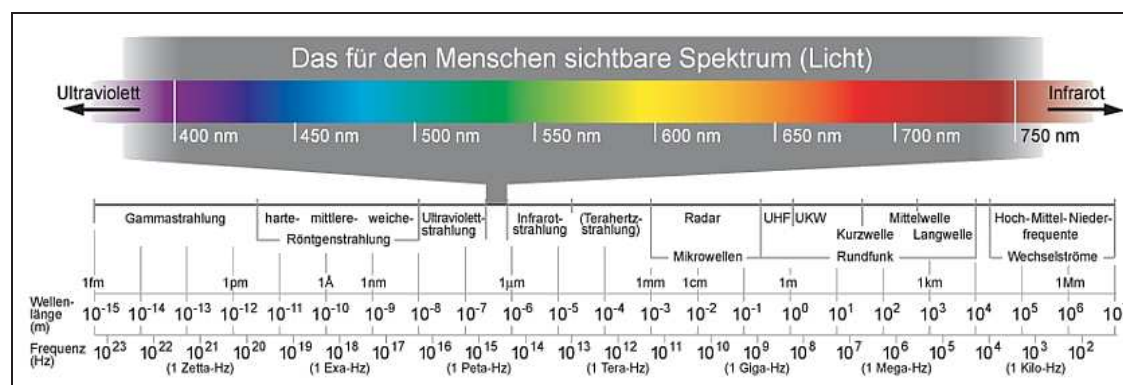


Abbildung 2.1: Das elektromagnetische Spektrum (© 2005 WIKIPEDIA)

nen passiv oder aktiv arbeiten: Passive Sensoren registrieren ausschließlich reflektierte Sonnenstrahlung bzw. emittierte Eigenstrahlung der betrachteten Objekte, aktive Systeme senden hingegen selbst ein „Mess-Signal“ (v. a. Laser- und Radarwellen) aus, dessen reflektierte Energie aufgezeichnet wird.

FE-Erzeugnisse kommen in unterschiedlichsten Bereichen zum Einsatz. Im zivilen Sektor reichen die Nutzungsmöglichkeiten vom einfachen Luftbild über Kartographie, Meteorologie, Geologie etc. bis hin zu Ökosystem-Forschung und Umweltmonitoring. Eine umfassende Aufzählung würde den Rahmen dieser kurzen Einführung übersteigen. Ein informativer Überblick ist ebenfalls in ALBERTZ (2001) zu finden.

2.2 Das Sensorensystem Landsat

Seit dem Start des ersten Landsat-Satelliten im Jahre 1972—zu dieser Zeit noch als „Earth Resources Technology Satellite“ bezeichnet—umkreisen ohne Unterbrechung Erdbeobachtungssysteme dieses Forschungsprogramms¹ die Erde. Zwar scheiterte die Inbetriebnahme des Landsat-6-Satelliten im Jahre 1993, aber die Langlebigkeit von Landsat 5 und der 1999 erfolgte Start von Landsat 7 stellten die Kontinuität des Programms sicher.

Mit einer Pixelgröße von 30×30 m zählen die Sensorensysteme der Landsat-5- und -7-Satelliten zu FE-Plattformen mittlerer räumlicher Auflösung. Typischerweise deckt eine Landsat-Vollszene eine Fläche von ca. 31.000 km^2 ab (183 km Scanstreifenbreite \times 170 km Länge). Es besteht jedoch auch die Möglichkeit, so genannte Viertelszenen zu kaufen. Die Preise für Vollszene bewegen sich zwischen 0 und etwa $0,0257 \text{ \$/km}^2$ —abhängig vom Aufnahmedatum und Grad der Korrektur/Vorverarbeitung der Szene (EURIMAGE, 2005). Seit dem Ausfall der Bewegungskompensation (siehe Seite 10) werden auch neu akquirierte Produkte zu reduziertem Preis angeboten.

Der im Rahmen *dieses* Projektes verwendete FE-Datensatz beinhaltet sowohl Landsat-5- als auch 7-Szenen. Beide Systeme umkreisen die Erde in einer Höhe von 705 km und überfliegen dieselbe Stelle in 16-tägigem Intervall. Die Kanäle 1–5 und 7 beider Systeme differieren hinsichtlich Auflösung und aufgezeichnetem Spektralbereich nicht. Unterschiede bestehen im Kanal 6 (thermisches Infrarot). Zudem bietet die neueste Landsat-Generation mit dem ETM⁺-Sensorensystem (Enhanced Thematic Mapper Plus) gegenüber dem TM-System von Landsat 5 (Thematic Mapper) einen zusätzlichen panchromatischen Kanal (Pan). Eine Zusammenstellung der Charakteristika beider Systeme bietet Tabelle 2.1 auf der nächsten Seite, während Abbildung 2.2 einen schematischen Überblick über die Subsysteme des Landsat-7-Satelliten vermittelt. In Abbildung 2.3 auf Seite 9 sind Landsat und QuickBird gegenübergestellt. Aufgrund der langen Historie der Landsat-Reihe existiert eine Vielzahl von Veröffentlichungen zu diesem Thema. Einige davon werden in Kapitel 4 auf Seite 21 aufgeführt. Es wird daher an dieser Stelle auf die Darstellung weiterer technischer Spezifika verzichtet.

¹ gemeinschaftlich betrieben durch National Aeronautics and Space Administration (NASA) und United States Geological Survey (USGS)

Tabelle 2.1: Geometrische und spektrale Auflösung der Instrumente TM und ETM⁺

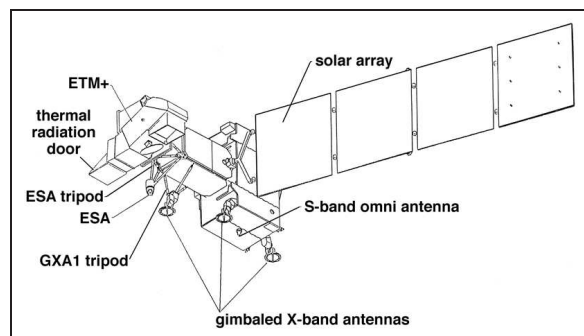
Kanal	Geometrische Auflösung	Spektrale Auflösung	Bereich des Spektrums
	in m	in nm	
1	30	450–520	sichtbares Blau
2	30	520–600	sichtbares Grün
3	30	630–690	sichtbares Rot
4	30	760–900	NIR
5	30	1550–1750	MIR
6	60 ^a /120 ^b	10400–12500	TIR
7	30	2080–2350	MIR
Pan ^a	15	500–900	Grün, Rot, NIR

Quelle: NASA (1998)

^a ETM⁺

^b TM

Abbildung 2.2
Landsat 7 mit Sensor ETM⁺ schematisch
(© NASA (1999))



2.3 Das Sensorensystem QuickBird

Instrumente zur Fernerkundung und ihr Einsatz unterlagen in den vergangenen Jahrzehnten in erheblichem Maße Veränderungen (HILDEBRANDT, 1996). Wurden bis in die 90er Jahre des letzten Jahrhunderts Satelliten zur Erderkundung hauptsächlich von staatlichen Organisationen² eingesetzt und administriert, wandelte sich dies u. a. mit den erfolgreichen Starts der kommerziellen Satelliten IKONOS und QuickBird. Einen umfassenden Überblick über staatliche Aktivitäten zur Integration von kommerziellen FE-Produkten in staatliche Programme bieten BIRK ET AL. (2003) am Beispiel der USA. Wesentliche Anstrengungen auf diesem Gebiet beinhalten die Verifizierung und Validierung der durch nicht-staatliche Anbieter bereitgestellten Daten.

Das hochauflösende System QuickBird wurde im Oktober 2001 in einen 450 km hohen Orbit gebracht (DIGITALGLOBETM, a). Die Betreuung erfolgt durch die Firma DIGITALGLOBETM, welche QuickBird in Kooperation mit KODAK, BALL AEROSPACE und FOKKER SPACE entwickeln ließ. QuickBird ist der Markenname des eigentlich als „Ball Aerospace Global Imaging System 2000“ bezeichneten Satelliten. Er besteht im Wesentlichen aus einem Transportfahrzeug³ und dem daran

² hochauflösende Systeme vornehmlich unter militärischer Verwaltung

³ „Ball Aerospace Commercial Platform 2000“

Tabelle 2.2: Geometrische und spektrale Auflösung des Satelliten QuickBird

Kanal	Geometrische Auflösung ^a	Spektrale Auflösung
	in m	in nm
Blau	2,44–2,88	450–520
Grün		520–600
Rot		630–690
NIR		760–900
Pan	0,61–0,72	445–900

Quelle: DIGITALGLOBE™ (a)

^a 0 bzw. 25° vom Nadir

montierten optischen System⁴ (BALL AEROSPACE & TECHNOLOGIES CORP., a,b,c). Abbildung 2.3 auf der nächsten Seite vermittelt einen Eindruck des 3,04 m langen Raumfahrzeuges.

Die Kamera des QuickBird-Systems ist ähnlich den Instrumenten von Landsat ein multispektraler Scanner. Es stehen vier Farbbänder (Blau, Grün, Rot und NIR) sowie ein panchromatischer Kanal zur Verfügung (Tabelle 2.2). Gewählter Orbit und verwendete Sensor- und Speichertechnik gestatten die Aufnahme von Einzel- oder Streifenszenen der Größen $16,5 \times 16,5$ km bzw. $16,5 \times 165$ km. Der mögliche Aufnahmekorridor erstreckt sich ca. 272 km zu beiden Seiten der projizierten Flugbahn; dieselbe Stelle am Boden wird wieder nach 1–3 Tagen überflogen. Je nach Nutzerbedürfnissen sind QuickBird-Szenen in drei Verarbeitungsstufen erhältlich, die sich hinsichtlich vorgenommener Korrekturen und Anpassungen unterscheiden. Im einfachsten Fall erfolgt nur eine Korrektur der sensor- und systemspezifischen Verzerrungen („basic“). Eine Georeferenzierung und Überführung in eine Kartenprojektion sowie radiometrische Korrekturen sind in der „standard“-Stufe enthalten. Für die Variante „orthorectified“ erfolgt schließlich noch eine Entzerrung von topographiebedingten Abbildungsfehlern.

Eine extrem hohe Auflösung des panchromatischen Kanals und die Möglichkeit der Kombination desselben mit den Multispektralkanälen schaffen Auswertungspotenziale auf einer Ebene, die bisher digitalen Luftbildern vorbehalten war. Die Kosten für dieses optische System sind höher als bei Nutzung von Landsat-Daten: 16–45 \$ kostet ein Quadratkilometer je nach Korrekturmethode, gewünschten Kanälen und Auftragsdringlichkeit (EURIMAGE, 2005); zudem ist QuickBird ein relativ junges System. Dies mögen Gründe dafür sein, dass es bis jetzt relativ wenige dokumentierte wissenschaftliche Anwendungsbeispiele (siehe Abschnitt 4.1.2 ab Seite 24) gibt.

2.4 Datenverfügbarkeit

Ein wesentlicher Gesichtspunkt bei der Entwicklung eines Inventurverfahrens mit Fernerkundungsphase ist die Datenverfügbarkeit der in Frage kommenden FE-Systeme entsprechend den Anwendungsbedürfnissen. Besteht die Möglichkeit zur Verwendung eines sensorunspezifischen Verfahrens, ist die Anfälligkeit gegenüber Systemausfällen seitens der Satellitensensoren natürlich geringer.

⁴ „Ball Aerospace High-resolution Camera 60“, in Zusammenarbeit mit KODAK

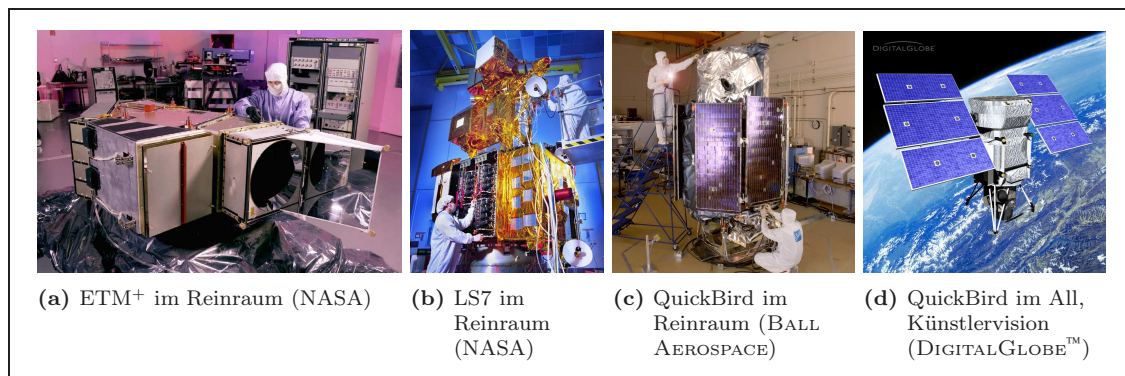


Abbildung 2.3: Die Systeme Landsat und QuickBird (© NASA (1999) bzw. DIGITALGLOBE™ (b))

Dies war für RICHARDS ET AL. (2000) eine der Prämissen bei der Entwicklung ihres Inventurverfahrens (siehe Seite 22).

Die im WEI-BWI-Projekt entwickelten Algorithmen zur Klassifikation und Inventarisierung forstlicher Zielgrößen gehen zwar nicht ausschließlich von *einem* System aus; bestimmte Eigenschaften müssten Alternativsensoren dennoch aufweisen. Dabei sind eine multispektrale Erfassung und eine vergleichbare Auflösung bei ähnlicher Flächenabdeckung an erster Stelle zu nennen. Zur Optimierung eines Stichprobenverfahrens sind überdies die anfallenden Kosten ein entscheidender Parameter (COCHRAN, 1977; KÖHL, 1994)—die Preise pro Flächeneinheit sollten sich also ebenfalls nicht signifikant unterscheiden. Zu den beispielhaft untersuchten Systemen Landsat und QuickBird existieren nur eingeschränkt empfehlenswerte Alternativen.

2.4.1 Landsat

„Bei der Durchführung von sehr großflächigen Kartierungen über mehrere Bildszenen hinweg ist die Datenverfügbarkeit allerdings weniger günstig. So ergaben Untersuchungen im Datenarchiv des Deutschen Zentrums für Luft- und Raumfahrt (DLR)⁵, dass die Datensätze von zwei bis drei Jahren benötigt werden, um flächendeckend für die BRD auswertbare Daten zu erhalten [...]. Dies bedeutet eine starke Einschränkung bei allen überregionalen Fernerkundungsanwendungen vor allem dann, wenn bei der Kartierung die Aktualität der Auswertung eine entscheidende Rolle spielt.“

SCHARDT (1990)

Dieses Zitat verdeutlicht die Herausforderung bei dem Vorhaben, Deutschland umfassend mit Satellitenszenen des Typs Landsat abdecken zu wollen. Deutschland umfasst laut den im Projekt verwendeten GIS-Daten eine Fläche von etwa 358.000 km², zu deren vollständigen Abdeckung insgesamt 32 Landsat-Vollszenen benötigt werden. Inzwischen ist die Datenbasis gewachsen, Daten des ETM⁺-Sensors stehen zur Verfügung. Trotzdem mussten bei der Kartierung der Bodenbedeckungsklassen im Rahmen des deutschen Teils des europäischen Projektes „CORINE Land Cover 2000“ sowohl Szenen von 1999 (9 Szenen) als auch 2001 (4 Szenen) herangezogen werden, da von den 32 notwendigen nur 19 aus dem Jahr 2000 verwendbar waren (KEIL ET AL., 2003).

⁵ durchgeführt von WINTER (1987)

Im Hinblick auf internationale Berichterstattungspflichten stellt die Verwendung von Daten aus drei unterschiedlichen Jahren sicher kein unüberwindliches Problem dar; die Kartierung großflächiger Schadereignisse wie bspw. Stürme wird hingegen erheblich erschwert, wenn von den betroffenen Gebieten keine aktuellen FE-Daten unmittelbar nach der Störung verfügbar sind.

Bei der Konzipierung von Inventurverfahren, die Stichproben verwenden, erübrigt sich der Zwang zur vollständigen Abdeckung der Landesfläche, dennoch kann nicht ausgeschlossen werden, dass bspw. bei einem systematischen Stichprobenmuster durch die wiederholte Überdeckung derselben Stelle durch Wolken oder andere atmosphärische Störungen eine Szenenverwendung unmöglich wird.

Landsat 7 erfüllte das ursprüngliche Missionsziel mit der Vollendung des fünften Einsatzjahres bereits im April 2004 (UNITED STATES GEOLOGICAL SURVEY, 2004). Die Zukunft von „Landsat 8“ ist ungewiss, da das Landsat-Programm nicht der erhoffte kommerzielle Erfolg wurde, den sich die USA versprochen hatten. Für Wissenschaftler ist es ungeachtet dessen eines der wichtigsten Standbeine der FE überhaupt, da beim Design der Landsat-Instrumente zahlreichen wissenschaftlichen Ansprüchen Rechnung getragen wurde (MARBURGER, 2004). Am 31. Mai 2003 fiel das Modul zur Kompensation der Vorwärtsbewegung des Satelliten (der Scanline Corrector (SLC)) während der Abtastung aus und konnte nicht wieder in Betrieb genommen werden. Damit verringerte sich die aufgenommene Fläche einer Szene um 25 %; störende Artefakte werden seitdem mit Hilfe älterer Szenen per Histogrammanpassung soweit als möglich beseitigt (UNITED STATES GEOLOGICAL SURVEY, 2003). Landsat 5 ist zwar ebenfalls noch einsatzfähig (nach mehr als 20 Jahren Betriebszeit!), liefert aber ebenso keine einwandfreien Daten mehr.

Diese Probleme unterstreichen die Notwendigkeit einer Verfahrensauslegung für unterschiedliche Sensorentypen, wie einführend bereits dargelegt wurde. Die technisch am ehesten verwandten Systeme SPOT⁶ (Frankreich) und IRS⁷ (Indien) offerieren eine etwas höhere Auflösung zu wesentlich höheren Preisen (SPOT, 2005; EUROMAP, 2005), was ihre Kompatibilität mit den im Projekt vorgeschlagenen Verfahren einschränkt. Abschließend muss erwähnt werden, dass die WEI-BWI-Ansätze auf eine etwaige Verbesserung der künftigen BWI III abzielen—die vielleicht frühestens in 10–12 Jahren stattfindet. Allerdings ist es schwierig, heute für diesen fernen Zeitpunkt die dann verfügbare FE-Technik abzuschätzen.

2.4.2 QuickBird

Bedenkt man, dass QuickBird ein kommerzielles System ist, verwundert nicht, dass die Dokumentation von Systemfehlern weniger transparent erfolgt, als es bei Landsat der Fall ist. Eine Einschätzung der Zuverlässigkeit der Datenversorgung fällt damit schwerer. QuickBird ist laut Herstellerangaben für eine 7-jährige Betriebszeit mit Treibstoff befüllt (DIGITALGLOBE™, a). Erfüllt der Satellit die Herstellererwartungen, steht er also bis mindestens 2008 zur Verfügung. Die Firma DIGITALGLOBE™ hat nun laut ihrem Internetauftritt den Anspruch, mit ihrem System ein sehr agiles und reaktionsschnelles Instrument zur Verfügung zu stellen. Inwieweit ein dadurch bedingter häufiger Wechsel in Ausrichtung und Lage den Treibstoffverbrauch bzw. die Lebensdauer beeinflusst, ist für den Datennutzer ungewiss. Zwar hat sich die Joint Agency Commercial Imagery Evaluation (JACIE) zum Ziel gesetzt, kommerzielle Satelliten dahingehend zu beurteilen, ob sie

⁶ Systeme Probatoire d'Observation de la Terre

⁷ Indian Remote Sensing Satellite

angegebene Leistungsdaten einhalten und eine so genannte „temporale“ Beständigkeit haben, für QuickBird hat diese Prüfung jedoch noch nicht stattgefunden. Das bereits im September 1999 gestartete IKONOS-System wurde bereits einer eingehenden Prüfung unterzogen (ZANONI ET AL., 2003).

Die Erfahrungen des Instituts für Waldwachstum und Forstliche Informatik der Technischen Universität Dresden mit dem QuickBird-Vertriebspartner EURIMAGE werfen ein weniger ermutigendes Licht auf die garantierte Lieferbarkeit von Daten. Im Projekt „Kombination von terrestrischen Aufnahmen und Fernerkundungsdaten-Daten mit Hilfe der k NN-Methode zur Kartierung von Waldökosystemen“ (BMBF, DLR) erwies sich die Akquisition einer Aufnahme des Tharandter Waldes innerhalb des gesetzten Zeitfensters als gänzlich unmöglich, obwohl in diesem Zeitraum stabile, wolkenfreie Wetterlagen vorherrschten, während die Datenbeschaffung im WEI-BWI-Projekt ebenfalls mit erheblichen Verzögerungen verbunden und schlussendlich nur ein Teil des spezifizierten Aufnahmegebietes als wolkenfreie Szene lieferbar war.

QuickBird steht mit seinem räumlichen Auflösungsvermögen gegenwärtig an der Spitze nicht-militärischer Satelliten. Geringfügig gröbere Auflösungen liefern die Instrumente von SPOT 5 und IKONOS.

2.4.3 Visueller Vergleich beider Systeme

Der vorhergehende Abschnitt beschäftigte sich mit der Datenverfügbarkeit beider Satellitensysteme. Zum Abschluss dieses Kapitels erfolgt an dieser Stelle noch eine kurze Gegenüberstellung beider FE-Produkte, um zu verdeutlichen, welchen Weg die Sensoren-Technik bis heute zurückgelegt hat. Abbildung 2.4 zeigt einen kleinen Ausschnitt aus dem Testgebiet „Pfälzerwald“.

Alle Datensätze wurden in Neuberechneter Form („resampled“) geliefert; aus diesem Grund entsprechen die Beispielabbildungen nicht den nativen Auflösungen der Sensoren. Es wird viermal derselbe Bildausschnitt (allerdings zu unterschiedlichen Zeitpunkten) gezeigt; der Versatz in der Position des Flugzeuges im QuickBird-Bild lässt sich mit der asynchronen Aufzeichnung von panchromatischen und multispektralen Kanälen und der Eigenbewegung des Satelliten erklären. Die multispektrale Darstellung erfolgt in Echtfarben, d. h., die Kanäle eins, zwei und drei wurden den Farben blau, grün bzw. rot zugeordnet.

QuickBird offeriert verglichen mit Landsat 7 naturgemäß eine enorme Fülle von Details. Dabei ist allerdings zu beachten, dass eine hohe räumliche Auflösung keine Universallösung für jedwe-

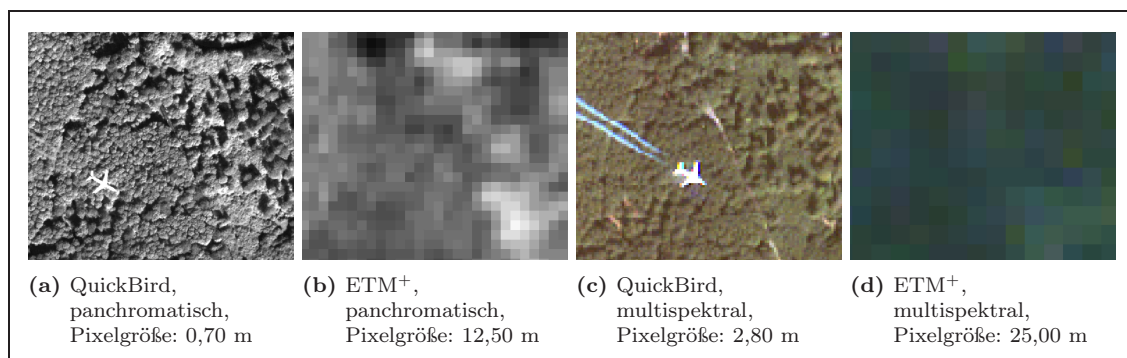


Abbildung 2.4: Vergleich der räumlichen Auflösung: ETM⁺ vs. QuickBird (© NASA bzw. DIGITALGLOBE™)

2 Satellitengestützte Fernerkundungssysteme

de Fragestellung darstellt. Auch heute haben Sensoren mit wesentlich größerer Auflösung ihre Daseinsberechtigung, da sie ungleich größere Gebiete pro Überflug kartieren können und einen mehrere Größenordnungen kleineren Preis haben (min. 40 km² pro investiertem Dollar bei Landsat versus 0,02 km²/\$ bei QuickBird im teuersten Fall).

Kapitel 3

Die logistische Regression

3.1 Einführung

Zur Analyse von stochastischen Zusammenhängen in Datensätzen existiert eine Vielzahl von Methoden (DRAPER und SMITH, 1998; ROTHMAN, 1998; LLOYD, 1999; SACHS, 2002). Zu den bekanntesten zählen Regressionsberechnungen, die eine Aussage zur *Art* dieser Zusammenhänge treffen—sie liefern eine funktionale Beschreibung in Form einer mathematischen Gleichung.

Ein klassisches Instrument ist die *lineare Regression*. Sie fußt auf der Methode der „kleinsten Abweichungsquadrate“, was bedeutet, dass die resultierende Gleichung die Summe der quadrierten Abweichungen zwischen vorhergesagten und beobachteten Werten minimiert. Dieser Regressions-typ wird häufig angewendet, u. a. deshalb, weil die Ergebnisse bzw. das erstellte Regressionsmodell einfach und intuitiv interpretierbar sind. Die sinnvolle Anwendung der linearen Regression setzt spezielle Bedingungen voraus, die auszugsweise Erwähnung finden sollen (HARTUNG und ELPELT, 1999; SACHS, 2002):

1. *Linearität* des Zusammenhangs; d. h., es wird von einer additiven Wirkung der Regressoren ausgegangen,
2. *Stetigkeit* des Regressanden; d. h., absolut- oder interval-skalierte kontinuierliche abhängige Variable,
3. *Unabhängigkeit* des Regressanden; d. h., die Beobachtungen der abhängigen (Zufalls-) Variable müssen voneinander unabhängig sein,
4. *Normalverteilung* des Regressanden; d. h., die Beobachtungen der abhängigen Variable müssen normalverteilt sein,
5. *Homoskedastizität*; d. h., die Varianz der Residuen sollten immer etwa gleich sein und den Mittelwert $\mu = 0$ haben,
6. *Beobachtungsfehlerfreiheit* der Regressoren; d. h., die erklärenden Variablen der Gleichung sind keine Zufallsvariablen und weisen keine Messfehler auf.

Eine geringfügige Verletzung dieser Regeln kann trotzdem zu vernünftigen Resultaten führen; bei größeren Abweichungen ist der Aussagewert der Ergebnisse allerdings ernsthaft in Frage gestellt.

Insbesondere nominal- oder ordinal-skalierte Daten entziehen sich der Analyse durch eine lineare Regression.

Mit der *logistischen Regression* existiert eine Alternative, die vor allem in der epidemiologischen Forschung bereits seit langem zur Anwendung kommt (HOSMER und LEMESHOW, 1989; KLEINBAUM und KLEIN, 2002). In diesem Forschungsbereich ergeben sich oft Fragestellungen der Form: „Wie wirkt sich die Ausprägung einer/mehrerer unabhängigen Variablen auf das Vorhandensein eines Merkmals (der abhängigen Variablen) aus?“ Die Besonderheit gegenüber der multiplen, linearen Regression besteht darin, dass dieses abhängige Merkmal¹ einerseits sehr häufig ordinal- oder nominal-skaliert ist und andererseits oftmals nur zwei Ausprägungen aufweist—z. B. *tot* vs. *lebend*, *kein Krebs diagnostiziert* vs. *Krebs diagnostiziert* oder, auf das Projekt übertragen: *Nichtwaldpixel* vs. *Waldpixel*. Man spricht von einer *binären Ausprägung*. Für den Fall einer abhängigen Variablen mit mehr als zwei Ausprägungsmöglichkeiten, bspw. *Nadel-*, *Misch-*, *Laubwald*, existiert eine Weiterentwicklung der binären logistischen Regression—die *multinomiale* logistische Regression. Beide Varianten verwenden, anders als die lineare Regression, als Schätzmethode das Verfahren der maximalen Mutmaßlichkeit mit iterativer Schätzung, im englischen Sprachraum besser bekannt als Maximum-Likelihood-Verfahren (ML). Erläuterungen dazu folgen im Abschnitt 3.2.2 auf Seite 18.

Die weite Verbreitung der logistischen Regression ist auf einige interessante Charakteristika dieses Konzepts zurückzuführen. So sind die Mindestanforderungen für eine Anwendung deutlich weniger umfangreich. Es wird keine Normalverteilung der Daten erwartet, die Variablen können nominal, ordinal² oder metrisch skaliert sein. Weiterhin bewirkt eine unterschiedliche Stichprobenrate innerhalb der Teilpopulationen mit bzw. ohne Ausprägung des Zielmerkmals keine negative Beeinflussung der Parameterschätzung (ausgenommen des konstanten Terms). Dadurch besteht die Möglichkeit, insbesondere seltene Merkmale durch eine höhere Stichprobenrate zuverlässiger zu erheben (HOSMER und LEMESHOW, 1989; BALTES-GÖTZ, 2004). Die verwendete mathematische Funktion bewirkt einen sigmoiden Kurvenverlauf und eine Beschränkung des Erwartungswertes der abhängigen Variablen auf $0 \leq E(Y = 1|X) \leq 1$, da eine Wahrscheinlichkeit modelliert wird. Einzelheiten dazu werden im folgenden Abschnitt zur zugrunde liegenden mathematischen Theorie der logistischen Regression erläutert. Die Kurvenform ist vom biologischen Standpunkt gesehen plausibel: Unterhalb eines bestimmten Schwellenwertes bewirkt eine Steigerung der Einflussgrößen wenig, bei Überschreitung dieses Schwellenwertes reagiert die abhängige Größe deutlich, um dann bei weiterer Steigerung nur noch vernachlässigbar anzusteigen (KLEINBAUM und KLEIN, 2002).³ Abbildung 3.1 auf der nächsten Seite illustriert die genannten Charakteristika an einem konkreten Beispiel. Zu beachten ist die Darstellung eines „negativen“ Wachstums in dem Sinne, dass mit steigenden Grauwerten die Wahrscheinlichkeit, dass ein solches Pixel Wald repräsentiert, sinkt. Die Ursache dafür ist die Eigenschaft des Waldes, ein dunkler Strahler zu sein, dessen Albedo meistens deutlich unter der anderer Landbedeckungsformen liegt (von Gewässern abgesehen). Die Trennung von *Wald* und *Nichtwald* fällt zwischen den Grauwerte 27 und 28, wenn man von einem symmetrischen Klassifizierungsschwellenwert $S = 0,5$ ausgeht.

Im Gegensatz zur linearen Regression erfolgt die Interpretation der gewonnenen Ergebnisse völlig anders, da kein direkter Einfluss einer Größe auf den Anstieg einer Regressionsgeraden

¹ im englischsprachigen Raum oft als *disease variable* bezeichnet

² Einige bekannte Statistiksoftware-Pakete nutzen allerdings nicht den vollen Informationsgehalt von ordinal-skalierten Daten bzw. verwenden dann spezielle Routinen.

³ weshalb einigen Wachstumsfunktionen eine logistische Funktion zugrunde gelegt ist

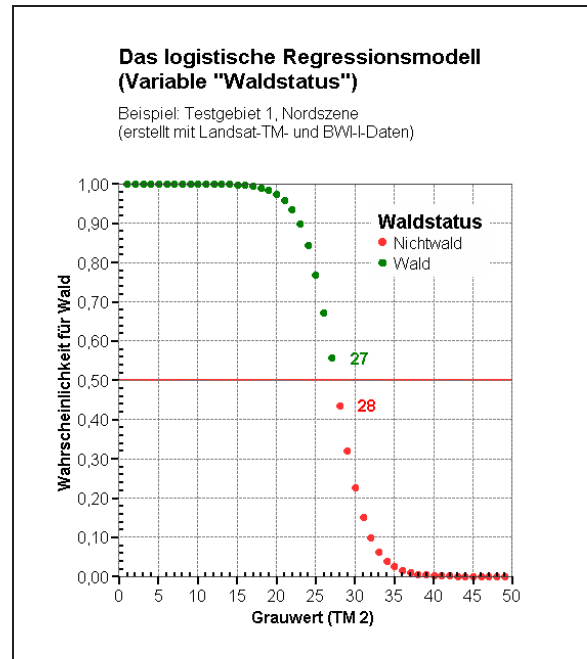


Abbildung 3.1
Logistisches Modell am Beispiel des Merkmals *Waldstatus*

ablesbar ist.⁴ Die Gründe dafür werden ebenfalls im folgenden Abschnitt dargelegt.

3.2 Theoretische Grundlagen

3.2.1 Aufbau von Generalized Linear Models

Die folgenden Aussagen stellen eine Zusammenfassung der entsprechenden Abschnitte der Arbeiten von HOSMER und LEMESHOW (1989), McCULLAGH und NELDER (1989), AGRESTI (1990), DOBSON (2001), LLOYD (1999) und KLEINBAUM und KLEIN (2002) sowie einiger der Statistik gewidmeten Internet-Seiten dar.

Die logistische Regression (wie die lineare Regression auch) lässt sich nach NELDER und WEDDERBURN (1972) einer ganzen Familie von Modellen mit gleichen Grundlagen zuordnen: den verallgemeinerten linearen Modellen, bekannter als Generalized Linear Models (GLM). Typisch für diese Familie ist der Modellaufbau aus drei Elementen:

- zufällige Komponente,
- systematische Komponente und
- Verbindung zwischen diesen beiden Bestandteilen.

Die *Zufallskomponente* besteht in der abhängigen Variablen mit einer Wahrscheinlichkeitsverteilung, die der Klasse der exponentiellen Verteilungen angehören muss. Der systematische Teil des GLM wird von einer *linearen* Verknüpfung der erklärenden Variablen gebildet, die als Prädiktor für die abhängige Variable fungiert. Eine so genannte *Link-Funktion* stellt das Bindeglied zwischen Erwartungswert der Zielgröße und linearem Prädiktor dar. Sie muss monoton⁵ und invertierbar sein.

⁴ Diese Eigenschaft wird häufig als ein Nachteil der logistischen Regression angesehen.

⁵ damit differenzierbar

Die Art dieser Verbindungsfunktion hängt meist eng mit der Wahrscheinlichkeitsverteilung der Zielgröße zusammen, nichts desto trotz besteht Flexibilität in der Wahl des Links. Diese Spezifikation eines GLM gestattet neben der Normalverteilung auch die Verwendung anderer exponentieller Verteilungen, z. B. Poisson-, Gamma- und Binomialverteilungen (LANE, 2002), weshalb für GLMs nicht die Forderung nach Normalität der abhängigen Variablen besteht.

Ein Beispiel soll die beschriebenen Eigenschaften darlegen. Zielgröße und damit zufällige Komponente des GLM sei der Waldstatus eines Landsat-Bildelements in Abhängigkeit von den mit dem Satellitensensor erfassten Grauwerten (GWs). Es ist somit die Wahrscheinlichkeit gesucht, dass ein Pixel mit gegebenen GWs in den jeweiligen Kanälen ein Waldpixel ist:

$$P(Y = 1|\text{GW}) = \pi(x) = \frac{e^{\beta\mathbf{X}}}{1 + e^{\beta\mathbf{X}}}. \quad (3.1)$$

Formel 3.1 zeigt die *Grundform* des verwendeten *logistischen Regressionsmodells*. $P(Y = 1|\text{GW})$ ist die bedingte Wahrscheinlichkeit für das Auftreten der Merkmalsausprägung „Wald“ bei einem gegebenen GW. Da *Waldstatus* mit 0 und 1 kodiert ist, kann diese bedingte Wahrscheinlichkeit gleich dem Erwartungswert $\pi(x)$ ⁶ gesetzt werden (HOSMER und LEMESHOW, 1989). $\beta\mathbf{X}$ ist die lineare Kombination aller zur Vorhersage herangezogenen Parameter und Koeffizienten; Erläuterungen dazu folgen weiter unten.

Da ein Pixel entweder als Wald oder als Nichtwald klassifiziert werden kann, existieren nur zwei mögliche Ausprägungen der Zielvariable—sie ist binomialverteilt. Eine normale lineare Regression würde an dieser Stelle ungeeignet sein, unter anderem schon deshalb, weil sich die Erfolgswahrscheinlichkeit des Ereignisses $Y = 1 = \text{Wald}$ definitionsgemäß nur zwischen 0 und 1 bewegen kann. Die Einhaltung dieser Bedingung kann ein einfaches Modell der Form

$$y = f(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p \quad (3.2)$$

nicht sicherstellen. Sowohl die Binomialverteilung als auch der in mancher Literatur als *Bernoulliverteilung* bezeichnete Spezialfall der Binomialverteilung⁷ sind Vertreter der exponentiellen Verteilungen. Selbige haben eine allgemeine Wahrscheinlichkeitsdichtefunktion der Form

$$f(y|\theta, \phi) = \exp \left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right]. \quad (3.3)$$

Es bedeuten

θ natürlicher oder *kanonischer* Parameter

ϕ Dispersions- oder Streuungsparameter⁸

$a(\cdot)$, $b(\cdot)$, $c(\cdot)$ verteilungsspezifische Funktionen.

Wird die allgemeine Gleichung 3.3 durch den konkreten Fall der Bernoulliverteilung ersetzt, erhält man:

$$P(Y = y) = \pi^y(1 - \pi)^{1-y} = \begin{cases} \pi & \text{wenn } y = 1 = \text{Wald} \\ 1 - \pi & \text{wenn } y = 0 = \text{kein Wald} \end{cases} \quad (3.4)$$

⁶ wird später zu π verkürzt verwendet

⁷ wenn die Anzahl „Ziehungen“ = $n = 1$

⁸ nur für einige Verteilungen relevant

mit

- y Waldstatus,
- π Erwartungswert der Zielvariablen.

Bringt man Gleichung 3.4 in die kanonische Form von Gleichung 3.3, ergibt sich

$$f(y) = \exp \left[\frac{y \ln \left(\frac{\pi}{1-\pi} \right) + \ln(1-\pi)}{1} \right]. \quad (3.5)$$

Es wird ersichtlich, dass der kanonische Parameter θ gleich $\ln \left(\frac{\pi}{1-\pi} \right)$ und der Dispersionsparameter ϕ gleich 1 ist.

Wie angedeutet, sollen Grauwerte eines oder mehrerer Kanäle der zu klassifizierenden Szene als erklärende Variablen fungieren. Der systematische Teil des GLM hat also, da eine lineare Kombination gefordert ist, folgende Grundgestalt (in Matrixschreibweise):

$$\eta = \boldsymbol{\beta} \mathbf{X} \quad (3.6)$$

mit

- η linearer Prädiktor,
- $\boldsymbol{\beta}$ $(p+1) \times 1$ -Vektor der zu schätzenden Parameter,⁹
- \mathbf{X} $n \times p$ -Designmatrix mit Grauwerten von p Kanälen in n Beobachtungen.

Die Link-Funktion $g(\cdot)$ als dritter Bestandteil des GLM soll die Verbindung zwischen den beiden bereits erläuterten Komponenten bewirken. Ihre Grundform sieht demnach wie folgt aus:

$$g(\pi) = \eta = \boldsymbol{\beta} \mathbf{X}. \quad (3.7)$$

Im konkreten Beispiel kommt die *Logit*-Funktion zum Einsatz: π wird also einer Logit-Transformation unterzogen.

$$g(\pi) = \ln \left(\frac{\pi}{1-\pi} \right) = \eta = \boldsymbol{\beta} \mathbf{X} \quad (3.8)$$

Bildet man aus $g(\cdot)$ die zugehörige Umkehrfunktion $g^{-1}(\cdot)$, erhält man wieder Gleichung 3.1:

$$\pi(x) = \frac{e^{\boldsymbol{\beta} \mathbf{X}}}{1 + e^{\boldsymbol{\beta} \mathbf{X}}}.$$

Die Logitfunktion¹⁰ wird im Falle der Binomialverteilung *kanonischer Link* genannt, da gilt: $\theta = \eta$. Der kanonische Parameter ist also gleich dem linearen Prädiktor. Durch die Logit-Transformation wird sichergestellt, dass der Prädiktor Werte zwischen $-\infty$ und $+\infty$ annehmen kann, ohne dass dies eine Auswirkung auf die Forderung nach einem Wertebereich zwischen 0 und 1 im Falle der abhängigen Variablen hat. Man erreicht durch eben diese Transformation, dass der Prädiktor eine lineare Gestalt erhält. Eine Konsequenz dieses Verfahrens ist, dass die Fehler keiner Normal-

⁹ Der Vektor $\boldsymbol{\beta}$ besteht ja nicht nur aus den unbekanntem Parameter $\beta_1 \dots \beta_p$, sondern auch der Konstante β_0 .

¹⁰ daher der Name *logistische Regression*

sondern einer Binomialverteilung mit Mittelwert 0 folgen. Ein weiterer Vorteil der Verwendung eines kanonischen Links sind vereinfachte Parameterschätzungen.

3.2.2 Schätzverfahren für Generalized Linear Models

Wie bereits angedeutet, kommt bei logistischen Regressionen das Maximum-Likelihood-Verfahren zur Anwendung, um die Koeffizientenschätzer zu ermitteln. Auf dieses soll nachfolgend kurz eingegangen werden, um den Einblick in die Theorie der GLM bzw. speziell der logistischen Regression abzurunden.

Nach HOSMER und LEMESHOW (1989) bestehen auch bei der Modellanpassung Parallelen zwischen der linearen und der logistischen Regression. Auf den ersten Blick kommt bei der linearen Regression natürlich die Methode der geringsten Abweichungsquadrate zum Einsatz. Geht man aber von einer Normalverteilung der Fehler aus, so ist dieses Verfahren praktisch ein Spezialfall des ML-Verfahrens, das auch für GLMs angewendet wird.

Vereinfacht gesagt wird bei der Modellanpassung jene Parameter-Kombination gesucht, die das Auftreten der *beobachteten* Daten als Gesamtheit am wahrscheinlichsten macht—daher der Name. Dazu werden ML-Gleichungen verwendet. Sie sind bei der linearen Regression wie das Regressionsmodell selbst linearer Natur und können daher ohne Schwierigkeiten über ihre ersten Ableitungen maximiert werden. Dies ist bei logistischen Modellen nicht möglich, da aus den nichtlinearen Mutmaßlichkeitsgleichungen Ableitungen resultieren, deren Verwendung nur einen iterativen Prozess der Maximierung zulässt. Für den Anwender statistischer Softwarepakete stellt dies indes kein Problem dar, da die Programmroutinen die Anpassung vornehmen.

Um den Vorgang der Modellanpassung zu erläutern, ist es praktisch, sich nochmals Gleichung 3.4 in Erinnerung zu bringen:

$$P(Y = y) = \pi^y(1 - \pi)^{1-y} = \begin{cases} \pi & \text{wenn } y = 1 = \text{Wald} \\ 1 - \pi & \text{wenn } y = 0 = \text{kein Wald} \end{cases}$$

Um Missverständnissen vorzubeugen, wird sie an dieser Stelle etwas erweitert für den univariaten Modellfall dargestellt, wie er zur Erstellung der Waldkarten zum Tragen kam:

$$\zeta(x_i) = P(Y = y) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}. \quad (3.9)$$

Mit ihrer Hilfe lässt sich der Beitrag eines jeden Wertepaares (x_i, y_i) mit $(i = 1 \dots n)$ zur ML-Funktion darstellen. Gemäß den Regeln der Wahrscheinlichkeitsrechnung werden statistisch unabhängige Ereignisse multipliziert, um zu einer Gesamtwahrscheinlichkeit (joint probability) zu gelangen (SACHS, 2002). Daraus resultiert:

$$l(\beta) = \prod_{i=1}^N \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}. \quad (3.10)$$

$l(\beta)$ bringt zum Ausdruck, dass der Vektor der Parameter $\beta_0 \dots \beta_p$ gesucht ist, welcher die Funktion maximiert. Zwar erscheinen die β_p in Gleichung 3.10 nicht, $\pi(x_i)$ wird jedoch mit ihnen berechnet (siehe Gleichung 3.1). Gleichung 3.10 lässt sich hinsichtlich des notwendigen Berechnungs-

aufwandes vereinfachen, wenn statt ihrer der Logarithmus naturalis der Mutmaßlichkeitsfunktion verwendet wird:

$$\ln[l(\boldsymbol{\beta})] = \sum_{i=1}^N y_i [\pi(x_i)] + (1 - y_i) [1 - \pi(x_i)]. \quad (3.11)$$

Um das Maximum für $\ln[l(\boldsymbol{\beta})]$ zu ermitteln, wird Gleichung 3.11 jeweils nach β_0 und β_1 abgeleitet und die resultierenden Gleichungen gleich Null gesetzt. Im multivariaten Fall ergeben sich nicht nur zwei, sondern $p + 1$ solcher Ableitungen.

Kapitel 4

Literaturüberblick

4.1 Anwendungsbeispiele für fernerkundungsgestützte Waldkartierung

4.1.1 Landsat

Die im Folgenden aufgeführten Beispiele für die Nutzung von FE-Daten auf Satellitenbasis sollen keineswegs eine erschöpfende Aufzählung darstellen. Sie sollen lediglich illustrieren, in welchen verschiedenen Gebieten FE-Daten dieser Art zum Einsatz kommen.

SCHARDT (1990) setzte sowohl Radar- als auch Landsat-5-Daten ein, um zu untersuchen, inwieweit diese sich für weit reichende Klassifizierungen im Forstbereich einsetzen ließen. Dazu wählte er drei Testgebiete in Deutschland, die aufgrund ihrer Besonderheiten jeweils unterschiedliche Themenschwerpunkte in den Untersuchungen repräsentierten. SCHARDT (1990) konnte zeigen, dass mit einem multitemporalen Ansatz (April- und Juli-Szenen eines bzw. direkt aufeinander folgender Jahre) und Einbindung eines digitalen Höhenmodells (DHM, international gebräuchlicher: Digital Elevation Model (DEM)) deutliche Verbesserungen sowohl bei visueller als auch automatischer Interpretation erzielbar waren. Auf spezielle Aspekte der Ergebnisse dieser Arbeit wird im Folgenden noch eingegangen.

Der Einsatz von Landsat-TM-Daten zur Modellierung von Schadstoffflüssen stand im Mittelpunkt der Forschung von SMIA TEK (1993). Ziel war die Gewinnung flächenbezogener Ergebnisse, die bekanntermaßen Probleme bereitet, wenn nur von punktförmigen Stichprobenflächen ausgegangen wird. Nach SMIA TEK ergibt sich eine Reihe von Besonderheiten aus der Arbeit mit FE-Daten. Unter anderem ist eine Klassifizierung als „Messverfahren“ für bestimmte Zielgrößen stets fehlerbehaftet¹—ein Umstand, dem man bei der Verwendung von am Boden gemessenen Daten aus dem Weg zu gehen glaubt.² SMIA TEK (1993) weist darauf hin, dass Klassifizierungen auf Basis des Maximum-Likelihood-Verfahrens oftmals verzerrt sind, da die daran geknüpfte Bedingung der Normalverteilung der Merkmalswerte nicht erfüllt ist. Am deutlichsten wirkt sich dieses Problem bei in Realität bi- oder multimodal verteilten Daten aus. Als Auswertemethode zur Landnutzungsklas-

¹ Ursachen können sein: Mischpixel, ungenaue Klassendefinitionen, ungerechtfertigte Approximation der Klassenbeschreibung mit einer Normalverteilung.

² obwohl ja auch bei klassischen Messverfahren wie der Kluppung, Entfernungs- und Höhenmessung Fehler nicht auszuschließen sind

sifizierung für die Landsat-Szene (von Baden-Württemberg) kam eine dreifache Stichprobe zum Einsatz. In einem ersten Schritt wurde eine Stichprobe zur Datenreduktion gezogen, eine weitere folgte als zweiter Schritt zum Zwecke der Trainingsdatenerfassung. Eine abschließende Stichprobe wurde gezogen, um den Klassifizierungsfehler schätzen zu können. Stichprobeneinheit war das einzelne Pixel. Grundsätzlich erkennt SMIA TEK dem Verfahren Potenzial zu, gibt aber zu bedenken, dass die Pixelzuordnung zu Landnutzungsklassen geringer relativer Häufigkeit problematisch ist.

Die Beurteilung der Qualität fernerkundungsbasierter Klassifikationen bzw. thematischer Karten ist ein elementarer Bestandteil im Produktionsprozess solcher Datensätze. Die aus den Karten abgeleiteten Informationen dienen in vielen Fällen der politischen Planung und Entscheidungsfindung. Diese Entscheidungen wiederum sind oftmals mit tief greifenden Konsequenzen verbunden, was eine eingehende Beschäftigung mit der Genauigkeit und Anwendbarkeit der Karten erfordert. SMITS ET AL. (1999) stellen einen Ansatz vor, der über eine reine Fehler- oder Konfusionsmatrix hinausgeht und die Kosten möglicher Fehlklassifikationen ebenso wie die originären Ziele des *Datennutzers* einzubinden sucht.

Im Rahmen des TREES-II-Projektes³ des Joint Research Centre (JRC) beschrieben RICHARDS ET AL. (2000) ein Verfahren zur Schätzung der Entwaldungsrate in den feuchten Tropenwäldern der Erde zwischen 1992 und 1997. Dabei kamen nicht nur Landsat-TM-, sondern auch SPOT- und AVHRR-Daten zum Einsatz. Dabei wurde folgendermaßen vorgegangen: Als erstes erfolgte eine Prästratifizierung des Tropengürtels der Erde basierend auf dem Waldanteil (geschätzt mit Advanced Very High Resolution Radiometer (AVHRR)) und bekannter Abholzungsaktivität (nach Experteneinschätzung). Um die Auswahl der Stichprobenflächen unabhängig vom verwendeten Sensorensystem gestalten zu können, wurde mit einem sog. „Area Frame“ gearbeitet. Da es sich bei dem TREES-II-Projekt um einen Untersuchungsgegenstand mit globaler Ausdehnung handelte, musste nach neuen Wegen gesucht werden, um die Abbildungsfehler bestehender Kartenprojektionen und Koordinatensysteme zu überwinden. Als geeignet wurde eine Tessellation von Hexagonen auf der Erdkugel befunden. Landsat-Daten wurden herangezogen, um die Entwaldungsrate in bedeutsamen Abholzungsregionen exakter schätzen zu können, da die vom AVHRR-System gelieferte Auflösung von einem Kilometer zu grob für diese Anwendung war.

Zur Kartierung der Landbedeckungsklassen Großbritanniens wurden 1990 die „Land Cover Map of Great Britain“ und 2000 die „UK Land Cover Map“ erstellt, welche hauptsächlich auf Landsat-Daten basieren; für die jüngere der Karten wurden z. T. auch IRS-Daten verwendet. FULLER ET AL. (2003) diskutieren anhand beider die Anforderungen an die Klassifikationsgenauigkeit thematischer Karten bei der Bilanzierung von Flächenveränderungen. Es wird deutlich, dass sich bei üblicherweise erreichbaren Kartengenauigkeiten von ca. 85–90 % und 2 bis 3 zu unterscheidenden Klassen Veränderungen auf mindestens 20 % der Fläche vollzogen haben müssen, um zu 75 % verlässliche Aussagen zu diesen Veränderungen zwischen den betrachteten Zeitpunkten ableiten zu können. Die Ergebnisse von FULLER ET AL. haben weit reichende Konsequenzen, da nach ihnen für das hier bearbeitete Projekt keine sinnvolle Veränderungsbilanz erstellbar ist (dazu mehr im Kapitel 7 ab Seite 57).

SAGISCHEWSKI ET AL. (2003) bearbeiteten ein Projekt zur Erstellung des Waldverzeichnisses Thüringen. Landsat-TM-Daten wurden dazu mit höher aufgelösten panchromatischen Informationen des indischen Satelliten IRS-1C fusioniert und anschließend objektorientiert klassifiziert. Bei

³ Tropical Ecosystem Environment Observations by Satellite

der Kartierung der Waldfläche konnten mit diesem Verfahren 94,3 bzw. 97,3 % Erstellergenauigkeit für die Klassen *Nichtwald* bzw. *Wald* erzielt werden. Das Verfahren wurde anhand zweier Forstämter auf seine Operationalität geprüft.

In der Schweiz wurden Daten verschiedener optischer Sensoren und Radarbänder untersucht, um Möglichkeiten zur Windwurfkartierung aufzuzeigen. SCHWARZ ET AL. (2003) konnten zeigen, dass im April aufgenommene ETM⁺-Szenen eine zu 86 % richtige Klassifizierung der Klassen *Wald*, *andere Vegetation* und *Windwurf* innerhalb des Waldes ermöglichen. Sie gehen von noch höheren Genauigkeiten für den Fall von Aufnahmen aus der Vegetationszeit aus. Höher aufgelöste IKONOS-Daten gestatteten eine Gesamtgenauigkeit von 85–88 % bei pixelbasierter Herangehensweise; das Maximum von 93 % wurde mittels manueller Interpretation der IKONOS-Szenen erreicht.

Einen sehr umfassenden Artikel zur Landbedeckungsklassifizierung mit Fernerkundungsmethoden hat GALLEGO (2004) publiziert. Er weist darin auf mögliche Hindernisse und Schwierigkeiten bei der Bearbeitung des genannten Themas hin. Betrachtet werden u. a. Verzerrungen durch einfache Pixelzählmethoden zur Flächenschätzung, Mischpixel und deren Implikationen bei Landnutzungsstatistiken und verschiedene Klassifikationsmethoden. Darüber hinaus werden Hindernisse bei der Verwendung von Fernerkundungsdaten zur Erstellung offizieller Statistiken aufgeführt. Wie FULLER ET AL. (2003) (siehe oben) warnt GALLEGO explizit vor unkritisch erstellten Veränderungsbilanzen, die erhebliche Fehler durch Fehlklassifikationen zu einem oder beiden betrachteten Zeitpunkten aufweisen können.

GHEBREMICAEL ET AL. (2004) wandten sich einer anderen Fragestellung zu: Kann man Kronenraumattribute von Schwarzholzakazie (*Acacia mearnsii*) aus Landsat-ETM⁺-Aufnahmen bestimmen? GHEBREMICAEL ET AL. (2004) nutzten für ihre Forschungen zwei Viertelszenen von Landsat 7. Als Basis diente die These, dass die Belaubung bzw. Benadelung von Bäumen als Ort der Photosynthese das entscheidende Bindeglied zwischen Umwelt und Biomasseproduktion sei. Eine geeignete Größe zum Ausdruck der Belaubungssituation eines Baumes/Bestandes ist der Blattflächenindex (BFI).⁴ Da eine direkte Ermittlung des BFI im Satellitenbild unmöglich ist, untersuchten GHEBREMICAEL ET AL. einige Vegetationsindizes (VIs) hinsichtlich ihrer Verwendbarkeit zur Bestimmung des BFI. Der Normalized Difference Vegetation Index (NDVI) lieferte die besten Ergebnisse, vor allem in Hinsicht auf Robustheit/Unempfindlichkeit gegenüber diversen störenden Einflussfaktoren. Es konnte ein einfaches, lineares Modell gefunden werden, das mit einem Bestimmtheitsmaß von $R^2 = 0,622$ und hoher Signifikanz ($p < 0,01$) besser als alle anderen VIs geeignet war, um den BFI vorherzusagen. Das „imperfekte“ Ergebnis wurde mit zahlreichen nicht kontrollierbaren Umweltfaktoren während der Bildakquisition und heterogenen Beständen begründet.

In Finnland wurden von MÄKELÄ und PEKKARINEN (2004) Untersuchungen zur Schätzbarkeit von Bestandesvolumina unter Verwendung von Landsat-TM-Bildmaterial angestellt. Als statistischer Ansatz kam die *k*NN-Methode zum Einsatz (*k* nächste Nachbarn bzw. Nearest Neighbours). Mehrere Ansätze wurden erprobt, um das Problem von Mischpixeln zu verkleinern. Es konnte gezeigt werden, dass aufgrund der geringen Auflösung des Landsat-Systems im Verhältnis zu den kleinräumigen Bestandesstrukturen nur unbefriedigende Schätzungen der Bestandesvolumina möglich sind. Bei einer Anwendung des Verfahrens in Gebieten mit größeren, homogenen Strukturen

⁴ Auch als Leaf Area Index (LAI) bezeichnet. Nach SCHÜTT ET AL. (1992) Verhältnis der einseitigen, gesamten Blattfläche eines Bestandes zur überdeckten Bodenfläche.

werden allerdings deutlich bessere Ergebnisse erwartet.

Kontextbezogene Klassifizierungsalgorithmen für Landsat-TM-Bilder wurden von MAGNUSSEN ET AL. (2004) am Beispiel zweier kanadischer Untersuchungsgebiete analysiert. Bei der Klassifizierung in wenige generelle Klassen (wie im WEI-BWI-Projekt vorgenommen) offerieren diese Methoden keine nennenswerten Vorteile. MAGNUSSEN ET AL. zeigen aber, dass bei einer höheren Klassenanzahl und gegebener spektralen Trennbarkeit durchaus Verbesserungen der Genauigkeit im Bereich von 4–6 % je Klasse möglich sind, insbesondere wenn die Ausgangsgenauigkeit der thematischen Karte ohne Nutzung von Kontextinformationen zwischen 50–80 % liegt.

Neben SMITS ET AL. (1999) nimmt FOODY (2005) eine weitere kritische Betrachtung heute üblicher Qualitätsprüfungsmethoden vor. Die Verwendung „globaler“, d. h. für eine ganze Karte oder Klassifizierung gültiger, Genauigkeitsmaße—abgeleitet aus Fehlermatrizen—birgt demnach die Gefahr einer unzutreffenden Verallgemeinerung der ermittelten Genauigkeiten. Abhilfe können räumlich begrenzte Analysen schaffen, die jeweils nur Ausschnitte des betrachteten Gebietes einbeziehen und so differenzierte Aussagen zur Variabilität der thematischen Genauigkeit gestatten.

4.1.2 QuickBird und andere hoch auflösende Systeme

SAWAYA ET AL. (2003) prüften IKONOS- und QuickBird-Daten im Zusammenhang mit hydrologischen Fragestellungen. Beispielhaft wurden drei der wichtigsten Parameter bei der Gewässereinschätzung untersucht: Wasserreinheit, Anteil wasserundurchlässiger („versiegelter“) Flächen und Gewässervegetation. Besondere Bedeutung haben diese Parameter bei der Städteplanung. Die Wasserreinheit ließ sich mit IKONOS-Daten durch ein log-transformiertes lineares Regressionsmodell in hohem Maße vorhersagen; zudem wurde festgestellt, dass IKONOS-Szenen in idealer Weise mit ETM⁺-Daten kombiniert werden können, wenn terrestrisch gewonnene Referenzdaten nicht ausreichend zur Verfügung stehen. Die Kartierung versiegelter Flächen gelang ebenso mit hoher Genauigkeit durch eine polynomische Gleichung 2. Grades. Ein pixelbasiertes Verfahren zur Verarbeitung der hoch aufgelösten Daten wurde dabei als sehr gut geeignet erachtet—eine Meinung, die andere Autoren nicht teilen (siehe HURD ET AL. (2005), LALIBERTE ET AL. (2004) und PUISSANT ET AL. (2005)). Die Analysen zur Kartierung von Wasservegetation mit IKONOS und QuickBird waren zum Druckzeitpunkt der Veröffentlichung noch nicht abgeschlossen, zeigten aber erste viel versprechende Ergebnisse. Zusammenfassend wurde herausgestellt, dass die Nutzung hoch auflösender Sensoren für Untersuchungen auf lokaler Ebene entscheidende Vorteile gegenüber vergleichbaren „analogen“ Luftbildern aufweisen (digitales Format, multispektrale Auflösung, Multitemporalität etc.), jedoch ungeeignet für die Bearbeitung großflächiger Untersuchungsgegenstände sind.

CLARK ET AL. (2004) untersuchten die Möglichkeiten zur Dokumentation von Absterbeprozessen im Tropischen Regenwald Costa Ricas mit hoch aufgelösten Satellitendaten. Dafür wurden zwei Szenen unterschiedlicher Sensoren der „VHR“-Klasse⁵ herangezogen—je eine IKONOS- und eine QuickBird-Szene. Zwischen beiden Szenen lag ein Zeitraum von etwas weniger als zwei Jahren. Es gelang erfolgreich, anhand eines Vergleich beider Szenen den Prozentsatz abgestorbener Bäume mit einer hohen Genauigkeit vorherzusagen (2,8 % per Fernerkundung geschätzt vs. 2,3 % per terrestrischer Aufnahme ermittelt).

Eine objektbasierte Klassifizierung von Luftbildern und einer QuickBird-Szene nahmen LA-

⁵ Im Sprachgebrauch der FE hat sich der Begriff Very High Resolution (VHR) für Sensorensysteme eingebürgert, die eine Auflösung von ca. 5 m oder besser bieten.

LIBERTE ET AL. (2004) vor, um die Ausbreitung verschiedener Straucharten in ehemalige Grasflächen New Mexicos zu dokumentieren. Der mit FE-Daten (anfangs nur Luftbilder) abgedeckte Zeitraum reichte bis 1937 zurück und ließ einen deutlichen Rückgang von Grasflächen zu Gunsten von Straucharten erkennen. Die objektbasierte Herangehensweise zeigte deutliche Vorteile gegenüber einem pixelbasierten Ansatz; trotzdem wurde der Anteil Sträucher an der Gesamtfläche unterschätzt. Insgesamt wurde dem Verfahren jedoch eine gute Eignung bestätigt. Im Hinblick auf QuickBird sind die Ergebnisse insofern bemerkenswert, als dass nur die Informationen des panchromatischen Kanals (mit auf 86 cm verringerter Auflösung) Eingang in die Klassifizierung fanden, eine zusätzliche Nutzung der multispektralen Informationen bei voller Auflösung also vermutlich noch mehr Potenzial aufweist.

Für einen direkten Vergleich von IKONOS mit QuickBird zogen WANG ET AL. (2004) je eine Szene beider Systeme von einem Mangrovenwald an der panamaischen Karibikküste heran und führten drei unterschiedliche Tests durch. Im ersten Schritt erfolgte lediglich ein visueller Vergleich verbunden mit einer statistischen Auswertung der Bilddaten. Die Schritte zwei und drei beinhalteten Klassifizierungsansätze auf spektraler, spektraler + texturaler und schließlich objektorientierter Basis. Es zeigte sich, dass die höhere räumliche Auflösung von QuickBird zu keinen besseren Klassifizierungsergebnissen führte. Bemerkenswert ist außerdem die Erkenntnis, dass der objektorientierte Ansatz mit *eCognition 3* praktisch keine Verbesserungen gegenüber der rein spektralen Klassifizierung erbrachte. Durchschnittliche Gesamtgenauigkeiten von ca. 70–75 % ließen sich auch ohne *eCognition* erreichen.

In den USA wird zur Zeit⁶ von HURD ET AL. (2005) eine Studie durchgeführt, um die Marschlandschaften am Long Island Sound (New York/Connecticut) zu erforschen. Für dieses Projekt wurden vier Zielstellungen erarbeitet, die unter anderem die Kartierung und Quantifizierung von Marschflächenzuwachsen und -verlusten und die Charakterisierung der Pflanzengemeinschaften beinhalten. Für die reine Flächenkartierung wurde der ETM⁺-Sensor ins Auge gefasst, während die detaillierte Bestimmung von Pflanzenarten mit QuickBird-Daten erfolgt. FE-Daten mit einem Auflösungsvermögen von z. T. < 1 m lassen sich nach heutigen Erkenntnissen durch einzelpixelweise Analyse nur unzureichend auswerten (siehe dazu auch PUISSANT ET AL. (2005), S. 734). Für die bisher erfolgten Klassifizierungen wurde daher kein pixelbasierter, sondern ein objektorientierter Ansatz gewählt, der neben spektralen auch objektspezifische Informationen in den Klassifizierungsprozess einfließen lässt. Erste Ergebnisse liegen bereits vor und legen den Schluss nahe, dass die Verwendung von 2,4-m-QuickBird-Szenen eine Abgrenzung verschiedener Pflanzenarten in den Marschlandschaften gestattet.

Stellvertretend für hoch auflösende Sensoren bearbeiteten PUISSANT ET AL. (2005) simulierte SPOT-5-Daten verschiedener Auflösungen (1, 2,5, 5 und 10 m). Sie verwendeten Grauwertübergangsmatrizen⁷, um eine kombinierte Spektral- und Texturanalyse einer Stadtszene von Strasbourg durchzuführen. Ziel war es festzustellen, ob eine Klassifikation in städtische Landbedeckungsklassen mit zusätzlichen Texturinformationen bessere Resultate liefert als ein rein spektraler Ansatz. Es wurde deutlich, dass mit zunehmender Auflösung der Nutzen einer Texturanalyse in Bezug auf eine Genauigkeitssteigerung der Klassifikation sinkt. Interessanterweise war die 1-m-Auflösung hier eine Ausnahme, bei der eine deutlichere Verbesserung (> 4 %) erzielbar war.

⁶ April 2005

⁷ Syn. „co-occurrence matrix“ oder „spatial greylevel dependence matrix“

4.2 Logistische Regression zur Klassifizierung von Fernerkundungsdaten

Es existiert eine Reihe von Untersuchungen mit der Zielstellung, FE-Daten mit Hilfe der logistischen Regression zu klassifizieren.

GOWDA ET AL. (2001) beschäftigten sich mit der Kartierung von Ackerbewirtschaftungsmethoden in Minnesota/USA auf Grundlage von Landsat-TM-Daten als Eingangsgrößen mehrerer logistischer Regressionsmodelle. Verschiedene Modellgleichungen, basierend auf der Forschung von DEVENTER ET AL. (1997), kamen zum Einsatz, um eine Klassifizierung in „konventionelle“ vs. „bewahrende“ Pflügestrategie zu ermöglichen. Zur Bewertung der Genauigkeit der Modelle wurden terrestrische Daten herangezogen. Das einfachste Modell, basierend auf Band 5 der Landsat-Szene, lieferte die höchste Klassifikationsgüte mit 77 % richtig zugeordneten Bewirtschaftungsformen. Eine grundsätzliche Eignung des Verfahrens wurde bestätigt, allerdings zeigte sich in Analogie zu anderen Untersuchungen, dass die testgebietsspezifischen Erkenntnisse bzw. Modellgleichungen nicht ohne weiteres auf andere Gebiete anwendbar sind.

Ebenfalls für die Landwirtschaft untersuchten BRICKLEMYER ET AL. (2002) die Verwendbarkeit von Landsat-ETM⁺-Daten zur Inventarisierung bzw. Kontrolle von Ackerbauflächen in Montana/USA. Die Motivation hinter dieser Forschungsarbeit bestand in der Tatsache, dass Ackerflächen unter Umständen als Kohlenstoffdioxid-Senken in Frage kommen. Inwieweit sie allerdings tatsächlich einen Beitrag zur so genannten Sequestrierung⁸ leisten, hängt u. a. davon ab, ob sie in der Brachzeit gepflügt werden oder nicht. Gepflügte Felder weisen eine deutlich höhere Mikroorganismenaktivität auf, die zu einer verstärkten Umsetzung organischer Substanzen im Boden und damit zu gesteigerter CO₂-Emission führen. Ungepflügte Felder hingegen behalten einen Großteil des während der Vegetationsphase eingelagerten CO₂ und qualifizieren sich damit besser als potentielle Senken von Kohlenstoffdioxid. BRICKLEMYER ET AL. (2002) wählten in einer Landsat-Szene eines Ackerbaugebiets Trainingsflächen für gepflügte und ungepflügte Felder aus, extrahierten die entsprechenden Grauwerte und erstellten daraus ein Regressionsmodell mit dem Pflügestatus als abhängiger Variable. Anschließend wurde die Satellitenszene basierend auf dem Modell klassifiziert. Die Ergebnisse zeigten, dass beide Feldtypen zu über 95 % richtig erkannt bzw. klassifiziert wurden.

FRASER ET AL. (2003) erforschten die Möglichkeiten der Kartierung von Waldbrandflächen mit Hilfe von Satellitendaten des SPOT-4-VEGETATION-Systems. Die Kartenerstellung erfolgte mit Hilfe eines multitemporalen Ansatzes und logistischen Modellen. Um das Rohergebnis (unkorrigierte Wahrscheinlichkeiten aus dem Modell) in eine Kartendarstellung mit dem Merkmal *Waldbrandfläche ja/nein* überführen zu können, wurden einige GIS⁹-gestützte Algorithmen zum Einsatz gebracht. Obwohl es sich als problematisch erwies, Brandflächen mit $A < 10 \text{ km}^2$ korrekt zu erkennen, sind die Ergebnisse insgesamt als sehr gut einzustufen. Fast alle größeren Brandflächen wurden korrekt klassifiziert; Defizite ergaben sich lediglich mit jenen Satellitenszenen, die größere Wolkenbereiche beinhalteten.

Um das Wachstum von jungen Sitka-Fichten-Beständen (*Picea sitchensis* Bong. Carr.) auf Plantagen kostengünstig und schnell erfassen zu können, untersuchten DONOGHUE ET AL. (2004)

⁸ C-Festlegung, also längerfristige Speicherung

⁹ Geographisches Informationssystem

sowohl Landsat-ETM⁺- als auch SPOT-4-HRVIR¹⁰-Daten mithilfe von GLMs. Die Einbindung zweier ähnlicher Systeme sollte zeigen, dass im Störfalle mit Ausfall eines Systems die Methodik nicht automatisch nutzlos würde. Das Hauptaugenmerk lag auf der Erfassung von fehl- bzw. unterentwickelten Beständen. Zur Kalibrierung der Modelle standen terrestrische Daten aus der GIS-Datenbank der letzten Inventuren durch die Forstverwaltung zur Verfügung. Auch DONOGHUE ET AL. (2004) entschieden sich bei der Grauwertextraktion für Mittelwerte aus 9×9 Pixeln im Zentrum der jeweiligen (homogenen) Probeflächen, um die Auswirkung von Lagedifferenzen zwischen aufgenommenen Bodendaten und Satellitenszenen zu verringern. Es konnte gezeigt werden, dass die straffste Beziehung zwischen der Höhe und spektralem Signal besteht, Brusthöhen-durchmesser, Grundfläche oder Dichte hingegen nur mittleren oder praktisch keinen Aussagewert besitzen.¹¹ Ebenfalls wurde deutlich, dass ein univariates Modell mit ETM⁺-Band 7 von Landsat völlig hinreichende Ergebnisse liefert und multivariate Ansätze mit Ratios bei zu vernachlässigender Steigerung der Genauigkeit nur eine unnötige Steigerung der Modellkomplexität mit sich bringen. Interessanterweise ergaben sich trotz höherer Auflösung des SPOT-Systems mit selbigem kaum bessere Ergebnisse, was auf die Homogenität der Bestände zurückgeführt wurde.

Als Bestandteil des „USGS Land Remote Sensing Program“ arbeiten PODRUZNY ET AL. (2004) zurzeit an einem Projekt zur Habitatkartierung des Grizzlybären in der weiteren Umgebung des Yellowstone-Nationalparks. Als Input dienen neben neun Szenen von Landsat ETM⁺ auch zahlreiche am Boden erhobene Daten zur zusätzlichen Ausstattung des Klassifizierungsmodells bzw. zur Validierung der gewonnenen Erkenntnisse. Als Klassifizierungsansätze werden sowohl logistische Regressionsmodelle als auch so genannte *Classification and Regression Trees* (CART)¹² in Erwägung gezogen. Das Hauptaugenmerk der Untersuchungen wird auf der Verbreitung und entsprechenden Kartierung der Baumart *Pinus albicaulis* ruhen, da bekannt ist, dass die Samen des Baumes einen großen Anteil der Nahrung des Grizzlybären ausmachen.

¹⁰ High Resolution Visible/Infrared

¹¹ Die Ausführungen legen allerdings nahe, dass tatsächlich der Kronenschlussgrad und damit letztendlich die mehr oder weniger sichtbare Bodenvegetation das Signal beeinflussen. Da jedoch eine ausgeprägte Abhängigkeit von Kronenschluss und Höhe beobachtbar ist, reicht die Erfassung der Höhe aus und liefert gute Ergebnisse.

¹² siehe dazu auch LAWRENCE ET AL. (2004)

Kapitel 5

Testgebiete und Datengrundlagen

5.1 Die Testgebiete im Überblick

Für das Projekt wurden fünf Testgebiete (TGs) festgelegt. Sie schließen in der Summe eine Fläche von mehr als 30.000 km² ein; dies entspricht etwa 8,5 % der Fläche Deutschlands. Der durchschnittliche, flächengewichtete Waldanteil beträgt ca. 42 % und liegt deutlich über dem Bundesmittelwert von rund 34 %. Mithin werden rund 11 % der deutschen Waldfläche durch diese fünf untersuchten Gebiete repräsentiert. Die Festlegung einer derartig großen Testfläche ist bei FE-Studien selten anzutreffen; stattdessen erfolgt vielfach eine Bearbeitung sehr spezifischer Fragestellungen. Da die Untersuchungen jedoch zeigen sollten, inwieweit das Design der Bundeswaldinventur durch die Einbindung von FE-Daten zu verbessern wäre, erfolgte die Auswahl der Testgebiete mit dem Ziel, möglichst unterschiedliche Waldverhältnisse in Deutschland abzudecken und zugleich potentielle Problemfälle in der FE-Phase einzuschließen. Die Testgebiete heißen:

- TG1 „Westliche Mittelgebirgsschwelle“
- TG2 „Nordostdeutsches Tiefland“
- TG3 „Östliche Mittelgebirgsschwelle“
- TG4 „Alpen“
- TG5 „Pfälzerwald“;

ihre Lage wird in Abbildung 5.2 auf Seite 31 verdeutlicht.

Vergleichende Informationen zur Alterstruktur und Baumartenverteilung liefert die Abbildung 5.1 auf der nächsten Seite. Überwiegend jüngere Bestände mit einem Alter von zumeist unter 75 a bilden den Wald in den ersten drei Testgebieten. Im Alpen-Testgebiet liegt eine flachere Verteilung mit Schwerpunkt auf älteren Beständen um 100 a vor. Eine zweigipflige Verteilung deutet sich im Pfälzerwald an; lokale Maxima liegen bei Altern von ca. 40 bzw. 75 a. Die Fichte ist die häufigste Baumart der westlichen und östlichen Mittelgebirgsschwelle sowie der Alpen, während im ostdeutschen Tiefland vorwiegend Kiefern stocken. Im Pfälzerwald ist das Verhältnis zwischen Nadel- und Laubbäumen ausgewogener. Alle vorkommenden Angaben zu Wuchsgebieten entstammen GAUER und ALDINGER (2005).

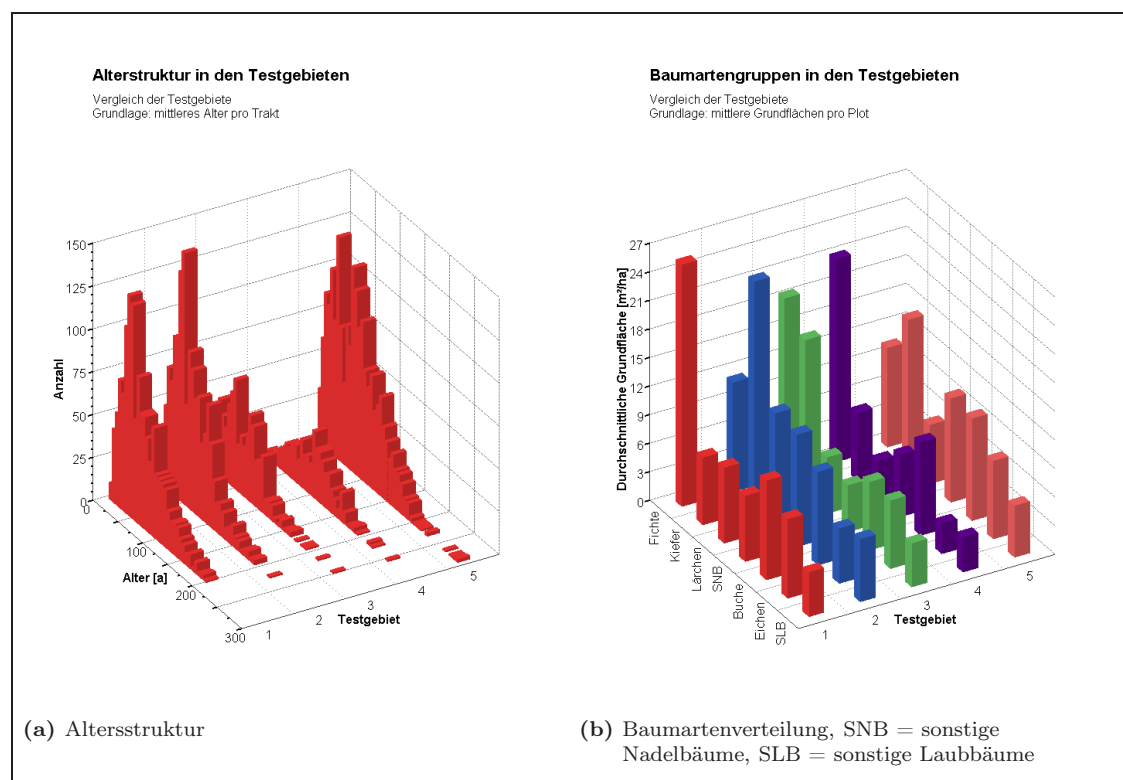


Abbildung 5.1: Altersstruktur und Baumarten der Testgebiete im Vergleich

5.1.1 Testgebiet 1: Westliche Mittelgebirgsschwelle

Das ca. 8.000 km² große Testgebiet 1 erstreckt sich über Teile Nordrhein-Westfalens, Hessens und Rheinland-Pfalz. Im Nordwesten des Gebietes befindet sich Dortmund; die von dort ausgehende Bundesautobahn BAB44 bildet den TG-Nordrand. Das Testgebiet umfasst die Naturparks Arnberger Wald, Homert, Ebbegebirge, Rothaargebirge, Bergisches Land und Rhein-Westerwald sowie mehrere größere Talsperren. Es weist einen Waldanteil von ca. 57 % auf. Dieser Wald verteilt sich im Wesentlichen auf die größeren, geschlossenen Waldgebiete der Naturparks. Nadel- und Nadelmischbestände dominieren die Wälder mit ca. 63 %. Die durchschnittliche Waldrandlänge von 53 m/ha spricht bezogen auf das WEI-BWI-Projekt für einen mittleren Fragmentierungsgrad. Von den fünf im TG vorkommenden Wuchsgebieten (WGs) sind zwei von flächenmäßiger Bedeutung: WG40 („Sauerland“) und westlich davon WG41 („Bergisches Land“).

5.1.2 Testgebiet 2: Nordostdeutsches Tiefland

Das zweite Testgebiet im Nordosten Deutschlands erstreckt sich über eine Fläche von etwas mehr als 9.600 km² und ist damit das größte der fünf Untersuchungsareale. Der Nationalpark Müritz gehört ebenso dazu wie der Naturpark Feldberg und Teile des Biosphärenreservats Schorfheide-Chorin. Zahlreiche Seen der Mecklenburger Seenlandschaft (u. a. die Müritz im Nordwesten) bewirken reich gegliederte Wälder mit einem Flächenanteil von insgesamt etwa 36 % des Testgebiets. Mit fast 73 % fällt die Dominanz der Nadel- und Nadelmischbestände noch deutlicher als im ersten Testgebiet aus. Die Waldareale sind gegenüber dem ersten Testgebiet weniger fragmentiert; dies

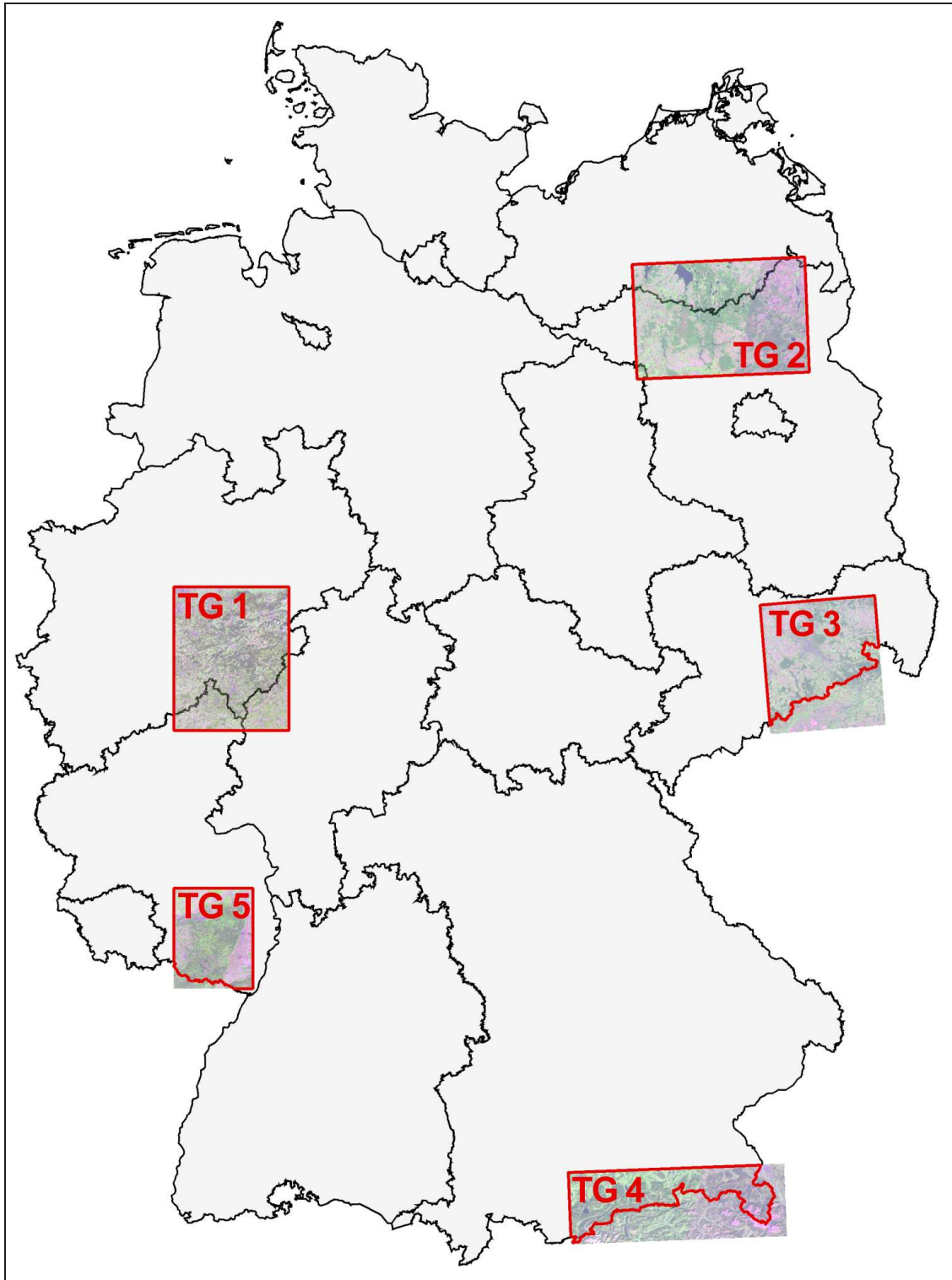


Abbildung 5.2: Lage der Testgebiete

äußert sich in einer Waldrandlänge von 47 m/ha. Sieben WGs sind im TG2 vereint; das „Mecklenburger Jungmoränenland“ (WG6) und südlich davon das „Ostmecklenburg-Nordbrandenburger Jungmoränenland“ (WG10) nehmen dabei den größten Anteil ein.

5.1.3 Testgebiet 3: Östliche Mittelgebirgsschwelle

Zur Untersuchung zergliederter Waldrandstrukturen und verstreuter Waldflächen wurde das dritte der Untersuchungsgebiete festgelegt. Es weist mit fast 60 m durchschnittlicher Waldrandlänge pro Hektar die stärkste Fragmentierung auf. Wesentliche Elemente auf den ca. 5.100 km² des TG sind das Erzgebirge, der Nationalpark Sächsische Schweiz, die Landschaftsschutzgebiete Friedewald, Dresdner Heide und Tharandter Wald sowie der ehemalige Truppenübungsplatz Königsbrück im Norden. Besonders Letzterer stellt durch seine der Heide ähnliche Gliederung eine Herausforderung für automatisierte Klassifizierungsverfahren dar. Auch hier stellen Nadelbaumbestände mit ca. 72 % das Gros des Waldes, welcher 30 % des TG bedeckt. In dieses TG fallen acht unterschiedliche Wuchsgebiete, die mit einer Ausnahme alle nennenswerte Anteile am Testgebiet haben.

5.1.4 Testgebiet 4: Alpen

Das topographisch anspruchsvollste aller Testgebiete ist TG4 mit einer Fläche von über 4.100 km². Es wird durch Salzburg im Nordosten, die Bundesautobahn BAB95 im Westen und die Landesgrenze im Süden begrenzt. Dieses TG wurde gewählt, um die Auswirkungen schwieriger Topographie zu analysieren. Die große Reliefenergie der Alpen mit den entsprechenden Höhenunterschieden von teilweise mehr als 1.000 m auf engem Raum sind dafür prädestiniert. Fast genau 60 % der Fläche sind bewaldet; dabei stehen mit fast 74 % einmal mehr die Nadelbestände im Vordergrund. Zusammen mit dem fünften Testgebiet verfügt es über die kompaktesten Waldflächen; die mittlere Waldrandlänge liegt bei rund 36 m/ha. Von einem winzigen Teil des Wuchsgebiets 80 abgesehen, befinden sich nur zwei WGs im Bereich des Alpen-Testgebietes. Es sind dies WG81 und 82 („Schwäbisch-Bayerische Jungmoräne und Molassevorberge“ bzw. „Bayerische Alpen“).

5.1.5 Testgebiet 5: Pfälzerwald

Das kleinste der Testgebiete, der Pfälzerwald, wurde vornehmlich deswegen festgelegt, da auf dieser Fläche Daten eines hoch verdichteten BWI-Netzes vorliegen. Der große zusammenhängende Komplex des Naturparks Pfälzerwald südlich der Bundesautobahn BAB6 bildet darin das bestimmende Element. Entsprechend sind fast 55 % der rund 3.600 km² großen Fläche bewaldet. Im Gegensatz zu den anderen TGs herrscht hier ein ausgeglichenes Verhältnis von Laub- und Nadelbeständen. Die aus dem verdichteten BWI-Netz resultierende große Stichprobenanzahl pro Fläche sollte zeigen, inwieweit dadurch bessere Auswertungen bzw. andere Verfahren mit hoch auflösenden Satellitensystemen möglich wären. Für die Untersuchungen im TG5 ist vor allem eines der fünf Wuchsgebiete interessant: WG72 „Pfälzerwald“, welches praktisch deckungsgleich mit dem gleichnamigen Gebirge ist.

5.2 Datengrundlagen

Die Fragestellungen des WEI-BWI-Projektes gestalteten die Auswahl der zu verwendenden Datengrundlage zu einem Optimierungsproblem. Bei einer kleinflächigen Studie bzw. einem regionalen Vorhaben mag es sinnvoll erscheinen, neben den Fernerkundungsdaten eine Fülle von Zusatzinformationen in den Klassifizierungsprozess einfließen zu lassen, um die erzielbare Genauigkeit zu steigern. Als Beispiele kommen digitale Forstgrundkarten, topographische Karten und andere (forstspezifische) GIS-Daten in Betracht. Die operationelle Umsetzung einer solchen Strategie ist im bundesweiten Maßstab jedoch nicht realisierbar, da neben den Kosten zwei weitere Faktoren zu berücksichtigen sind. Einerseits müssten die Daten idealerweise für das gesamte Inventurgebiet vom gleichen Zeitraum stammen, andererseits müssten Erhebung und Auswertung der Zusatzdaten vereinheitlicht erfolgen. In der Praxis können diese Bedingungen nicht erfüllt werden.

5.2.1 BWI-Daten

Wie einführend erwähnt erfolgte die Projektbearbeitung in enger Kooperation mit dem Institut für Waldökologie und Waldinventuren der BFH in Eberswalde, welches Auszüge der BWI-I/II-Datenbank für alle fünf TGs zur Verfügung stellte. Diese Auszüge lagen in Form mehrerer MS-Access-Datenbankdateien vor. Die darin enthaltenen Informationen gestatteten die Spezifikation der logistischen Regressionsmodelle, eine Visualisierung von Rohdaten und Ergebnissen über entsprechende Koordinaten in GIS sowie unterschiedlichste Auswertungen. Tabelle 5.1 auf der nächsten Seite enthält einen Überblick zur Anzahl und Art von Traktecken pro Testgebiet.

5.2.2 Landsat-TM- und -ETM⁺-Szenen

Primäre Quelle der FE-Daten waren die Systeme der Landsat-Serie. Um neben reinen Klassifizierungen auch Veränderungsberechnungen zwischen BWI I und II durchführen zu können, wurden für beide Zeitpunkte entsprechende Szenen beschafft. Dabei waren Umweltbundesamt und DLR von großer Hilfe, da sie die notwendigen Szenen unentgeltlich zur Verfügung stellten.¹

Die BWI-I-Daten wurden zur Analyse der TM-Szenen von 1989 bis 1992 verwendet, die BWI-II-Daten für die ETM⁺-Szenen von 1999 bis 2001. Sämtliche Landsatdaten waren bereits für das CLC-Projekt mittels „Nearest Neighbour Resampling“ neu berechnet worden, so dass die Pixelgröße der multispektralen Bänder 25 m, die des panchromatischen Kanals von ETM⁺ 12,5 m betrug.² Wegen der angestrebten Veränderungsrechnungen war eine Anpassung der TM-Daten auf die orthorektifizierten ETM⁺-Daten notwendig. Dieser Schritt erfolgte ebenfalls bereits beim DLR mittels Wahl geeigneter Passpunkte und einer polynomischen Anpassung 2. Grades.

Die Szenenbezeichnungen nach dem Worldwide Reference System 2 des USGS für Landsat (WRS-2) sowie Aufnahmezeitpunkte und Sonnenstand sind in Tabelle 5.2 auf Seite 35 zusammengefasst; sie entstammen den zugehörigen Metadatensätzen. Selbigen war zu entnehmen, dass die für das CLC2000-Projekt erworbenen Originalszenen (ETM⁺) ausnahmslos die höchste Qualitätseinstufung („9“ auf einer Skala von 0–9) erhielten. Für Landsat 5 (TM) standen diese Angaben

¹ Die verwendeten Szenen entstammen dem schon erwähnten CORINE Landcover Projekt (CLC), in dem Europas Bodenbedeckungsklassen nach einem einheitlichen Interpretationsschlüssel kartiert wurden. Diese Kartierung ist für die Zeitpunkte 1990 und 2000 abgeschlossen.

² Diese Veränderung erfolgte ohne jeglichen Informationsgewinn, sondern diente lediglich einer besseren Auswertbarkeit.

5 Testgebiete und Datengrundlagen

Tabelle 5.1: Übersicht zu Anzahl und Art der BWI-Traktecken pro Testgebiet.

Die Feldaufnahmen zur BWI I begannen 1986 auf dem Gebiet der damaligen BRD. Daten auf dem Gebiet der neuen Bundesländer (TG2 und 3) liegen erst seit der BWI II vor.

Art der Traktecke	TG1	TG2	TG3	TG4	TG5	Summe
Bundeswaldinventur I						
Nichtwaldecken	865	n/v	n/v	386	436	1687
Waldecken ^a	1099	n/v	n/v	550	486	2135
Waldecken ohne Wald ^b	70	n/v	n/v	41	24	135
Summe	2034	n/v	n/v	977	946	3957
Bundeswaldinventur II						
Nichtwaldecken	1249	2809	1756	390	1631	7835
Waldecken ^a	1449	1478	720	565	1918	6130
Waldecken ohne Wald ^b	69	61	21	22	51	224
Summe	2767	4348	2497	977	3600	14189

Quelle: BWI-Datenbankauszug der BFH

^a Wald im spektralen Sinne, ohne Blößen und Nichtholzboden

^b Blößen und Nichtholzboden (ebenfalls „Wald“ laut BWI-Nomenklatur)

nicht zur Verfügung.

Für alle fünf Testgebiete wurden jeweils zwei Szenen³ der TM- und ETM⁺-Sensoren geliefert, die in der Summe mit mehr oder weniger großen Überlappungsbereichen die Testgebiete vollständig abdeckten. Diese Szenen hatten hinsichtlich der Abmessungen mit den Originalszenen von Landsat wenig gemein, da sie im CLC-Projekt verändert worden waren. Leider war es nur für TG3 möglich, beide Szenen je Sensor vom selben Aufnahmetag zu akquirieren. Alle anderen Szenen unterscheiden sich hinsichtlich des Aufnahmezeitpunktes zum Teil erheblich, bei den TM-Daten liegen im Extremfall über 3 Jahre zwischen den Aufnahmezeitpunkten (Tabelle 5.2 auf der nächsten Seite). Damit geht die Gefahr einher, unterschiedliche atmosphärische Bedingungen anzutreffen, die einen ganzheitlichen Klassifizierungsansatz erschweren. Vom phänologischen Standpunkt betrachtet können überdies die Aufnahmen vom Mai bzw. späten September nicht als optimal angesehen werden.

Visuell sind die TM-Daten insgesamt schlechter als die Landsat-7-Szenen einzustufen. Z. T. ist eine deutliche Streifenbildung sichtbar, auch wirken sie weniger kontrastreich. Nach Angaben des DLR (Briefwechsel) gelang im TG4 die geometrische Anpassung der TM-Daten auf die ETM⁺-Szenen nicht vollständig. Dadurch ergibt sich besonders im Gebirge stellenweise ein Versatz von 50 m. Über TG5, welches verrauschter als der Rest der TGs wirkt, liegt sichtbarer Dunst.

Die ETM⁺-Szene 196/24 (TG1 Norden) weist im Nordosten Quellbewölkung auf; ebenso sind in Szene 192/27 (TG4 Osten) zahlreiche kleine Wolken über Bergen sichtbar. Der Überlappungsbereich zwischen Ost- und Westszene ist in den Alpen allerdings so groß, dass weite Teile des TG4 durch die (wolkenfreie) Westszene (193/27) abgedeckt werden. Szene 192/25 (TG3 Süden) zeigt

³ georeferenziert auf das Gauß-Krüger-System

Tabelle 5.2: Im WEI-BWI-Projekt verwendete Landsat-Daten

TG	Sensor	Szene im WRS-2		Aufnahmedatum		Sonnenstand ^b	
		Path	Row	Tag	Zeit ^a	Höhe in °	Azimut in °
1	TM ^c	196	24	25.05.1989	n/v	n/v	n/v
		196	25	05.08.1992	n/v	n/v	n/v
	ETM ⁺	196	24	15.05.2000	10:13:23	54,5	150,5
		196	25	05.07.2001	10:11:02	58,0	141,7
2	TM	193	23	07.07.1989	n/v	n/v	n/v
		194	23	19.05.1992	n/v	n/v	n/v
	ETM ⁺	193	23	14.08.2000	09:53:25	48,2	151,2
		194	23	04.09.1999	10:01:27	42,0	156,4
3	TM	192	24	26.07.1991	n/v	n/v	n/v
		192	25	26.07.1991	n/v	n/v	n/v
	ETM ⁺	192	24	24.09.2000	09:47:00	35,7	159,0
		192	25	24.09.2000	09:47:23	36,9	158,0
4	TM ^c	192	27	07.08.1991	n/v	n/v	n/v
		193	27	30.08.1991	n/v	n/v	n/v
	ETM ⁺	192	27	26.08.2001	09:46:30	48,8	147,3
		193	27	13.09.1999	09:56:46	43,5	153,8
5	TM ^c	195	26	07.09.1989	n/v	n/v	n/v
		196	25	05.08.1992	n/v	n/v	n/v
	ETM ⁺	195	26	15.08.2001	10:04:43	50,9	145,8
		196	25	05.07.2001	10:11:02	58,0	141,7

Quelle: Metadaten der FE-Datensätze

^a Greenwich Mean Time (MEZ – 1 h)

^b relativ zum wahren Szenenzentrum;

Höhe: $0^\circ \hat{=}$ Sonne am idealen Horizont, $90^\circ \hat{=}$ Sonne im Zenit

Azimut: $0 \triangleright 90 \triangleright 180 \triangleright -90 \triangleright -180^\circ \hat{=}$ N \triangleright O \triangleright S \triangleright W \triangleright S

^c ohne Band 6 (TIR)

entlang der Elbe gut sichtbare Artefakte, d. h., in regelmäßigen Abständen (ca. 600 m) erscheint das Bild streifenweise um 1 bis 2 Pixel versetzt. Möglicherweise ist dieser Effekt auf die innere Geometrie der Szene und deren Korrektur zurückzuführen.

5.2.3 QuickBird-Szenen

Das 4×4 km-Raster der BWI wurde gewählt, um auf Bundesebene statistisch abgesicherte Aussagen treffen zu können. Einige Bundesländer verdichteten bei der zweiten BWI dieses Netz auf $2,83 \times 2,83$ bzw. 2×2 km, um für kleinere Auswertungseinheiten ebenso abgesicherte Ergebnis-

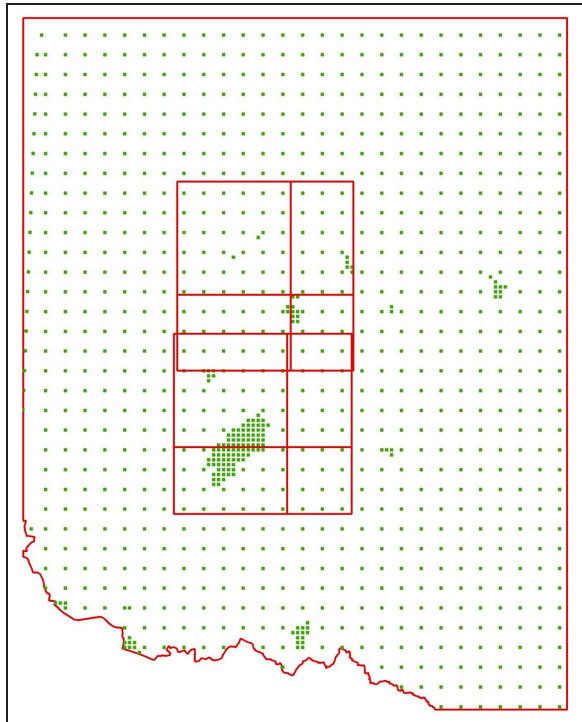


Abbildung 5.3

Die Lage der 8 QuickBird-Kacheln innerhalb des Pfälzerwaldes (TG5).

Die grünen Punkte repräsentieren das verdichtete Stichprobennetz der BWI II in Rheinland-Pfalz (2×2 km). Die Trakthäufungen entstehen durch eine weitere Erhöhung der Stichprobendichte in ausgewählten Gebieten ($0,5 \times 0,5$ km; bspw. in der Kernzone des Biosphärenreservats).

se zu erzielen (POLLEY ET AL., 2004).⁴ In Teilen von Rheinland-Pfalz wurde das BWI-II-Netz noch weiter verdichtet, so dass ein $0,5 \times 0,5$ km-Netz resultierte. Dadurch ist auf dem relativ kleinen Gebiet des Pfälzerwaldes ein großer Stichprobenumfang verfügbar. Diese Situation sollte genutzt werden, um zu eruieren, inwieweit höher auflösende Sensorensysteme wie QuickBird zur Unterstützung einer Regionalinventur geeignet sind.

Zur Abdeckung des spezifizierten Bereiches wurden 2×4 Kacheln pro panchromatischem und multispektralem Produkt geliefert, welche im UTM-System referenziert waren (siehe auch Abbildung 5.3). Nachstehende Angaben gelten für alle Kacheln. Selbige entsprechen der Produktebene „LV2A“, die eine radiometrische Korrektur und eine Entzerrung mittels eines groben DHM⁵ beinhaltet. Zum Resampling der Daten wurde die Cubic-Convolution-Methode angewendet. Die Aufzeichnung der Daten erfolgte am 05.09.2004 gegen 10:35 Uhr bei einem Sonnenstand von ca. $161^\circ/46^\circ$ (Azimut/Höhe); dabei stand der Satellit auf $235^\circ/78^\circ$ (Azimut/Höhe). Aus dieser Konstellation erhebt sich kein Problem durch störende Reflexionen. Der panchromatische Kanal liegt in einer Auflösung von 70 cm pro Pixel vor, während die Pixelgröße der Bänder Blau/Grün/Rot/NIR 280 cm beträgt. Verglichen mit Landsat ist die radiometrische Auflösung der QuickBird-Daten zumindest nominal siebenmal größer: Jedes Band nimmt spektrale Informationen in einem Bereich von 2048 Graustufen auf, was einer Quantisierung von 11 Bit (statt 8 Bit bei Landsat) entspricht.

5.2.4 Orthophotos

Neben den bisher aufgeführten Daten wurden weitere Informationsquellen genutzt, um die Klassifikationsergebnisse überprüfen und bewerten zu können. Für diese Aufgaben und zur Untersuchung

⁴ Auf diese Weise wurden parallel zur Bundeswaldinventur *Landeswaldinventuren* durchgeführt.

⁵ Die Metadaten enthalten den Hinweis auf ein „coarse“ DEM zur Korrektur topographiebedingter Verzerrungen.

kombinierter Klassifikationsansätze wurden digitale Orthophotos angeschafft. Diese lagen in unterschiedlichen Auflösungen vor. Für die westliche Mittelgebirgsschwelle (TG1) betrug diese ca. 94 cm; befliegen worden war das Gebiet zwischen 1996 und 1999. Teile der östlichen Mittelgebirgsschwelle (TG3) waren mit 80-cm-Orthophotos aus dem Zeitraum 1996–2001 abgedeckt. Im alpinen Raum (TG4) lagen so genannte Quicklooks auf einer CD des Bayerischen Landesvermessungsamtes mit einer Auflösung von ca. 5,7 m vor, die laut Impressum 1999 zusammengestellt worden war. 1-m-Daten standen im fünften Testgebiet zur Verfügung, die den Jahren 1998 und 1999 entstammen. Aus Kostengründen konnten keinerlei Orthophotos für das nordostdeutsche Tiefland (TG2) akquiriert werden.

5.2.5 Topographische Karten

Als weiteren Ansatz zur Überprüfung und statistischen Analyse (bspw. Stratifizierungsansätze) kamen topographische Karten in Betracht. Für das Projekt und diese Arbeit stand die DTK25-V des Bundesamtes für Kartographie und Geodäsie in Einzelebenen für jedes Testgebiet zur Verfügung. Zusätzlich wurde die CD TOP200 angeschafft, die bundesweit eine topographische Karte in Maßstäben von 1 : 8 Mio. bis 1 : 200.000 bereitstellt. Mit ihrer Hilfe vereinfachte sich die Orientierung innerhalb der Testgebiete.

Kapitel 6

Methodik

Dieses Kapitel ist der Methodik gewidmet, welche zur Bearbeitung der einführend dargelegten Fragestellungen gewählt wurde. Soweit merkmals- bzw. sensorenspezifische Besonderheiten zu beachten waren, wird auf die methodischen Unterschiede hingewiesen, ansonsten gelten die Ausführungen für alle Merkmale und beide Sensorentypen gleichermaßen.

6.1 Datenvorbereitung für TM- und ETM⁺-Szenen

6.1.1 NDVI-Berechnung, Grauwertextraktion, Datensynthese

Mit dem Spatial Modeler von Erdas Imagine erfolgte als erstes die Berechnung und Einbindung des Vegetationsindex' NDVI als siebentes bzw. achtes Band¹ in allen Szenen. Auch mussten die relevanten Informationen im Datenbankauszug bedarfsgerecht zusammengestellt werden: Die Datenbank hielt die Koordinaten der BWI-Aufnahmepunkte im Gauß-Krüger-System vor; dabei konnten pro TG mehrere Meridianstreifen auftreten. Bspw. lagen von den 2767 BWI-II-Traktecken des TG1 168 im 2er-Streifen, der Rest im 3er-Streifen. Im Gegensatz dazu lagen die FE-Daten in szenenweit einheitlichen Projektionen vor; TG1 z. B. im dritten Meridianstreifen. Daher wurden alle Koordinaten im Datenbankauszug auf jenen Streifen umgerechnet, welcher für die Projektion der entsprechenden FE-Szene festgelegt worden war. Anschließend wurden Koordinatenlisten zusammengestellt, die die Rechts- und Hochwerte aller Traktecken pro TG enthielten. Diese Listen wiederum bildeten die Voraussetzung für die anschließende Grauwertextraktion aus den Satellitenszenen mit dem Pixel-To-ASCII-Modul von Erdas.

Die erzeugten Grauwerttabellen wurden Access-gerecht konvertiert in die Datenbank importiert und mit den Trakteckeninformationen verknüpft. Zusätzlich wurden alle notwendigen Informationen zu den wichtigsten Zielgrößen (soweit möglich) für jede Traktecke zusammengestellt, so dass ein kompletter Datensatz resultierte, der für die BWI-Traktecken sowohl die terrestrisch erhobenen als auch jene mittels Fernerkundung gewonnenen Daten vereinte. Angaben zu Waldstatus, Totholz und Baumartenmischungen ließen sich aus der Datenbank selbst entnehmen; für die Größen Stammdurchmesser, Baumhöhe, Alter, Baumanzahl pro Hektar und Vorräte waren SAS-basierte Hochrechnungen notwendig (RIEDEL, 2006).

¹ siehe Übersichtstabelle 5.2 auf Seite 35: einige Szenen waren ohne Band 6 geliefert worden

6.1.2 Ausschluss von Gewässerflächen

Spektral betrachtet ähneln (nicht durch Schwebstoffe getrübt) Gewässer, also größere Flüsse und Seen, insbesondere im ETM⁺-Band 2 sehr der Bedeckungsklasse Wald. Konsequenterweise wurden in einem Vorab-Klassifizierungstest sämtliche Gewässer überwiegend als Wald klassifiziert, was zu einer Waldflächenüberschätzung führen würde. Ein gängiger Ansatz zur Vermeidung dieses Problems ist die Entfernung von Gewässern aus der untersuchten Szene *vor* der Klassifizierung (ITTEN ET AL., 1992), der auch hier angewendet wurde.

In den TM-Szenen führte eine einfache Schwellenwert-Regel im Band 5 nicht zum gewünschten Ergebnis.² Um den Ausschluss von Schattengebieten zu vermeiden, hätte der Schwellenwert so niedrig gewählt werden müssen, dass die meisten Gewässerpixel ebenso wenig entfernt worden wären. Es wurde stattdessen auf ein logistisches Modell (LogReg) mit mehreren Bändern als Eingangsgrößen zurückgegriffen. Dazu erfolgte die Definition von Trainingspixeln für die Klassen *Wasser*, *Schatten* und *Sonstiges*. *Schatten* und *Sonstiges* wurden zur Klasse *Nichtwasser* zusammengefasst, da ja nur die Gewässerpixel entfernt werden sollten.

Uni- und multivariate Ansätze wurden auf ihre Eignung getestet. Nach Vergleich der Klassifizierungsgenauigkeiten wurden Modelle mit den Bandkombinationen 1-3-4-5-NDVI bzw. 1-2-3-4-5-7-NDVI eingesetzt. Ein einheitliches Modell hätte die erreichbare Genauigkeit vermindert. Zur Kalibrierung der Modelle wurden jeweils gleiche Anzahlen Wasser- und Nichtwasserpixel verwendet. Die Modelle trennen die Trainingsdaten ausnahmslos mit Klassifizierungsgenauigkeiten von über 99 %, was maßgeblich der Auswahl der Trainingspixel geschuldet sein dürfte. Wendet man sie über den Spatial Modeler auf die gesamte Szene an, zeigt ein visueller Vergleich, dass sie Pixel in Schlagschattenbereichen zum Teil als *Wasser* erkennen, während Pixel vor allem im Flachwasserbereich von Seen mitunter nicht als Wasserpixel klassifiziert werden. Eine Quantifizierung des Klassifizierungsfehlers über visuelle Vergleiche hinaus war nicht möglich, da sie mit erheblichem manuellem Aufwand verbunden gewesen wäre. Zur Unterdrückung des Salz- und Pfeffereffekts wurde nach der Klassifizierung ein 3×3 -Medianfilter eingesetzt.

Das Testgebiet „Pfälzerwald“ nahm eine Sonderstellung ein; Wasserpixel treten dort zu spärlich auf, um ein abgesichertes eigenes LogReg-Modell zu erstellen. Die stattdessen erfolgte Anwendung des Modells von TG4 erbrachte gute Ergebnisse beim Entfernen der Wasserflächen.

Die Modelle zur Klassifizierung von Gewässern wurden mit dem Spatial Modeler umgesetzt, so dass Gewässer-Masken entstanden, die nur die Wasserflächen zeigten. Im abschließenden Schritt erfolgte per Verschneidung der Ausschluss (eine Maskierung) der unerwünschten Flächen in den Originalszenen.

Für die ETM⁺-Szenen war die Definition eines Schwellenwertes im Band 5 im Spatial Modeler ausreichend, um eine gute Gewässerentfernung zu erreichen. Dieser Schwellenwert musste allerdings auch szenenweise ermittelt werden; eine einheitliche Regel hätte zu keinem zielführenden Ergebnis geführt. Die Regeldeklaration für Band 5 konnte direkt in die Modelle zu Wald/Nichtwald-Klassifizierung integriert werden, was eine separate Verschneidung von Gewässerkarten und Originalszenen erübrigte.

² Im Spektralbereich des Bandes 5 treten Gewässer sehr deutlich gegenüber Vegetationsflächen hervor. Für Schattenbereiche gilt dies allerdings nicht.

6.2 Datenvorbereitung für QuickBird-Szene

6.2.1 NDVI-Berechnung, Grauwertextraktion, Datensynthese

Der normalisierte Vegetationsindex wurde für die QuickBird-Szene nachträglich in der Datenbank berechnet, um für die Parametrisierung der Regressionsmodelle zur Verfügung zu stehen.

Zur spektral-basierten Auswertung von hoch auflösenden Sensoren existieren unterschiedliche Auffassungen (siehe Literaturübersicht). Um den Einfluss der Pixelgröße auf das Klassifizierungsergebnis untersuchen zu können, erfolgte eine Neuberechnung der Szene mit den Pixelgrößen 5, 15 und 25 m unter Verwendung der Resamplingverfahren (LEICA GEOSYSTEMS, 2003):

- Nearest Neighbor,
- Bilinear Interpolation,
- Cubic Convolution und
- Bicubic Spline.

Laut *Erdas Field Guide* (LEICA GEOSYSTEMS, 2003) weist jedes dieser Verfahren spezielle Eigenschaften auf, die es für bestimmte Einsatzzwecke optimal erscheinen lässt. Betrachtet man allerdings die Gegenüberstellungen der Vor- und Nachteile jedes Verfahrens, so ist eine eindeutige Entscheidung nicht zu treffen. Zum Zwecke des Vergleiches wurden daher alle 3 Umrechnungen des Originaldatensatzes mit allen vier Verfahren durchgeführt, um über den Einfluss der Pixelgröße hinaus auch den der Resamplingmethode aufzeigen zu können. In der Summe wurden also 13 Datensätze ausgewertet: sowohl die Originalversion der QuickBird-Szene mit 3 m Auflösung als auch die durch vier unterschiedliche Berechnungsverfahren erstellten Szenen mit den Pixelgrößen 5, 15 und 25 m.

Nach der Erstellung der 13 Datensätze kam die gleiche Vorgehensweise wie bei den Landsat-Szenen zum Einsatz: Extraktion der Grauwerte mit den BWI-Koordinatenlisten und Verknüpfung mit den relevanten Informationen aus der Datenbank.

6.2.2 Ausschluss von Gewässerflächen

In der QuickBird-Szene als Ausschnitt des Testgebietes „Pfälzerwald“ wurde auf eine Gewässerentfernung verzichtet. Sie enthielt nur wenige, sehr kleine Wasserflächen mit einem Anteil von unter 0,1 ‰ der Gesamtfläche, deren Entfernung keinen signifikanten Einfluss auf die Klassifizierungsgenauigkeit ausgeübt hätte.

6.3 Logistische Regression zur Kartierung von Waldvorkommen

6.3.1 Zusammenstellung der Datensätze

Für die szenenweise Analyse der Zusammenhänge zwischen terrestrisch und durch FE erhobenen Daten wurden über Datenbankabfragen die benötigten Datensätze aus dem Gesamtbestand bereitgestellt. Aus diesen je eine Szene umfassenden Datensubsets wurden danach sämtliche Datensätze

(= Pixel) entfernt, bei denen in allen Bändern gleichzeitig der Grauwert 0 auftrat. Dieser Schritt war zur Vermeidung überflüssigen Rechenaufwandes notwendig, da jede Szene die Ausmaße des entsprechenden Testgebietes hatte, Teile davon jedoch nur mit „Nullpixeln“ gefüllt waren.³ So wurden die Untersuchungen nur auf den jeweils mit Daten besetzten Teil der Szenen eingeschränkt. Auf diese gefilterten Abfragen wurde in SPSS per Datenbankabfrage zugegriffen.

6.3.2 Vorbereitende Untersuchungen

In Voruntersuchungen wurden sowohl die Grauwerte als auch die intervallskalierten Zielgrößen auf Normalverteilung getestet und Streudiagramme mit Sensorkanälen und Zielgrößen erstellt. Die Ergebnisse bestätigten die Notwendigkeit nichtparametrischer Verfahren, da in den seltensten Fällen von einer Normalverteilung ausgegangen werden konnte.

6.3.3 Uni- und multivariate Analysen

Der hier verfolgte Ansatz zur Erstellung der LogReg-Modelle ist den Empfehlungen von HOSMER und LEMESHOW (1989) angelehnt. Zunächst erfolgte eine univariate Betrachtung der Zusammenhänge zwischen beobachtetem Waldstatus und Grauwert eines Pixels. Die verwendete SPSS-Prozedur LOGISTIC REGRESSION nahm dabei *Waldstatus*—in den Ausprägungen 0 = Nichtwald bzw. 1 = Wald—als abhängige Variable auf, während der jeweils betrachtete Sensorkanal als erklärende Variable fungierte. Das Ziel dieser Voruntersuchung war die Identifizierung von Sensorbändern, die keinen Beitrag zur Vorhersage des Waldstatus' aufweisen würden. Kanäle, die an dieser Stelle bereits sehr geringe Klassifizierungserfolge bzw. wenig signifikante Resultate lieferten, wurden von den weiteren Betrachtungen ausgeschlossen.

Jene Bänder, denen ein maßgeblicher Beitrag zur Klassifizierung zuzusprechen war, wurden zusammen in multivariaten Modellen überprüft. Kanäle, deren Koeffizienten nicht signifikant verschieden von Null waren, wurden in diesem Schritt aus dem Modell entfernt. Auf diese Weise ergab sich jeweils ein „schlankes“ Modell ohne überflüssige Eingangsgrößen. Die Signifikanz der einzelnen Koeffizienten war jedoch nicht das einzige Kriterium zur Aufstellung der Modelle. Die Entwicklung möglichst einheitlicher Ansätze würde die operationelle Implementierung der Methoden in einer bundesweiten Anwendung erheblich vereinfachen. Zu diesem Gesichtspunkt folgen im Ergebnisteil weitere Überlegungen.

6.3.4 Beurteilung und Interpretation der Regressionsmodelle

Bevor die Modellfindung abgeschlossen war, wurde das favorisierte Modell in Bezug auf Anpassung und Angemessenheit geprüft. MENARD (2001) schlägt dafür eine dreiteilige Strategie vor. Als erstes wird anhand ausgewählter Teststatistiken das Modell als Ganzes bewertet. So lässt sich feststellen, in welchem Maße es sich den empirischen Daten anpasst und korrekte Klassifizierungsergebnisse liefert. Im zweiten Abschnitt erfolgt eine differenzierte Beurteilung der Einzelbeiträge der verwendeten unabhängigen Variablen zur Qualität des Modells (in Form ihrer Koeffizienten). Zu dieser Thematik existieren unterschiedliche Auffassungen: für manche Wissenschaftler gilt die

³ Z. B. besteht die Nordszene des Testgebietes „Östliche Mittelgebirgsschwelle“ nur bis zur Elbe aus richtigen Bildpixeln; südlich davon befinden sich ausschließlich Pixel mit Grauwerten = 0. Dieser Teil des TG wird von der Südszene abgedeckt.

Maxime des einfachsten Modells, andere wiederum vertreten die Meinung, alle logisch sinnvollen Variablen in die Spezifikation aufzunehmen, selbst wenn einige davon nicht signifikant zur Verbesserung des Modells beitragen (HOSMER und LEMESHOW, 1989; KLEINBAUM und KLEIN, 2002). Argumente lassen sich für beide Ansichten finden. Zählt die leichte Implementierbarkeit und kurze Rechenzeiten, sollte einem sparsamen Ansatz der Vorzug gegeben werden. Geht es dagegen hauptsächlich um eine umfassende Abbildung von möglicherweise kausalen Zusammenhängen, erscheint ein erweiterter Parametersatz durchaus angebracht. Ein dritter und letzter Schritt besteht darin zu prüfen, ob die Form des Modells adäquat und die damit verbundenen Anforderungen erfüllt sind (als logistic regression diagnostics bezeichnet).

Zur Bearbeitung dieser Fragen liefert SPSS einige diagnostische Größen standardmäßig als Teil der Routinen LOGISTIC REGRESSION und NOMINAL REGRESSION, andere mussten nachträglich ermittelt werden. Insbesondere zur Einzelbeurteilung der Koeffizienten und Datensätze war die Berechnung zusätzlicher Prüfgrößen notwendig.

6.3.4.1 Schritt 1: Gesamtmodell

Die tabellarische Übersicht des Klassifikationsergebnisses eines betrachteten Modells ist von zentraler Bedeutung für seine Beurteilung. In Form einer Kontingenztabelle werden die vorhergesagten den Referenzwerten gegenüber gestellt. In der Fernerkundung werden solche Kontingenztabelle oft als Konfusions- oder Fehlermatrizen bezeichnet, allerdings mit dem Unterschied, dass die Zuordnung von beobachteten und klassifizierten Daten einer Transponierung der SPSS-Struktur entspricht. In der SPSS-Ausgabe wird der Waldstatus der terrestrisch erhobenen Daten (ground truth) in die Zeilen, der aus der Modellvorhersage in die Spalten gesetzt. Weiterhin zu beachten ist: Die Genauigkeitsangaben beziehen sich *ausschließlich* auf das verwendete Modell und die ihm zugrunde liegenden Parametrisierungsdaten.⁴ Es wird *keine* Aussage zur Genauigkeit der gesamten Klassifikation einer Landsat- oder QuickBird-Szene mit Hilfe dieses Modells getroffen! Eine solche Bewertung ist nur mit in der Fernerkundung üblichen Methoden anhand des klassifizierten Endergebnisses möglich (SMITS ET AL., 1999; LILLESAND ET AL., 2004; FOODY, 2005). Mehr zum Thema der Qualitätssicherung und Ergebnisvalidierung wird im Abschnitt 6.3.6 ab Seite 49 ausgeführt.

Tabelle 6.1 auf der nächsten Seite zeigt eine solche Klassifikationstabelle. Die schwarz dargestellten Zahlen sind Teil der SPSS-Standardausgaben. Nach LILLESAND ET AL. (2004) lassen sich aus den Informationen der Tabelle weitere Aussagen ableiten (in roter Schrift: nachträgliche Einfügungen). Insbesondere die Betrachtung der Klassifizierungsgenauigkeit erfolgt bei FE-Fragestellungen oftmals unterteilt: Es werden Erstellergenauigkeit, Nutzergenauigkeit und Gesamtgenauigkeit unterschieden. Damit wird den unterschiedlichen Zielstellungen von Kartenerstellern und -nutzern Rechnung getragen. Die Prozentangaben der SPSS-Tabelle stellen die Erstellergenauigkeit dar. Es wird also die Anzahl richtig klassifizierter Pixel ins Verhältnis zur realen Anzahl Pixel in der betrachteten Klasse gesetzt. Im Beispiel wurden von 100 Nichtwaldpixeln 99 richtig zugeordnet—dies entspricht einer Erstellergenauigkeit von 99 %. Für die Klasse Wald ergibt sich

⁴ Nach BALTES-GÖTZ (2004) und WUENSCH (2005) sind die Angaben von SPSS bezüglich der Klassifizierungsqualitäten eines logistischen Modells optimistisch. Verursacht wird die Überschätzung der Genauigkeit dadurch, dass SPSS kein Jackknife-Verfahren zur Schätzung der Koeffizienten verwendet.

Tabelle 6.1: SPSS-Ausgabe: Beispiel einer Kontingenztabelle mit dem Klassifizierungsergebnis

		Vorhersage ^a		
		Waldstatus		% Richtige (E)
Beobachtung ^b		0	1	
	Waldstatus	0	99	1
	1	14	95	87,2
% Richtige (N)		87,6	99,0	(G) 92,8

^a laut Regressionsmodell^b laut BWI-Datenbank

E = Ersteller-, N = Nutzer-, G = Gesamtgenauigkeit

eine solche von

$$95/(14 + 95) \times 100 \% \approx 87,2 \%$$

Dem Ersteller ist es also wichtig zu erfahren, wie gut eine Merkmalsklasse identifizierbar ist. Den Nutzer einer Karte oder Klassifikation interessiert dagegen eher, mit welcher Sicherheit ein als Wald klassifiziertes Pixel auch wirklich dieser Klasse angehört. Hier wird ein anderes Verhältnis gebildet: Anzahl richtig klassifizierter Pixel zu Anzahl aller dieser Klasse zugeordneten Pixel. In Tabelle 6.1 beträgt die Nutzergenauigkeit für Nichtwald

$$99/(99 + 14) \times 100 \% \approx 87,6 \%,$$

für Wald

$$95/(1 + 95) \times 100 \% \approx 99,0 \%$$

Die Summe der Elemente auf der Hauptdiagonalen von links oben nach rechts unten geteilt durch die Gesamtzahl der klassifizierten Pixel ergibt die Gesamtgenauigkeit, im Beispiel

$$194/209 \times 100 \% \approx 92,8 \%$$

Als Ergänzung der Kontingenztabelle empfiehlt MENARD (2001) eine Maßzahl zur Vorhersageeffizienz des Modells. Derartige Indizes besitzen stets die gleiche Grundstruktur:

$$\text{Vorhersageeffizienz} = \frac{(\text{Fehler ohne Modell}) - (\text{Fehler mit Modell})}{(\text{Fehler ohne Modell})}. \quad (6.1)$$

Die Berechnung des Terms *Fehler ohne Modell* hängt vom verwendeten Index ab; *Fehler mit Modell* ist indes stets die gleiche Größe—es ist die Summe aller Fehlklassifikationen des Modells. Für Klassifikationsmodelle wie die hier untersuchten eignet sich der Index τ_p , dessen *Fehler-ohne-Modell*-Term mit

$$\sum_{i=1}^N f_i \left[\frac{(N - f_i)}{N} \right] \quad (6.2)$$

berechnet wird (MENARD, 2001). N ist der Stichprobenumfang, im vorliegenden Fall also die

Anzahl Pixel, zu denen Informationen in der BWI-Datenbank vorliegen. f_i bezeichnet die Anzahl Fälle innerhalb der Klasse i . Setzt man Gleichung 6.2 in 6.1 ein, resultiert

$$\tau_p = \frac{\left\{ \sum_{i=1}^N f_i \left[\frac{(N-f_i)}{N} \right] \right\} - n_{\text{Fehlklassifikationen}}}{\sum_{i=1}^N f_i \left[\frac{(N-f_i)}{N} \right]}. \quad (6.3)$$

Die Spannweite von τ_p hängt von den Randsummen der Kontingenztabelle ab; im Falle einer vollständig richtigen Klassifizierung ergibt sich jedoch immer $\tau_p = 1$. Negative Werte legen den Schluss nahe, dass der Modelleinsatz schlechtere Vorhersagen erbringt, als per Zufall möglich wären.

Andere Prüfgrößen wurden zwar ebenfalls im Modellfindungsprozess berücksichtigt, waren in ihrer Bedeutung allerdings den beiden oben vorgestellten Ausgaben untergeordnet. Dazu zählen zum einen das Pseudobestimmtheitsmaß R_L^2 , zum anderen die nachstehend aufgeführten (HOSMER und LEMESHOW, 1989; MENARD, 2001). D_0 ist die mit -2 multiplizierte logarithmierte Likelihood⁵ des Modells, wenn nur die Konstante, jedoch keine Sensorinformationen (als unabhängige Variablen) verwendet werden. Für ein Modell mit binär ausgeprägter abhängiger Variablen ergibt sich folgende Gleichung (MENARD, 2001):

$$D_0 = -2 \{ (n_{Y=1}) \ln [P(Y = 1)] + (n_{Y=0}) \ln [P(Y = 0)] \}. \quad (6.4)$$

$n_{Y=1}$ und $n_{Y=0}$ sind dabei die Fallanzahlen mit $Y = 1$ bzw. $Y = 0$ im Parametrisierungsdatensatz und $P(Y = 1)$ bzw. $P(Y = 0)$ die Wahrscheinlichkeiten dafür, dass Y gleich 1 bzw. 0 ist.

Unter Verwendung der unabhängigen Variablen ergibt sich ein neues Modell, dessen $-2LL$ -Ergebnis als D_M bezeichnet wird. Subtrahiert man D_M von D_0 , ergibt sich die Modellstatistik G_M , die eine Aussage über den modellverbessernden Einfluss der unabhängigen Variablen macht. Im SPSS-Ergebnisteil wird sie als Modell- χ^2 bezeichnet. Mit Hilfe dieser Angaben lässt sich das Pendant zum Bestimmtheitsmaß R^2 aus der linearen Regression R_L^2 über

$$R_L^2 = \frac{G_M}{G_M + D_M} \quad (6.5)$$

ermitteln.

Die im letzten Absatz erwähnten Maße der Anpassungsgüte (goodness of fit) haben für ein Klassifikationsmodell nur nachrangige Bedeutung, da das Hauptziel in der richtigen Vorhersage einer Merkmalsklasse besteht, die tatsächlich vorhergesagte Wahrscheinlichkeit jedoch uninteressant ist, so lange trotzdem eine richtige Zuordnung erfolgt.

6.3.4.2 Schritt 2: Koeffizienten

Auch für die Analyse der einzelnen Prädiktoren bzw. ihrer Koeffizienten liefert SPSS umfassende Informationen. Für die Modellauswahl wurden folgende Prüfgrößen berücksichtigt (nach MENARD (2001)):

- Wald-Statistik. Sie gibt Aufschluss darüber, ob der betrachtete Prädiktor das Modell verbessert. Die Nullhypothese lautet, dass der zugehörige Koeffizient nicht signifikant verschieden

⁵ das Ergebnis von Gleichung 3.10 auf Seite 18, meistens als $-2LL$ geschrieben

von Null ist.⁶

- Konfidenzintervalle für exponenzierten Koeffizienten. Schließt das Konfidenzintervall 1 ein, ist der Aussagewert des entsprechenden Prädiktors zweifelhaft.
- Veränderung der Modellgröße $-2LL$, wenn Prädiktor entfernt wird. Wenn sich durch das Weglassen des Prädiktors eine signifikante Verschlechterung der Modellanpassungsgüte ergibt, leistet diese Variable im Umkehrschluss einen nennenswerten Beitrag zur Modellverbesserung.

Für eine umfassende Betrachtung wurden neben der Festlegung der Modellvariablen durch den Nutzer (SPSS-Routine ENTER) die schrittweisen Routinen STEPWISE FORWARD und BACKWARD angewendet, um vor allem bei den multinomialen Regressionsmodellen eine Entscheidung zur Signifikanz einzelner Sensorenkanäle treffen zu können. Dabei testete SPSS selbständig, welche Variablenkombination ohne Redundanzen zu einem signifikanten Ergebnis führte.

6.3.4.3 Schritt 3: Angemessenheit des Modells

In der abschließenden Modellierungsphase wurden einige typische Fehlerquellen geprüft, um Fehlspezifikationen und numerische Probleme zu vermeiden. Mit einer „Box-Tidwell-Transformation“ genannten Umformung des Modells (HOSMER und LEMESHOW, 1989) erfolgte der Test auf Abweichungen der Modelle vom linearen Zusammenhang zwischen Prädiktorenkombination und abhängiger Variable. Dabei wurden die Modelle um Terme der Form $(x_i) \ln(x_i)$ erweitert. Ergaben sich im modifizierten Modell für diese Produkte aus Grauwert und Logarithmus naturalis des Grauwertes signifikante Koeffizienten, lag eine Verletzung der Linearitätsannahme vor.

Der Einfluss jedes einzelnen Datensatzes auf die Schätzung der Modellkoeffizienten und Ausreißer wurden über Speicherung bzw. Berechnung folgender Größen ermittelt (MENARD, 2001):

- studentisierte Residuen: Teil der SPSS-Ausgaben; kritische Fälle: $x \leq -3$ oder $x \geq 3$
- Einflussstatistik h_j : Teil der SPSS-Ausgaben; kritische Fälle: $h_j \gg \bar{h}$
- standardisierte Variante DBETA der Cook-Distanz d_{Cook} : Berechnung über

$$DBETA = \frac{d_{Cook}}{(1 - h_j)^2};$$

kritische Fälle: $DBETA \geq 1$.

Die Auswertung der obigen Indikatoren sollte Modellverbesserungen ermöglichen, denen in der Praxis allerdings Grenzen gesetzt waren. Fälle, die sich als besonders einflussreich erwiesen und zudem falsch klassifiziert waren, wurden im GIS mit hinterlegtem Satellitenbild einer eingehenden Betrachtung unterzogen. Mitunter waren Lagefehler für die Fehlklassifikation verantwortlich, in diesen Fällen deuteten die BWI-Informationen zwar auf Wald, im Satellitenbild lagen die Punkte jedoch im Nichtwald. Solche eindeutigen Fälle wurden aus dem Datenkollektiv entfernt und das

⁶ Die Wald-Teststatistik liefert bei großen Koeffizienten, wie es für die Modellkonstante in dieser Arbeit regelmäßig der Fall war, vergrößerte Standardfehler. Es resultiert die Annahme der Nullhypothese $h_0: b_0 = 0$, obwohl selbige u. U. verworfen werden müsste.

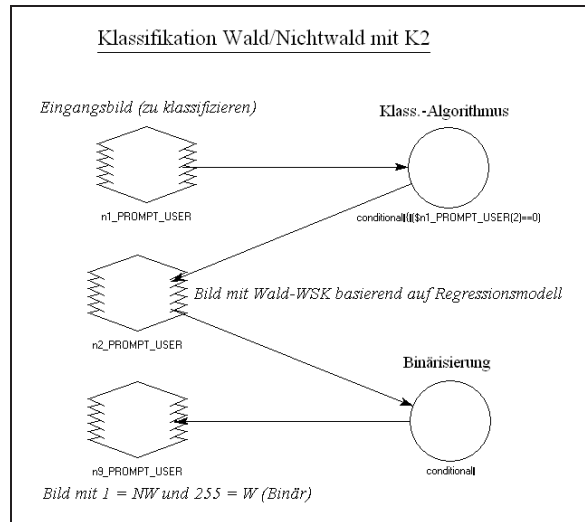


Abbildung 6.1
Der Spatial Modeler von Erdas zur Kartierung der Waldfläche.

Die Ausgangsszene wird im ersten Schritt in Wahrscheinlichkeiten für das Auftreten eines Waldpixels klassifiziert. Dieses Graustufenbild wird abschließend rekodiert, so dass alle Wahrscheinlichkeiten $\leq 0,5$ als Nichtwald (Grauwert 1), alle darüber als Wald (Grauwert 255) klassifiziert werden. Der Szenenbereich ohne Daten wird mit dem Grauwert 0 versehen.

Modell neu angepasst. Bei weniger eindeutigen Situationen wurde keine Entfernung der Datensätze vorgenommen. Außerdem muss auch hier darauf hingewiesen werden, dass bei den Modellen die Klassifizierungsgenauigkeit einen deutlich höheren Stellenwert hatte als die exakte Anpassungsgüte, weshalb den studentisierten Residuen nur in extremen Fällen Gewicht beigemessen wurde.

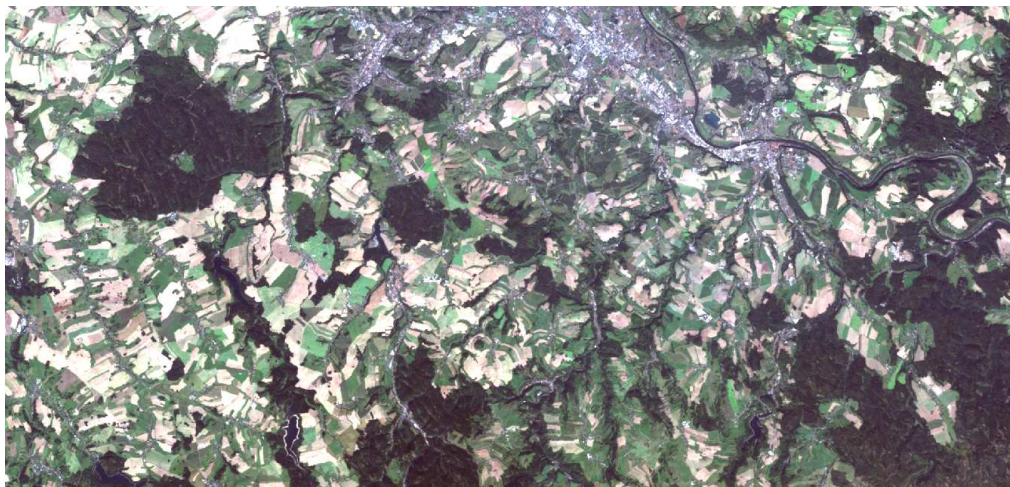
6.3.5 Kartenerstellung

Nach der Aufstellung der Modelle wurde der Spatial Modeler von Erdas genutzt, um aus den Originalszenen klassifizierte Karten zu gewinnen (siehe Abbildung 6.1). Zunächst wird in Stufe I festgelegt, dass der Bildbereich ohne Informationen mit 0 codiert wird. Dann werden Wasserflächen mit zwei einfachen Schwellenwerten im Band 5 der Originalszenen entfernt bzw. automatisch auf den Status *Nichtwald* mit dem Grauwert 1 gesetzt. Im weiteren Verlauf wird mit dem zuvor ermittelten LogReg-Modell in Abhängigkeit des Grauwertes von Band 2 errechnet, wie groß die Wahrscheinlichkeit (WSK) eines Pixels ist, der Klasse *Wald* anzugehören. In Stufe II wird die Wald-Wahrscheinlichkeitskarte in eine binäre Waldkarte überführt.

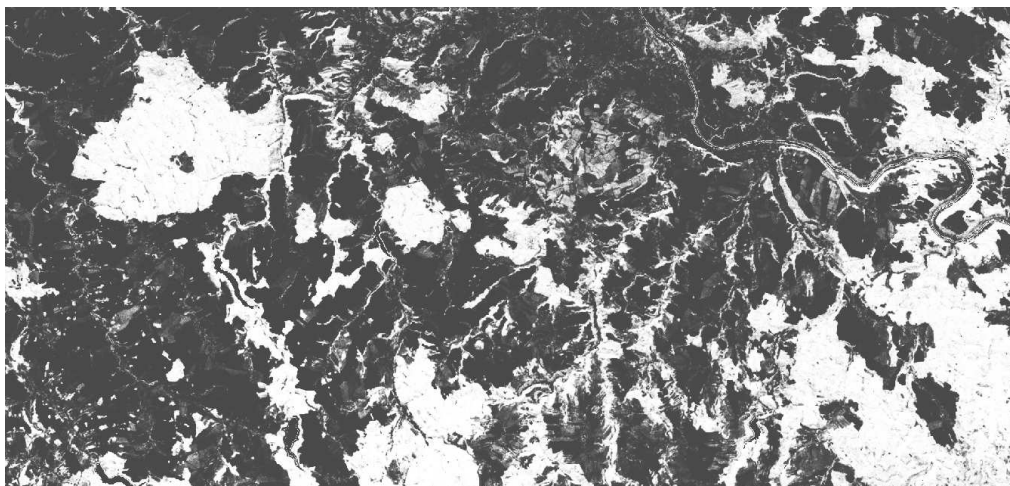
Auf Pixelbasis beruhende Klassifikationsalgorithmen ergeben oftmals eine Karte mit ausgeprägtem Salz-und-Pfeffer-Muster (LILLESAND ET AL., 2004). Zur Glättung dieses Effektes wurde eine Medianfilterung der Karte mit einem 3×3 Pixel großen Fenster bzw. Kernel vorgenommen. Der Filter ordnet alle im 3×3 -Pixel-Fenster vorkommenden Grauwerte der Größe nach und weist dem zu filternden Zentralpixel jenen Wert zu, der sich im Zentrum dieser Verteilung befindet. Das zentrale Pixel des Filterkerns wird im gegebenen Falle also immer jener Klasse zugeordnet, die am häufigsten im Filterkernel auftaucht (LEICA GEOSYSTEMS, 2003).⁷

Abbildung 6.2 verdeutlicht die Ausführungen zur Kartenerstellung. Die Ausschnitte zeigen einen Teil des Testgebiets 3 in Sachsen. Im linken oberen Teil befindet sich der Tharandter Wald, die Stadt östlich davon ist Dresden. Insgesamt wird eine Fläche von ca. 50×25 km dargestellt.

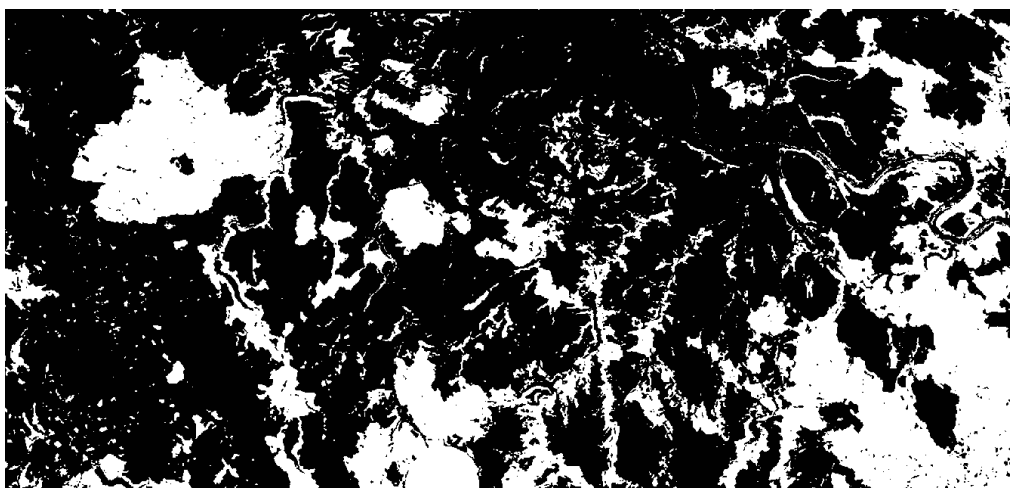
⁷ Nach LILLESAND ET AL. (2004) basiert die Kalkulation dabei immer auf den Originalwerten der umgebenden Pixel vor der Änderung durch den Filter.



(a) Originalszene: ETM⁺, multispektral, Echtfarbenabdarstellung



(b) Waldkarte I: Darstellung der Wald-Wahrscheinlichkeit (WWSK), hellere Töne bedeuten höhere WWSK



(c) Waldkarte II: Binärdarstellung, weiß: Waldpixel, schwarz: Nichtwaldpixel

Abbildung 6.2: Entstehung einer Waldkarte über LogReg-Klassifizierung einer ETM⁺-Szene

6.3.6 Validierung

6.3.6.1 Regressionsmodelle

Zur Überprüfung der numerischen Stabilität der LogReg-Modelle wurde das Verfahren der Kreuzvalidierung (BENGIO und GRANDVALET, 2004; BORTZ, 2005) angewendet. Dabei erfolgt eine Teilung der ursprünglichen Datenkollektive in k etwa gleich große, nicht überlappende Blöcke, so dass kein Fall in zwei Blöcken gleichzeitig auftaucht. Die als *k-fold cross validation* bezeichnete Methode verwendet jedes der k Teilkollektive einmal als Validierungsdatensatz, um das mit den verbliebenen 9 Teilkollektiven erstellte Modell zu validieren. Die Festlegung der Blockanzahl k entscheidet automatisch über die Anzahl Fälle m pro Block, da die folgende Beziehung gilt (BENGIO und GRANDVALET, 2004):

$$m = n/k.$$

n —die Anzahl Datensätze—ist fix, da durch die BWI vorgegeben. Je größer k gewählt wird, umso kleiner wird die Anzahl Fälle pro Block. Um in allen validierten Testgebieten aussagekräftige Ergebnisse zu erzielen, wurde daher $k = 10$ festgelegt, wodurch eine hinreichende Anzahl Fälle pro Block gewährleistet wurde.

Die Betrachtung der 10 Klassifizierungsergebnisse gab Aufschluss über die Stabilität der Modellparametrisierungen. Die zur Kartenerstellung verwendeten LogReg-Modelle basierten allerdings auf den gesamten zur Verfügung stehenden Datensätzen, um keine Informationen ungenutzt zu lassen.

6.3.6.2 Karten

„Another area that is continuing to receive increased attention by remote sensing specialists is that of classification accuracy assessment. Historically, the ability to produce digital land cover classifications far exceeded the ability to meaningfully quantify their accuracy. In fact, this problem sometimes precluded the application of automated land cover classification techniques even when their cost compared favorably with more traditional means of data collection. The lesson to be learned here is embodied in the expression *A classification is not complete until its accuracy is assessed.*“

LILLESAND ET AL. (2004)

Wie bei allen statistischen Auswertungen ist es auch für eine Klassifizierung notwendig, Angaben zur Genauigkeit bzw. Verlässlichkeit der Ergebnisse zu liefern, damit sich der Nutzer im Kontext seiner Problemstellung über die Verwendbarkeit der Daten klar werden kann. Das obige Zitat betrifft zwar vornehmlich Karten auf der Grundlage von automatischen Klassifizierungsalgorithmen, behält jedoch für den hier vorliegenden Fall ebenso seine Gültigkeit.

Nach JENSEN (2005) werden die Referenzdaten (also die ground truth) idealerweise im Rahmen terrestrischer Stichproben gewonnen. Steht diese Option nicht zur Verfügung, müssen höher aufgelöste FE-Daten als Substitut herangezogen werden. In diesem Sinne wurden für drei der fünf Testgebiete—namentlich für die „westliche“ und „östliche Mittelgebirgsschwelle“ sowie den „Pfälzerwald“—Luftbilder beschafft, die sich für die Validierung der Waldkarten eignen.

Für die Genauigkeitsbewertung ist die Festlegung von Stichprobenpunkten unerlässlich, an denen ein Vergleich der Klassifikation mit den Referenzdaten erfolgt. JENSEN (2005) empfiehlt

in Anlehnung an FITZPATRICK-LINS (1981) die folgende Formel zur Berechnung des minimal notwendigen Stichprobenumfangs:

$$n = \frac{Z^2 pq}{E^2} \quad (6.6)$$

mit

n	Anzahl Stichprobenelemente zur Validierung
p	erwartete Kartengenauigkeit in Prozent
q	$100 - p$
E	vertretbarer Fehler bei der Genauigkeitsbewertung
Z	= 2 bei einer Irrtumswahrscheinlichkeit von 5 % ausgehend von einer Standardnormalverteilung.

Bei einer angenommenen Kartengenauigkeit von 85 %⁸ und einem tolerierbaren Fehler von 5 % ergibt sich ein Stichprobenumfang von

$$n = \frac{2^2 \times 85 \times 15}{5^2} = 204, \quad (6.7)$$

der durch die Szenengröße nur einmal unterschritten wurde: im Norden des Pfälzerwaldes. Die Südszene deckt jedoch das fünfte Testgebiet zu über 95 % ab. Ausgehend von diesen Überlegungen wurde ein systematisches Gitternetz mit einem Punktabstand von 3×3 km entworfen, dessen Ursprung ausgehend vom jeweils nordwestlichsten Pixel der Testgebiete um 250 m nach Osten und Süden verschoben wurde (siehe Abbildung 6.3 auf der nächsten Seite). Damit sollte ein Zusammenfallen der Validierungsstichprobenpunkte mit jenen der BWI verhindert werden. Die Punkte lieferten die Mittelpunktkoordinaten für 25×25 m große Polygone (ein Landsat-Pixel), in denen die Waldfläche des Luftbildes betrachtet wurde. Als richtig klassifiziert wurden alle Waldpixel eingestuft, bei denen deutlich mehr als die Hälfte des zugehörigen Polygons mit Wald bedeckt war. Umgekehrt wurden alle Nichtwaldpixel für richtig befunden, bei denen eindeutig weniger als 50 % Waldbedeckung im Luftbild zu verzeichnen war. In Situationen, in denen ein Polygon etwa zur Hälfte mit Wald bedeckt war und damit eine gesicherte Entscheidung nicht zu treffen war, wurde das entsprechende Waldmaskenpixel als *unsicher* eingestuft. Aus dem Vergleich der Vorhersagen des Modells und der Informationen des Luftbildes erfolgte anschließend die Aufstellung einer Fehlermatrix. Die Aussagen der Fehlermatrix sind als konservativ zu bezeichnen, da alle unsicheren Pixel im Sinne eines Worst-Case-Szenariums als falsch klassifiziert eingestuft wurden.

6.3.7 Verfahrenstest ohne terrestrische Daten

Anhand der beiden ETM⁺-Szenen des Testgebietes „Westliche Mittelgebirgsschwelle“ wurde neben der oben beschriebenen Methode beispielhaft untersucht, wie gut sich das Waldvorkommen unter Verzicht auf terrestrische Daten klassifizieren bzw. kartieren lässt. Die Entscheidung für dieses TG fiel aufgrund der Schwierigkeiten bei der univariaten Klassifizierung der sehr zeitig aufgenommenen Nordszene (siehe Abschnitt 5.2.2 auf Seite 34). Ein Vergleich sollte zeigen, ob eine größere Anzahl visuell ausgewählter Trainingspixel und die Verwendung mehrerer Kanäle eine bessere Trennung der Feldflächen vom Wald ermöglichen würden. Im Bildverarbeitungsprogramm

⁸ ausgehend von den ungenauesten Klassifizierungsergebnissen der LogReg-Modelle

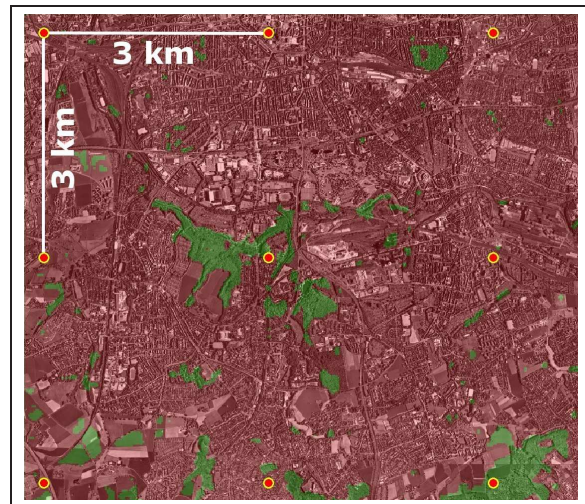


Abbildung 6.3

3 × 3-km-Netz zur Validierung der Waldmaske.

Die Abbildung zeigt das Validierungsnetz im nordwestlichsten Ausschnitt des ersten Testgebietes. Über dem Luftbild ist die halbtransparente Waldmaske eingefügt; rötliche Bereiche bezeichnen Nichtwaldpixel, grüne entsprechend Waldpixel.

wurden dabei Trainingsgebiete für die Klassen *Nichtwald* und *Wald* festgelegt. Um dem logistischen Regressionsmodell eine breite Datenbasis zur Verfügung zu stellen, wurde darauf geachtet, unterschiedliche Waldstrukturen und verschiedenste Nichtwaldflächen zu erfassen. Dazu gehörten insbesondere auch Gewässer, da dieser Ansatz ohne den zusätzlichen Schritt des Ausschlusses von Wasserflächen auskommen sollte. Unterstützend wurde ein zweiter, räumlich mit dem ETM⁺-Fenster verbundener Erdas-Bildbetrachter mit dem Orthophoto des TG1 geöffnet, um Fehler bei der Spezifikation der Trainingsgebiete zu minimieren. Das Verfahren wäre allerdings auch ohne diese zusätzliche Kontrollmöglichkeit anwendbar.

Aus den Trainingsgebieten wurden die Grauwerte aller Kanäle (inklusive NDVI) extrahiert und in einer Tabelle zusammengeführt. Da dabei keine Bindung an ein Stichprobennetz gegeben war, konnten zehntausende Punkte Eingang in die Regressionsanalysen finden. Mit Hilfe der ermittelten Modellgleichung—diesmal unter Einbeziehung des kompletten Prädiktorensatzes—wurde abermals eine Waldkarte erstellt, die mit den BWI-Informationen verglichen wurde.

6.3.8 Zum Einfluss der Testgebietswahl

Versucht man, einen Zusammenhang zwischen den Genauigkeitsangaben von Satelliten-basierten Klassifikationen und den dabei untersuchten Gebietsgrößen herzustellen, fällt auf, dass extrem hohe Genauigkeiten von über 90–95 % häufig mit Testgebietsflächen von wenigen Quadratkilometern assoziiert sind (vgl. IVITS und KOCH (2003); KOCH ET AL. (2003); CLARK ET AL. (2004); LALIBERTE ET AL. (2004); PUISSANT ET AL. (2005) u. a.). Größere Gebiete bis hin zu ganzen Ländern werden überwiegend mit geringeren Genauigkeiten um 85 % oder weniger klassifiziert (WYNNE ET AL., 2000; RACK, 2001; BLINN, 2005). Ausnahmen gibt es: SAGISCHEWSKI ET AL. (2003) konnten auf der Fläche zweier Thüringer Forstämter Wald objektbasiert mit über 94 % Gesamtgenauigkeit kartieren; ebenso war es in der Schweiz möglich, pixelbasiert die gesamte Waldfläche des Landes mit fast 92%iger Übereinstimmung zwischen Klassifikation und Referenz zu kartieren (BUNDESAMT FÜR STATISTIK, SEKTION GEOINFORMATION, 2004).

Größe und Lage eines Testgebiets haben demnach einen maßgeblichen Einfluss auf das Klassifizierungsergebnis, und nicht immer lassen sich Erkenntnisse aus Pilotstudien oder Kleinstgebietana-

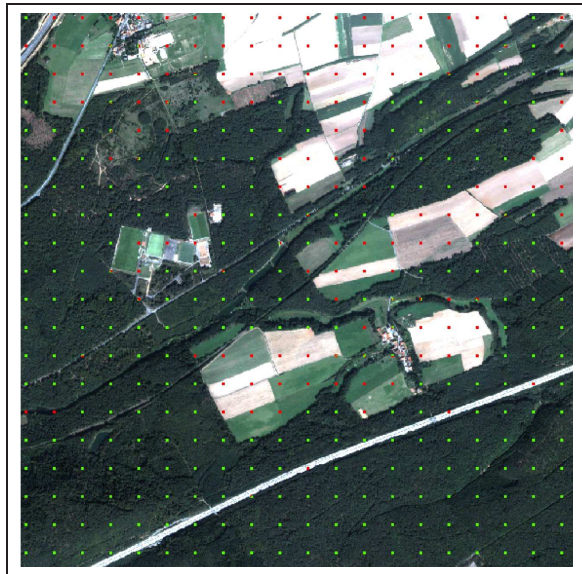


Abbildung 6.4
Kleinsttestgebiet mit Validierungsnetz.

Die Abbildung zeigt das 3×3 km große Gebiet im Norden der QuickBird-Szene. Ebenfalls dargestellt sind die 225 m^2 großen Validierungspolygone im Abstand von 150 m.

lysen übertragen, um zu operationellen Verfahren für landesweite Kartierungen zu gelangen. Zur exemplarischen Quantifizierung der Wirkung eines sehr kleinen Testgebietes auf das Kartierungsergebnis wurde im Norden der im Pfälzerwald gelegenen QuickBird-Szene (TG5) eine 3×3 km große Fläche ausgewählt. Diese besteht zu rund zwei Dritteln aus Waldbeständen, die eine geringe Fragmentierung aufweisen (siehe Abbildung 6.4).

In dieses Gebiet fallen nur zwei BWI-Trakte mit je 4 Plots, was diese Informationsquelle zur Gewinnung von Trainingsdaten ausscheiden ließ. Die hohe geometrische Auflösung von 3 m gestattete es jedoch, zuverlässig große Wald- und Nichtwaldbereiche im Satellitenbild für die Grauwertextraktion zu delinieren, so dass anschließend insgesamt ~ 227.000 Pixel für die Klasse *Nichtwald* und ~ 497.000 Pixel für *Wald* zur Verfügung standen. Gemessen an der Gesamtpixelzahl von $\sim 1.000.000$ wurden somit fast 75 % der Fläche in den Trainingsgebieten erfasst.

Das Vorgehen bei der Modellierung und Auswertung folgte der oben dargelegten Methodik, die auch für die anderen Testgebiete angewendet wurde. Die Parametrisierung der Regressionsmodelle wurde mit zwei Datensätzen durchgeführt. Einmal gingen alle Trainingsdaten in den Anpassungsprozess ein, für einen zweiten Lauf wurden 2×21.000 per Zufallsauswahl gewählte Fälle (also nur noch ca. 4 % der Gesamtfläche) genutzt. Anschließend konnten mit dem Erdas Modeler die Waldmasken erstellt werden.

Zur Validierung der Klassifikation war der panchromatische QuickBird-Kanal geeignet. Dessen Qualität (und Auflösung) überstieg jene der für das Gebiet vorliegenden Orthophotos (keine Artefakte, Rauschen, Verzerrungen), außerdem konnten so Ungenauigkeiten durch Lageabweichungen ausgeschlossen werden. Es wurde ein Quadratraster generiert, auf dem alle 150 m ein 15×15 m großes Testpolygon zu liegen kam. In der Summe wurden 400 solche Polygone überprüft, was bei

$$15 \text{ m} \times 15 \text{ m} = 225 \text{ m}^2$$

Polygongröße in

$$400 \times 225 \text{ m}^2 = 90.000 \text{ m}^2$$

Validierungsfläche (1 % der Testgebietsgröße) resultierte. Für jedes dieser Polygone wurde der Wald im panchromatischen Bild deliniert; anschließend Testpolygone und erfasste Waldpolygone fusioniert. Damit stand ein Vektordatenatz für die Validierung zur Verfügung, der alle benötigten Informationen zu Wald und Nichtwald enthielt. Selbiger wurde mit den beiden vektorisierten Waldmasken verschnitten, um einen Vergleich zwischen Referenz und Klassifikation zu ermöglichen. Die Ergebnisse werden auf zwei Arten präsentiert:

1. *Flächenhafte Überprüfung*: In jedem Testpolygon werden tatsächliche und klassifizierte Wald- und Nichtwaldfläche verglichen.
2. *Punktüberprüfung per Mehrheitsentscheid*: Für jedes Testpolygon wurde jene Klasse ermittelt, die mehr als 50 % Fläche einnimmt. Dem gesamten Testpolygon wird im Folgenden der Status der dominierenden Klasse zugewiesen. Auch für die Waldmasken wird diese Klasse im Bereich des Testpolygons ermittelt und dann mit der Referenz verglichen.

6.4 Logistische Regression zur Kartierung von Waldmischungen

Ein weiteres nominales Merkmal, dessen Klassifizierbarkeit bewertet wurde, war die Baumartenmischung. Über die thematische Auflösung *Nadelwald—Mischwald—Laubwald* hinaus erfolgten jedoch keine Untersuchungen, da bereits auf dieser Ebene die Grenzen des Verfahrens erreicht waren. Als *Nadelwald* wurden alle Bestände mit einem Nadelbaumanteil von mindestens 75 % eingestuft; gleiches gilt in entsprechender Weise für *Laubwald*. Alle anderen Plots, deren Mischanteile zwischen diesen beiden Bereichen lagen, gingen als *Mischwald* in die Berechnungen ein. Sämtliche Regressionen wurden ausschließlich mit BWI-II- und ETM⁺-Daten durchgeführt.

Die Vorgehensweise zur Klassifizierung des Merkmals *Mischung* ähnelt in weiten Teilen der des vorhergehenden Abschnittes. Auch hier wurde zur Vorbereitung der Regressionsberechnungen eine Reihe von graphischen Darstellungen zusammengestellt, um die spektrale Trennbarkeit der drei Mischungsklassen einschätzen zu können.

Im Gegensatz zum *Waldstatus* wurde jedoch auch eine andere Regressionsroutine verwendet. ITTEN ET AL. (1992) berichten von zwei Verfahrensalternativen zur Separierung der Mischungsklassen:

1. Zweistufige Klassifizierung durch (i) Unterteilung in *Nadelwald* und *Nichtnadelwald* mittels Kanal 4 und Kantenschärfung sowie anschließend (ii) innerhalb von *Nichtnadelwald* eine Trennung von *Laub-* und *Mischwald* mittels Subtraktion des Kanals 2 von Kanal 4.

ODER

2. Einstufige Klassifizierung mit den Informationen des Kanals 4 allein.

Inspiziert durch diese Empfehlungen wurden zwei Lösungswege beschritten. Zum einen erfolgte die Anpassung eines multinomialen LogReg-Modells, welches gleichzeitig alle drei Klassen als Ausprägungen der abhängigen Variablen *Mischung* verarbeiten konnte. Der zweite Weg bestand in der getrennten Anpassung zweier binomialer LogReg-Modelle; in der ersten Stufe zur Abtren-

nung von *Nadelwald*, in der zweiten Stufe innerhalb des *Nichtnadelwaldes* zur Unterscheidung der verbleibenden zwei Klassen.

In Analogie zur Vorgehensweise bei der QuickBird-Szene wurde der Umfang jeder Klasse über Zufallsstichproben an die Fallanzahl der am schwächsten mit Daten besetzten Klasse angepasst.⁹ Diese Unterteilung in nur drei Merkmalsausprägungen führte zu dem oft erwähnten Problem geringer Fallanzahlen innerhalb mancher Klassen, das eine stabile Modellanpassung erschwert. Versuche mit den ursprünglichen Klassenstärken endeten dafür beinahe immer damit, dass der Mischwald entweder überhaupt nicht erkannt wurde, also den anderen beiden Klassen zugeschlagen wurde oder aber nur vereinzelte Mischwaldpixel richtig identifiziert wurden. Vor diesem Hintergrund offenbart sich ein Nachteil des LogReg-Verfahrens: Die zur Verfügung stehenden Informationen werden ineffizient genutzt, wenn man sich zumindest näherungsweise auf vergleichbare Klassengrößen beschränken muss, um Modellverzerrungen zu vermeiden.

Bedingt durch die niedrigen Klassifikationsgenauigkeiten erscheint eine Anwendung des Verfahrens für eine künftige Bundeswaldinventur ungeeignet. Die Erfahrung mit der Waldkartierung im vorhergehenden Abschnitt zeigte, dass die Kartvalidierung stets schlechtere Ergebnisse erbrachte als es die reine Klassifizierungsleistung des logistischen Modells nahe legte. Da bereits letztere bei der Mischungskartierung sehr heterogen war und insgesamt auf niedrigem Niveau lag, wurden Validierungsbemühungen über die Standardausgaben von SPSS hinaus ebenso wenig unternommen wie eine Umsetzung der Modelle in Karten.

6.5 Logistische Regression zur Kartierung von Totholzvorkommen

Im Gegensatz zur automatisierten Waldflächenerfassung aus Fernerkundungsdaten ist die Kartierung von Totholzvorkommen ein weniger intensiv bearbeitetes Forschungsgebiet. Während sich Quellen zur Verwendung von Farbinfrarotphotos für diese Aufgabe finden lassen, mangelt es an Untersuchungen zur *automatischen* Klassifizierung von Totholzvorkommen mit Satellitendaten.

FUCHS (2003) verwendet den Differenzkanal [Rot]–[NIR] zur Klassifizierung von Farbinfrarot-Luftbildern, mit dessen Hilfe er über Schwellenwerte Totholz identifizierte. Der gleichen Informationsgrundlage bedienten sich auch FREI ET AL. (2003). Sie stellten eine Nichtklassifizierbarkeit von liegendem Totholz in Luftbildern mit 1 m Auflösung fest. Erst die Nutzung von 20-cm-Luftbildern offerierte die Möglichkeit zur Kartierung.

Die einzigen Hinweise auf Landsat-basierte Klassifizierungsverfahren finden sich in COENRADIE (1991) und STÜMER (2004). COENRADIE kartierte sowohl stehendes als auch liegendes Totholz im Bayerischen Wald mit Hilfe von Schwellenwerten in unterschiedlichen Infrarot- und Differenzkanälen des TM-Sensors, während STÜMER ausschließlich stehendes Totholz mit einem *k*NN-Verfahren erfasste. Die erreichten Genauigkeiten bewegten sich zwischen 60 und 74 %, allerdings ist die *k*NN-basierte Methode nicht auf einen LogReg-Algorithmus übertragbar.

⁹ Meistens war dies die Klasse *Mischwald*.

6.5.1 Zusammenstellung der Datensätze

Die BWI-Datenbank enthält eine eigenständige Tabelle zu aufgenommenen Totholzfunden, in der neben Baumartengruppe und Art des abgestorbenen Materials (stehend oder liegend) auch der Zersetzungsgrad sowie die Maße der Stammstücken/Stubben erfasst sind. Über eine Gruppierungsfunktion wurde sichergestellt, dass jede Traktecke mit Totholz nur einmal in die Berechnungen einging, selbst wenn es mehrere Totholzfunde auf diesem Plot gab. Dadurch reduzierte sich die Anzahl von fast 3200 Totholzfunden auf ca. 1700 Ecken mit dem Status *Totholz vorhanden*. Zur Sicherstellung ausgewogener Schätzungen wurden auch bei diesem Merkmal gleiche Klassenstärken verwendet; dazu mussten wiederum per Zufallsstichproben aus den größeren Kollektiven ohne Totholzfunde die passende Anzahl Traktecken gezogen werden.

6.5.2 Uni- und multivariate Analysen, Validierung

Voruntersuchungen mit dem oben erwähnten Differenzkanal Rot–NIR in einem univariaten Modell ließen keinen straffen Zusammenhang erkennen; daher wurden anschließend alle Regressionsgleichungen mit sämtlichen Landsat-Kanälen sowie dem NDVI durchgeführt. Die Aufteilung der 1700 Traktecken durch das szenenweise Auswerten resultierte wie beim Merkmal *Mischung* mitunter in sehr kleinen Stichprobenumfängen. Die Stabilitätstests in Form von Kreuzvalidierungen ähnlich wie beim Merkmal *Waldstatus* kamen aus diesem Grund nur für ein Gesamtmodell zum Einsatz, welches auf allen Totholz-Ecken und einer gleichen Anzahl Traktecken ohne Totholzfunde basierte.

Kapitel 7

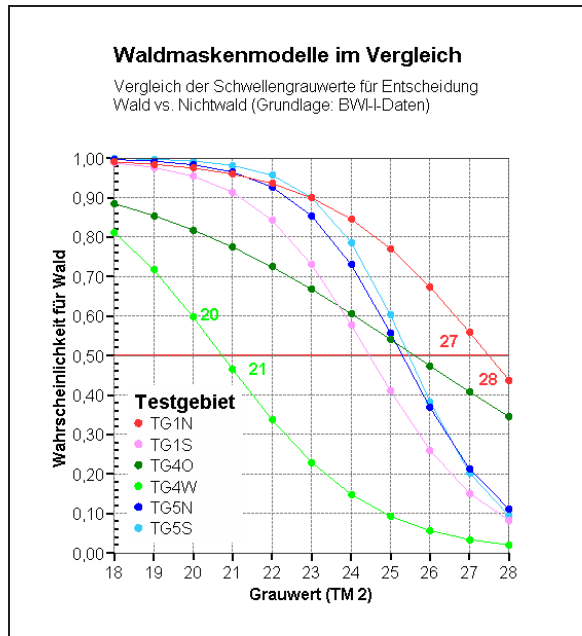
Ergebnisse

Die Darstellung der Ergebnisse erfolgt gegliedert nach untersuchten Klassifizierungsmerkmalen; innerhalb dieser nach Fernerkundungssystem und Testgebieten (soweit erforderlich). Zu beachten ist, dass die BWI-Datenbank für die erste BWI keine Daten aus den neuen Bundesländern vorhält. Betrachtungen zu Landsat 5 (Aufnahmezeitpunkt um 1990) beziehen sich also nur auf die in den alten Bundesländern gelegenen Testgebiete „Westliche Mittelgebirgsschwelle“ (TG1), „Alpen“ (TG4) und „Pfälzerwald“ (TG5). Um die Nachvollziehbarkeit der Untersuchungsergebnisse zu verbessern, ist deren Diskussion an entsprechender Stelle im Ergebnisteil eingebettet, anstatt in einem eigenen Kapitel niedergelegt zu sein.

7.1 Merkmal *Waldstatus*

Es wurde für das Merkmal *Waldstatus* ein möglichst einheitliches Klassifizierungsverfahren gesucht, um die Anwendung in einem etwaigen bundesweiten Einsatz zu vereinfachen. Diese Maxime galt für alle Testgebiete und beide Landsat-Versionen. Trotz des Strebens nach einem einheitlichen Modell ist als Fazit der Modellierung festzuhalten, dass nicht einmal einheitliche Modelle für beide Szenen *eines Testgebietes* verwendet werden konnten—zu groß waren die Unterschiede. Eine testgebietsübergreifende Klassifikationsgleichung zu verwenden war unter diesen Umständen erst recht unpraktikabel. Diese beiden Schlussfolgerungen werden in Abbildung 7.1 auf der nächsten Seite verdeutlicht. Das Diagramm zeigt die Waldwahrscheinlichkeit eines Landsat-5-Pixels in Abhängigkeit vom Grauwert des Kanals 2 für die Testgebiete 1, 3 und 5. Entscheidend ist der Grauwert, bei dem die Wahrscheinlichkeitsschwelle von 0,5 unterschritten und das Pixel nicht mehr als Wald klassifiziert wird. Die größten Unterschiede bestehen zwischen TG1N und TG4W; aber auch innerhalb eines Testgebietes treten deutliche Unterschiede zwischen den Szenen auf. Eine Ausnahme bildet der Pfälzerwald mit sehr ähnlichem Kurvenverlauf beider Szenen, was auf die homogenen Waldstrukturen zurückzuführen sein dürfte.

ITTEN ET AL. (1992) empfehlen für die Klassifikation von Waldflächen das TM-Band 2 (Grün). Dieser Empfehlung spricht nichts entgegen, wie die Modellvergleiche ergaben. Zwar lassen sich unter Einbeziehung weiterer oder aller Bänder als Prädiktoren des logistischen Regressionsmodells bessere Klassifikationsergebnisse erzielen, jedoch steht im Falle der ETM⁺-Aufnahmen in 7 von 10 Szenen der Genauigkeitsgewinn von weniger als 1 % in keinem Verhältnis zum gesteigerten

**Abbildung 7.1**

Modellvergleich aller TM-Szenen für das Merkmal *Waldstatus*. Erläuterungen im Text.

Die Verbindungslinien sollen lediglich die Übersichtlichkeit erhöhen, nicht jedoch implizieren, dass es nichtganzzahlige Grauwerte gibt.

Aufwand bei der Implementierung der Modelle in Erdas. Zudem führt die Verwendung aller Kanäle zu Redundanzen und damit zu insignifikanten Koeffizienten-Schätzern.

Für drei Szenen ergaben sich deutlichere Unterschiede. Die sehr zeitig im Mai aufgenommene Nordszene des ersten Testgebietes konnte mit den zusätzlichen Informationen der anderen Kanäle genauer klassifiziert werden. Vor allem die schwer vom Wald trennbaren Wiesenflächen wurden zahlreicher der richtigen Klasse *Nichtwald* zugeordnet.

Vergleichbar sind die Ergebnisse der TM-Szenen zu interpretieren. Bis auf die Nordszene des ersten TG, die ähnlich zeitig aufgenommen wurde wie ihr ETM⁺-Pendant, resultiert die Einbindung weiterer Landsat-Bänder entweder nur in einer Steigerung der Gesamtgenauigkeit von < 1 % oder sogar in einer Verringerung um 0,3 bzw. 0,1 % im Pfälzerwald (Nord- bzw. Südszene).

Die Validierung der Waldkarten durch den Vergleich mit Orthophotos wurde nur für die ETM⁺-Szenen der Testgebiete 1, 3 und 5 vorgenommen. Für die anderen beiden TGs bzw. Landsat 5 standen keine Luftbilder geeigneten Aufnahmedatums zur Verfügung. Die überwiegende Zahl der Untersuchungen wurde allerdings ohnehin an den ETM⁺-Szenen durchgeführt; die Ergebnisse der TM-Szenen werden der Vollständigkeit halber dargestellt. Aus diesem Grund beschränkt sich auch die Prüfung der Stabilität per Kreuzvalidierung auf die Modelle für Landsat 7.

7.1.1 Landsat TM

7.1.1.1 Regressionsmodelle

Zum besseren Überblick sind die Klassifizierungsmodelle für alle sechs Landsat-5-Szenen in Tabelle 7.1 auf der nächsten Seite zusammengefasst. Zur Interpretation der Werte sei hier die Modellgleichung 3.1 von Seite 16 nochmals dargestellt:

$$P(Y = 1|GW) = \pi(x) = \frac{e^{\beta X}}{1 + e^{\beta X}}. \quad (3.1)$$

Tabelle 7.1: Koeffizienten der Modelle für Merkmal *Waldstatus* (TM-Daten)

Szene	b_0	b_{k_2}	Szene	b_0	b_{k_2}
Westl. M'schwelle Nord	13,402	-0,488	Westl. M'schwelle Süd	16,643	-0,680
Alpen Ost	15,685	-0,641	Alpen West	11,097	-0,535
Pfälzerwald Nord	19,444	-0,769	Pfälzerwald Süd	22,708	-0,892

b_0 = Modellkonstante, b_{k_2} = Koeffizient für Landsat-Kanal 2

Setzt man zum Beispiel die Werte für den Norden der westlichen Mittelgebirgsschwelle in die Gleichung 3.1 ein und vereinfacht, so resultiert

$$P(Y = \text{Wald}|k_2) = \frac{1}{1 + e^{-(13,402 - 0,488k_2)}}. \quad (7.1)$$

In den Spatial Modeler von Erdas eingefügt, wurden auf diese Weise für alle drei Testgebiete Waldkarten erstellt.

Durch die Anwendung der dargestellten Modelle auf die Parametrisierungsdaten ergaben sich die in Tabelle 7.2 auf der nächsten Seite aufgezeigten Klassifikationsgenauigkeiten. Ebenfalls in dieser Tabelle enthalten sind die Werte für die Vorhersageeffizienz τ_p jedes Modells. Im ersten Testgebiet konnten mit dem beschriebenen 1-Kanal-Modell die wenigsten Trainingspixel korrekt den beiden Klassen zugewiesen werden. Besonders die Nordszene erwies sich wegen des frühen Aufnahmezeitpunktes (Mai) als schwierig zu klassifizieren. Dies zeigt sich an der mit 76,4 % geringsten Erstlergenauigkeit für die Klasse *Nichtwald*. Entsprechend ergab sich aus der Klassifizierung auch der niedrigste Wert für die Vorhersageeffizienz ($\tau_p = 0,704$). Die weniger fragmentierten Testgebiete in den Alpen und im Pfälzerwald ließen sich mit höherer Genauigkeit klassifizieren. Die wenigsten Fehlzuordnungen—gemessen an den Trainingsdaten—erfolgten durch das Modell für die Nordszene des Pfälzerwaldes ($\tau_p = 0,886$).

Wie angedeutet konnte in Ermangelung geeigneter Orthophotos keine Validierung der Waldmaskenqualität vorgenommen werden. Selbst ein summarischer Vergleich der Waldflächen laut Waldmasken mit jenen aus der BWI schlug fehl, da die eigens für das Projekt angelegte testgebietspezifische Datenbank für den Zeitraum der BWI I Fehler enthielt.

7.1.2 Landsat ETM⁺

7.1.2.1 Regressionsmodelle

Die durch die ETM⁺-Aufnahmen repräsentierte BWI II hält in der Datenbank Informationen für alle Bundesländer vor. Die Ergebnisse erweitern sich hier also um die im Osten gelegenen Testgebiete 2 (Nordostdeutsches Tiefland) und 3 (Östliche Mittelgebirgsschwelle); außerdem ermöglichten die Luftbilder eine über die reine Modellbewertung hinausgehende Validierung der resultierenden Waldmasken. Zunächst sei jedoch in Tabelle 7.3 auf der nächsten Seite eine Übersicht der Modellkoeffizienten gegeben. Auch für diese Modelle wurden in SPSS anhand der Eingangsdatensammlungen Kontingenztafeln mit den Klassifizierungsergebnissen erstellt, die in Tabelle 7.4 auf Seite 61 wiedergegeben sind.

7 Ergebnisse

Tabelle 7.2: Klassifikationsergebnisse der Modelle für Merkmal *Waldstatus* (TM-Daten)

Beobachtung ^b	Vorhersage ^a			% korr. (E)	Beobachtung ^b	Vorhersage ^a			% korr. (E)
	Waldstatus					Waldstatus			
	0	1			0	1			
Testgebiet 1 Nord ($\tau_p = 0,704$)				Testgebiet 1 Süd ($\tau_p = 0,750$)					
Waldstatus	0	396	122	76,4	Waldstatus	0	363	58	86,2
	1	59	690	92,1		1	54	423	88,7
% korr. (N)		87,0	85,0	(G) 85,7	% korr. (N)		87,1	87,9	(G) 87,5
Testgebiet 4 Ost ($\tau_p = 0,791$)				Testgebiet 4 West ($\tau_p = 0,785$)					
Waldstatus	0	285	51	84,8	Waldstatus	0	237	46	83,7
	1	35	497	93,4		1	23	347	93,8
% korr. (N)		89,1	90,7	(G) 90,1	% korr. (N)		91,2	88,3	(G) 89,4
Testgebiet 5 Nord ($\tau_p = 0,886$)				Testgebiet 5 Süd ($\tau_p = 0,862$)					
Waldstatus	0	166	9	94,9	Waldstatus	0	396	40	90,8
	1	7	110	94,0		1	24	472	95,2
% korr. (N)		96,0	92,4	(G) 94,5	% korr. (N)		94,3	92,2	(G) 93,1

^a laut Regressionsmodell

^b laut BWI-Datenbank

% korr. = korrekte Klassifikationen in Prozent,
gegliedert in E = Ersteller-, N = Nutzer-, G = Gesamtgenauigkeit

Tabelle 7.3: Koeffizienten der Modelle für Merkmal *Waldstatus* (ETM⁺-Daten)

Szene	b_0	b_{k_2}	Szene	b_0	b_{k_2}
TG1 Nord	9,572	-0,183	TG1 Süd	21,462	-0,421
TG2 Ost	18,908	-0,395	TG2 West	17,983	-0,421
TG3 Nord	23,051	-0,575	TG3 Süd	21,762	-0,548
TG4 Ost	16,194	-0,350	TG4 West	15,362	-0,396
TG5 Nord	22,069	-0,410	TG5 Süd	25,058	-0,494

b_0 = Modellkonstante, b_{k_2} = Koeffizient für Landsat-Kanal 2

Es fallen deutliche Unterschiede zwischen den Testgebieten auf, z. B. die geringe Genauigkeit in der Nordszene des TG1 „Westliche Mittelgebirgsschwelle“ für die Klasse *Nichtwald* mit 74,3 % (E) und die wenigen Fehlklassifikationen im TG5 „Pfälzerwald“ ($\tau_p = 0,907$). Die TGs 2 „Nordostdeutsches Tiefland“ und 4 „Alpen“ reihen sich ins Mittelfeld ein, während die beiden Modelle für TG3 „Östliche Mittelgebirgsschwelle“ ebenfalls eine hohe Vorhersageeffizienz τ_p von $\geq 0,84$ besitzen. Außerdem wird deutlich, dass bis auf wenige Ausnahmen die Klasse mit dem größeren Stichprobenumfang stets mit der höheren Genauigkeit aufwartet. Dieser Effekt konnte auch bei den anderen Merkmalen beobachtet werden und hängt damit zusammen, dass die Schätzalgorith-

Tabelle 7.4: Klassifikationsergebnisse der Modelle für Merkmal *Waldstatus* (ETM⁺-Daten)

	Vorhersage ^a				Vorhersage ^a				
	Waldstatus		% korr. (E)		Waldstatus		% korr. (E)		
Beobachtung ^b	0	1			Beobachtung ^b	0			1
Testgebiet 1 Nord ($\tau_p = 0,694$)				Testgebiet 1 Süd ($\tau_p = 0,837$)					
Waldstatus	0	427	148	74,3	Waldstatus	0	733	92	88,8
	1	57	743	92,9		1	43	791	94,8
% korr. (N)		88,2	83,4	(G) 85,1	% korr. (N)		94,5	89,6	(G) 91,9
Testgebiet 2 Ost ($\tau_p = 0,729$)				Testgebiet 2 West ($\tau_p = 0,733$)					
Waldstatus	0	2471	332	88,2	Waldstatus	0	1869	290	86,6
	1	192	1284	87,0		1	110	1039	90,4
% korr. (N)		92,8	79,5	(G) 87,8	% korr. (N)		94,4	78,2	(G) 87,9
Testgebiet 3 Nord ($\tau_p = 0,840$)				Testgebiet 3 Süd ($\tau_p = 0,845$)					
Waldstatus	0	1168	70	94,3	Waldstatus	0	915	46	95,2
	1	32	397	92,5		1	47	391	89,3
% korr. (N)		97,3	85,0	(G) 93,9	% korr. (N)		95,1	89,5	(G) 93,4
Testgebiet 4 Ost ($\tau_p = 0,800$)				Testgebiet 4 West ($\tau_p = 0,809$)					
Waldstatus	0	260	45	85,2	Waldstatus	0	253	41	86,1
	1	30	459	93,9		1	22	353	94,1
% korr. (N)		89,7	91,1	(G) 90,6	% korr. (N)		92,0	89,6	(G) 90,6
Testgebiet 5 Nord ($\tau_p = 0,907$)				Testgebiet 5 Süd ($\tau_p = 0,888$)					
Waldstatus	0	719	36	95,2	Waldstatus	0	1501	130	92,0
	1	25	554	95,7		1	68	1850	96,5
% korr. (N)		96,6	93,9	(G) 95,4	% korr. (N)		95,7	93,4	(G) 94,4

^a laut Regressionsmodell^b laut BWI-Datenbank

% korr. = korrekte Klassifikationen in Prozent,
 gegliedert in E = Ersteller-, N = Nutzer-, G = Gesamtgenauigkeit

men das Modell bei Imbalancen zwischen n_{Wald} und $n_{\text{Nichtwald}}$ stärker an jener Klasse mit dem größeren Stichprobenumfang ausrichten.

Die Ursachen für die heterogenen Ergebnisse sind unterschiedlicher Natur. Vom Standpunkt der Modellierung gesehen ergibt sich bspw. im TG1 Nord ein Nachteil durch den wesentlich breiteren Überlappungsbereich zwischen Wald und Nichtwald (lt. BWI) im Band 2 als in TG5 Nord, um zwei Extreme zu zeigen (siehe Abbildung 7.2 auf der nächsten Seite). Die Darstellung ist aus statistischer Sicht nicht korrekt, da Histogramme von diskreten Werten nicht als kontinuierliche Linie aufgetragen werden dürfen. In der SPSS-Ausgabe ist der gezeigte Modus jedoch illustrativer als ein dreidimensionales Säulendiagramm.

Verantwortlich für die schlechtere Trennbarkeit der Waldstatusklassen im TG1 Nord ist u. a.

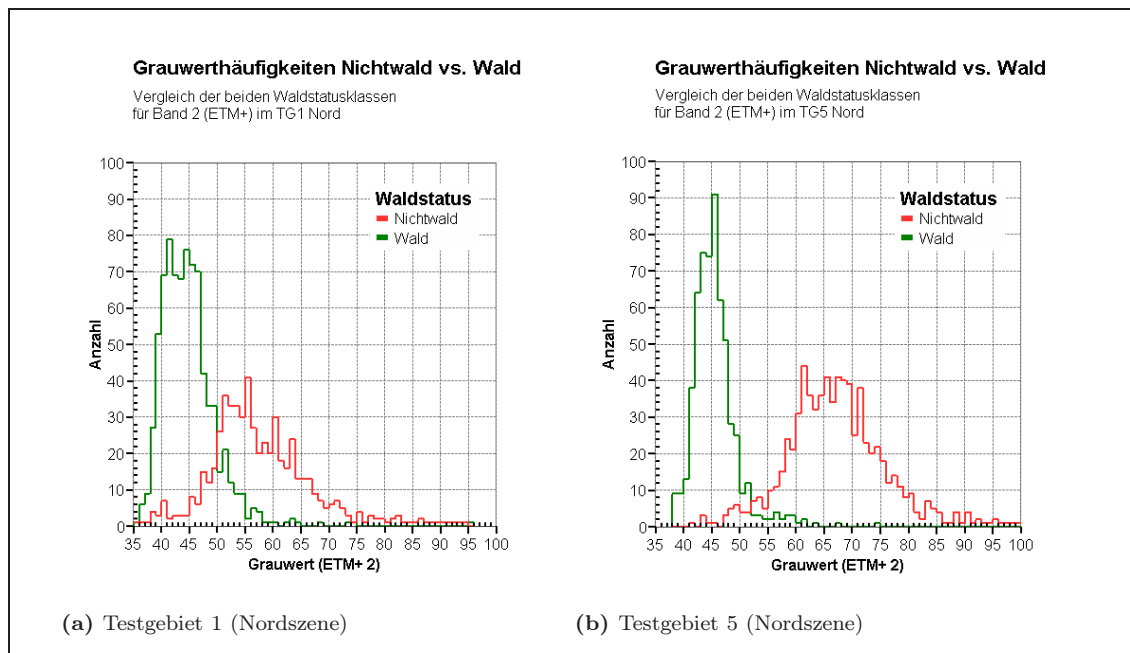


Abbildung 7.2: Unterschiedliche Überlappung im ETM⁺-Band 2

das Aufnahmedatum der ETM⁺-Szene: 15. Mai 2005. In diesem phänologischen Abschnitt besteht eine ausgeprägte Verwechslungsgefahr¹ zwischen ergrüntem Feldern bzw. Wiesenflächen und dem Wald. Entsprechend erklärt das Modell fälschlicherweise Wiesen- und Feldareale teilweise zu Wald (siehe Abbildung 7.3 auf der nächsten Seite).

Weitere Probleme ergaben sich aus der Verwendung eines einzelnen Prädiktors in Form des grünen Landsat-Bandes. Wie oben ausgeführt überwog das Streben nach einfacher Implementierbarkeit die letzten Steigerungsmöglichkeiten der Genauigkeit. Diese univariate Herangehensweise gleicht im Grunde einem Schwellenwertfahren, wie bereits auf Seite 15 angedeutet wurde. Selbiges bringt es mit sich, dass Waldflächen, die auch nur einen Grauwert über der Klassifizierungsschwelle der Klasse *Wald* liegen, falsch zugeordnet werden. Ein Beispiel dafür ist ebenfalls in Abbildung 7.3 auf der nächsten Seite gegeben. Sie zeigt einen Ausschnitt des Testgebietes 3, in dem ein Teil des Waldkomplexes als Nichtwald erkannt wurde. Eine Überprüfung der entsprechenden Pixel in der Landsat-Szene zeigte minimal höhere Grauwerte im Band 2 als bei den umliegenden Waldpixeln.

7.1.2.2 Validierung der Regressionsmodelle

Die Kreuzvalidierung der parametrisierten Modelle sollte aufzeigen, ob sich mit dem beschriebenen Verfahren numerisch stabile logistische Regressionsgleichungen erstellen lassen, die unabhängig von der Datenbasis eine konsistente Schätzung der Waldfläche zulassen. Da für die Kartierung des Waldvorkommens lediglich der Kanal TM/ETM⁺ 2 verwendet wurde, ist die Visualisierung der Validierungsergebnisse unkompliziert. Gesucht ist jener Schwellenwert eines jeden Modells, ab dem ein Pixel nicht mehr der Klasse *Wald* zugeordnet wird. Würden die zehn untersuchten Teilmodelle pro Testgebiet den gleichen Grauwert liefern, wäre eine ausreichende Stabilität erreicht.

¹ im spektralen Sinne

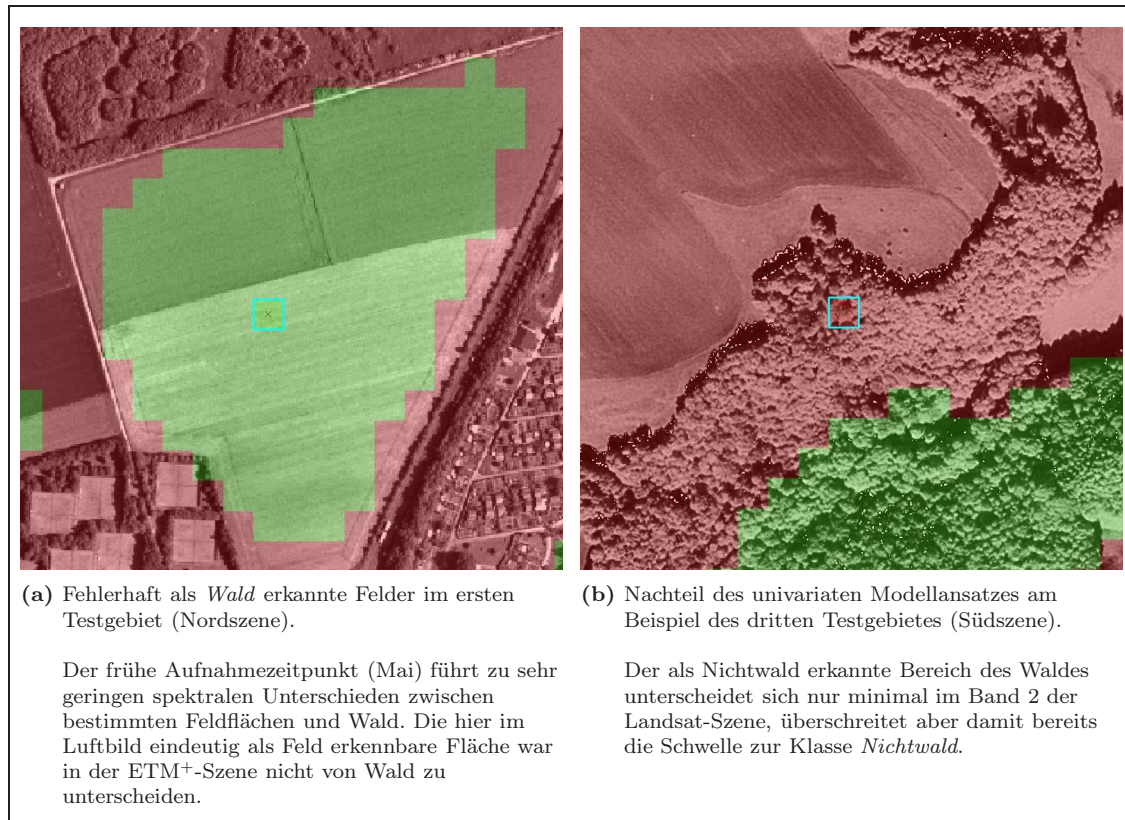


Abbildung 7.3: Mögliche Probleme bei der Waldkartierung. Rötlich eingefärbt erscheinen die als Nichtwald klassifizierten Flächen, grünlich der Wald.

Ein solches, ideales Ergebnis war nicht erzielbar, wie Abbildung 7.4 auf Seite 65 unterstreicht. Die Diagramme zeigen einen kleinen Ausschnitt der logistischen Modellfunktion, wie sie bereits auf Seite 15 vorgestellt wurde. Das Augenmerk liegt auf jenen Grauwerten, zwischen denen die Wahrscheinlichkeit eines Pixels, Wald aufzuweisen, unter 0,5 sinkt. Diese Schwelle ist durch eine rote Linie markiert. Die zehn sich zum Teil stark überlagernden Kurven repräsentieren die zehn bei der Kreuzvalidierung angepassten Modelle mit unterschiedlicher Datengrundlage. VD 1 im Diagramm bedeutet, dass bei diesem Modell die Teilpopulation 1 nicht im Modellanpassungsprozess verwendet, sondern als Validierungsdatensatz zurückgehalten wurde. Das Modell VD 2 kam dementsprechend ohne die Daten aus Teilpopulation 2 zustande usw. Der Kurvenverlauf unterstellt eine Möglichkeit zur Interpolation zwischen den Grauwerten, was praktisch unmöglich ist, da die Grauwerte einer Satellitenszene diskrete Ganzzahlen sind. Trotzdem wurde diese Darstellung gewählt, um die Unterschiede zwischen den Modellen hervorzuheben. So lange der Grenzgrauwert für die Klasse *Wald* gleich bleibt, gibt es keine Unterschiede in den erstellten Karten und damit der Waldflächenschätzung.

In Szene TG1N wurde bei neun von zehn Teilmodellen die Klassifizierungsgrenze bei Grauwert 52 erreicht, d. h. alle Pixel mit höheren Grauwerten wurden zu Nichtwald erklärt. Ein Teilmodell lieferte Grauwert 51. In dieser Szene gibt es 289.305 Pixel mit GW 52, d. h., bei einer Pixelfläche von $25 \times 25 \text{ m} = 625 \text{ m}^2$ würde eine Unterschätzung der Waldfläche² um ca. 18.000 ha erfolgen.

² in der Rohkarte ohne Medianfilter

In der Südszene des gleichen Testgebietes ergab sich das schlechteste Resultat; je fünf Teilmodelle trennten zwischen GWs 50 und 51, die andere Hälfte erst bei 51/52. Die vorhandenen 110.196 Pixel mit GW 51 ergäben also eine Differenz von knapp 7.000 ha zwischen beiden Fraktionen.

Beide Szenen des Testgebietes 2 lieferten eine sehr stabile Trennung beider Waldstatusklassen bei GW 47/48 im Osten und 42/43 im Westen.

In TG3 wurden nur wenig schlechtere Werte erzielt. Lediglich ein Teilmodell in der Nordszene sortierte statt bis GW 40 nur bis einschließlich 39 ein entsprechendes Pixel in die Klasse *Wald*. Es würden daher gegenüber den anderen neun Teilmodellen 137.790 Waldpixel mit einer Gesamtfläche von knapp 9.000 ha fehlen. Im Süden des TG lag die Grenze einheitlich zwischen 39 und 40.

Im Alpentestgebiet konnte ein ähnlich gutes Ergebnis erreicht werden wie in TG2, wenn auch der (interpolierte) Kurvenverlauf darauf schließen lässt, dass z. B. im Falle des Validierungsdatensatzes 6 nur knapp der nächst höhere Grauwert verfehlt wurde. Die Grenzen lagen im Osten bei 46/47 und im Westen bei 38/39.

Von den Modellen des Pfälzerwaldes lieferte nur eines in der Nordszene einen anderen Grauwert: Nummer 10. Dort wurde die Grenze nicht bei 53/54, sondern erst bei 54/55 gezogen, was zu einer Überschätzung von reichlich 1.300 ha Waldfläche führen würde. Im Süden wurde stets bei 50/51 getrennt.

Diese Ergebnisse unterstreichen nochmals die Tatsache, dass eine szenenübergreifende Klassifizierung mit einem einheitlichen Modell praktisch nicht vertretbar ist. Natürlich hätte der Versuch einer radiometrischen Anpassung der Szenen aufeinander unternommen werden können, allerdings gibt es bezüglich dieses Bildverarbeitungsschrittes unterschiedliche Ansichten. HALL ET AL. (1991) beschreiben eine Methode der radiometrischen Rektifikation zweier Szenen des gleichen Ausschnitts zu unterschiedlichen Zeitpunkten. Sie stellen explizit heraus, dass es sich dabei um eine *relative* Variante handelt, die für den vorliegenden Fall keine Vorteile böte. Erst die Einbindung von schwer zu akquirierenden atmosphärischen Kennwerten und Sensorkalibrierungsdaten würde eine *absolute* Anpassung ermöglichen, die für ein einheitliches Regressionsmodell notwendig wäre. Zwei weitere Verfahren zur relativen radiometrischen Normalisierung werden in OVER ET AL. (2003) und CANTY ET AL. (2004) vorgestellt, sind wegen des relativen Ansatzes allerdings ebenso wenig verwendbar wie der HALL'sche Vorschlag. BEAUBIEN ET AL. (2001) erwähnen zwar eine auf Sensorkalibrierungsdaten basierende Methodik zur absoluten Normalisierung; diese Sensordaten sind aber u. U. nicht verfügbar bzw. werden während der Laufzeit der FE-Sensoren verändert.

7.1.2.3 Validierung der Karten

Einige der beobachteten Probleme bei der Klassifizierung von Waldflächen wurden bereits im vorhergehenden Abschnitt beleuchtet. Neben den für die Modellierung relevanten Einschränkungen war es natürlich auch interessant, welche Genauigkeiten die resultierenden Karten bzw. Klassifikationen aufweisen würden. Die Ergebnisse waren insbesondere deswegen wichtig, da über die Validierung per Luftbilder Referenzdaten Eingang fanden, die nicht bereits in der Parametrisierungsphase verwendet wurden.

Es lässt sich festhalten, dass die Erkenntnisse im Wesentlichen mit den Modellklassifizierungsgenauigkeiten harmonieren. Wiederum wurde im Pfälzerwald (TG5) mit 92,3 % bzw. 94,6 % Gesamtgenauigkeit in der Nord- bzw. Südszene die höchste Anzahl Pixel richtig erkannt. Eine mögliche Erklärung dafür ist das große zusammenhängende Waldgebiet des Pfälzerwaldes selbst,

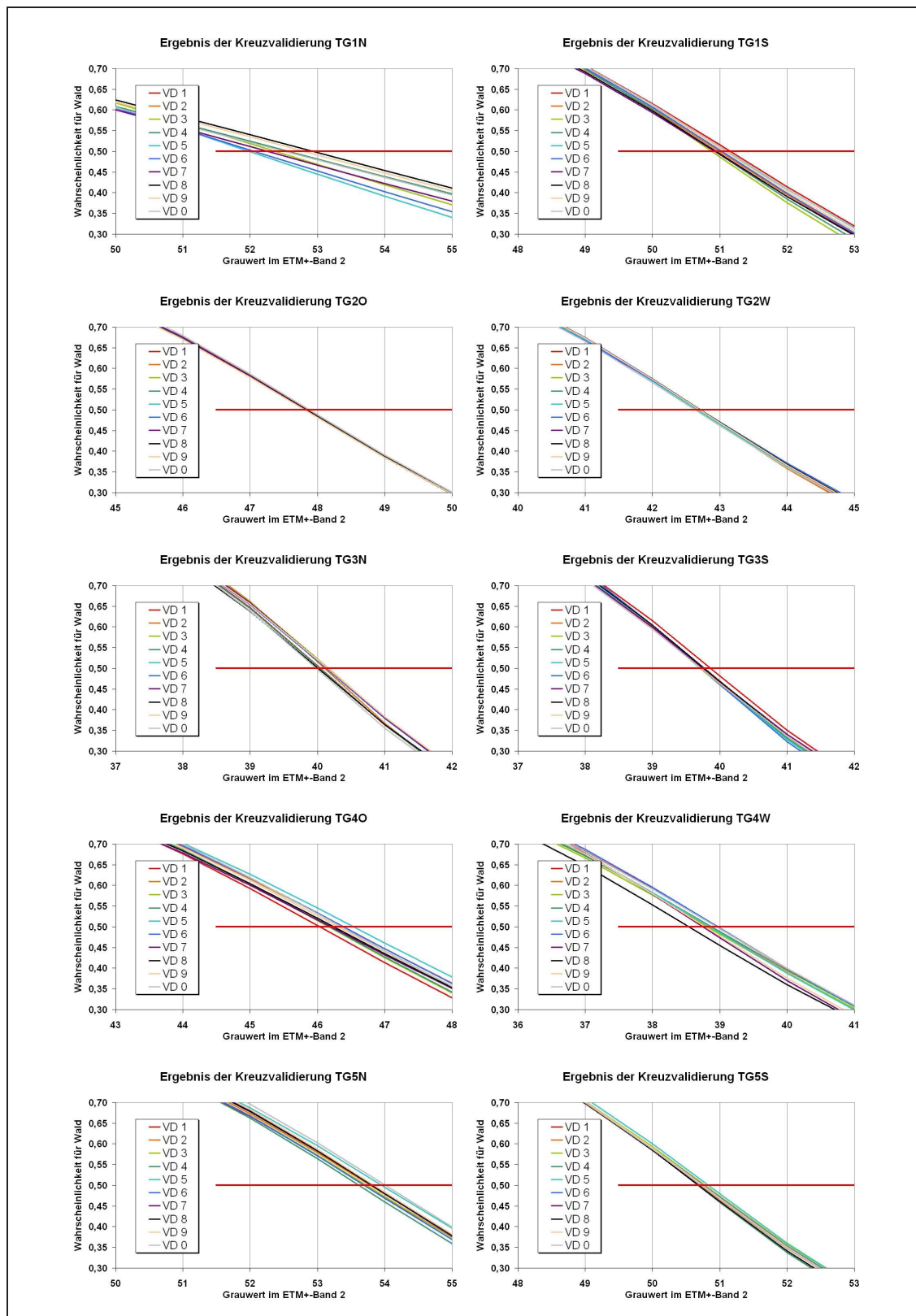


Abbildung 7.4: Ergebnisse der Kreuzvalidierung der ETM⁺-basierten Modelle. Erläuterungen im Text.

welcher in Bezug auf kleinflächige Parzellierung und zahlreiche Mischpixel die geringsten Probleme aufwirft. Wie auch bei den Regressionsmodellen reiht sich das dritte TG „Östliche Mittelgebirgsschwelle“ an zweiter Stelle der Klassifizierungsgüte ein. In der Nordszene wurden 91,6 % der Pixel, in der Südszene 90,8 % den richtigen Klassen zugeordnet. Die schlechtesten Werte wurden im ersten Testgebiet, der „Westlichen Mittelgebirgsschwelle“, erreicht. Zusätzlich zur oben erwähnten Problematik des frühen Aufnahmezeitpunktes ergaben sich Klassifizierungsschwierigkeiten durch Quellbewölkung im Nordosten der Szene. Bei der Überprüfung fielen mehrere Polygone in den Bereich von Wolken, deren große Albedo zwangsläufig zu einer Klassifizierung als *Nichtwald* führte, obwohl es Waldpixel hätten sein müssen. Die Klassifizierung der Südszene war weder durch ein zu frühes Aufnahmedatum noch durch Bewölkung erschwert, was das deutlich bessere LogReg-Modell-Ergebnis von rund 92 % gegenüber 85 % im Norden erklärt. Bei der Validierung stellte sich eine wesentlich kleinere Differenz auf niedrigerem Gesamtniveau ein—nämlich 84,7 % im Norden versus 86,5 % im Süden. Die kleinflächige Gliederung des Testgebietes dürfte dabei einen nicht unerheblichen Einfluss auf die erzielte Klassifizierungsgenauigkeit ausüben.

Auf Seite 40 wurde bereits erwähnt, dass nach Abschluss der Klassifikationen ein Medianfilter eingesetzt wurde. Zwar lässt sich auf diese Weise keine Umsetzung der komplexen BWI-Walddefinition erreichen, dennoch zeigte die parallel zur Validierung durchgeführte visuelle Überprüfung der Waldmaske, dass der Filter in vielen Fällen den erwünschten Effekt hatte. Versprengte Waldpixel wurden dadurch ebenso eliminiert wie mitten im Wald gelegene Nichtwaldpixel. An Waldkanten konnte es aber auch dazu kommen, dass durch den Filter Waldpunkte zu Nichtwald erklärt wurden, die an sich völlig korrekt erkannt worden waren (siehe auch Abbildung 7.5 auf der nächsten Seite). Außerdem ist es wichtig zu bemerken, dass die Validierung die richtige Zuordnung zur Klasse *Wald* nicht im Sinne der BWI überprüfen konnte, weil dazu im Luftbild allein keine ausreichenden Informationen zu erfassen sind. Vielmehr wurde hauptsächlich vom spektralen Standpunkt aus bewertet, ob ein Pixel richtig eingeordnet wurde. Das bedeutet, dass z. B. Polygone über ausreichend großen Baumgruppen an Siedlungsrandern—wenn als Wald klassifiziert—für richtig befunden wurden, auch wenn Baumformationen in Städten von der Walddefinition gemäß der BWI ausgenommen sind. Umgekehrt wurden als Nichtwald deklarierte Freiflächen in Waldbeständen ebenso als (spektral) richtig eingestuft, obwohl sie möglicherweise in der BWI-Datenbank als Wald geführt werden.

7.1.3 Klassifizierung und Kartierung ohne terrestrische Daten

7.1.3.1 TG1 Nordszene

In der Nordszene wurden ca. 58.000 Nichtwald- und 65.000 Waldpunkte ausgewählt und extrahiert, was einem Verhältnis von rund 47 zu 53 % entspricht und hinsichtlich einer möglichen einseitigen Optimierung des Regressionsmodells keine Probleme aufwarf. Interessanterweise waren bei der großen Zahl Eingangsdaten alle sieben ETM⁺-Bänder und der NDVI signifikant an der Verbesserung des Modells beteiligt—ein Ergebnis, das deutlich von den obigen Analysen abwich. Die Verwendung der wesentlich weniger zahlreichen BWI-Punktdateien (< 1/10 der obigen Punktzahl) führte in 8 von 10 Fällen zumindest zu insignifikanten Koeffizienten bei ETM⁺-Band 1, bei 6 Modellen auch für Band 5. Nur im Westen des Nordostdeutschen Tieflandes waren ebenfalls alle Koeffizientenschätzer signifikant. Die Ursache ist mit hoher Wahrscheinlichkeit im größeren

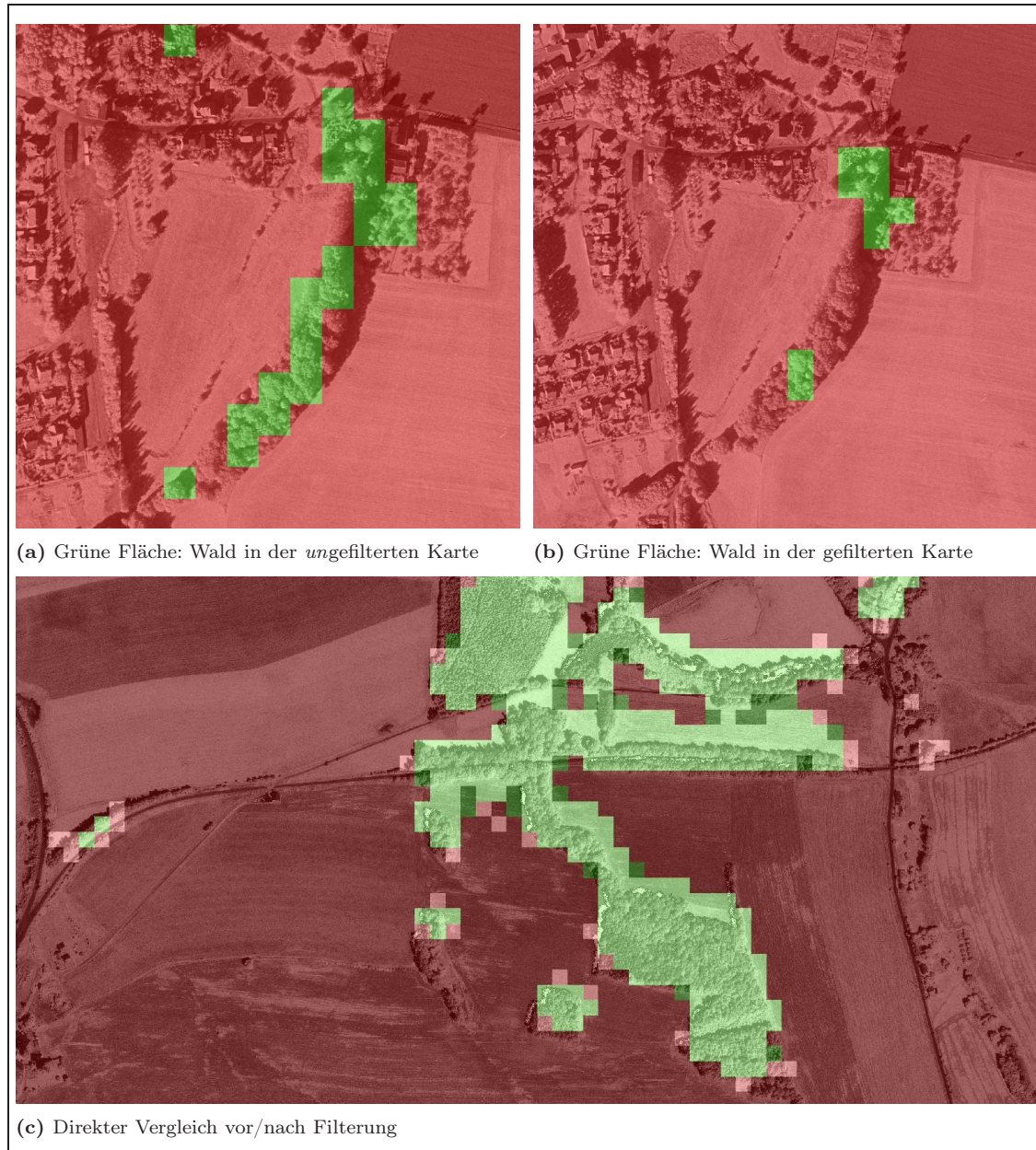


Abbildung 7.5: Auswirkung des Medianfilters auf die Waldklassifizierung.

Zu Abbildung (c):

Es werden vier Farben eingesetzt. **Hellgrün = Wald** bzw. **dunkelrot = Nichtwald** symbolisieren Übereinstimmung in der Klassifikation der unfilterten und gefilterten Karte. Die **dunkelgrünen** Flächen wurden durch den Filter von Nichtwald zu Wald, die **hellroten** von Wald zu Nichtwald.

Stichprobenumfang zu suchen, da neunmal mehr Pixel zwangsläufig auch zu wesentlich mehr Grauwertkombinationen führen sollten, was wiederum die Einbeziehung von mehr Informationen sinnvoll erscheinen lässt. Eine operationelle Umsetzung wird jedoch aus den bereits ausgeführten Gründen erschwert.

Wie schon bei der Entfernung der Gewässerflächen führte die Auswahl der Trainingspixel zu einer quasi-vollständigen Trennbarkeit der beiden Klassen mittels Regression: 98,0 % der Nichtwald- und 98,8 % der Waldpixel wurden durch das Regressionsmodell richtig zugeordnet. Die Übertragbarkeit dieses Resultats in die Praxis hält sich in engen Grenzen, weswegen mit dem ermittelten Modell und dem Modeler von Erdas eine Waldkarte erstellt wurde, die dann punktuell mit den BWI-Informationen verglichen wurde. Dabei zeigte sich, dass dieser Klassifizierungsansatz vergleichbare Genauigkeiten liefert wie das BWI-gebundene Verfahren. Mit 87,7 % Gesamtgenauigkeit sowie 78,1 % für Nichtwald und 94,6 % für Wald (jeweils Erstellergenauigkeit) liegt das Resultat zwischen den BWI-Modellen mit Kanal 2 und allen Kanälen als Prädiktoren. Offenbar ist auch die große Anzahl bewusst ausgewählter Pixel nicht geeignet, die Hindernisse der Nordszene hinsichtlich des frühen Aufnahmezeitpunktes und der Bewölkung zu umgehen.

7.1.3.2 TG1 Südszene

Im Süden des Testgebietes 1 wurden geringfügig kleinere Stichprobenumfänge ausgewählt: ca. 51.000 Nichtwald- und 57.000 Waldpunkte gingen in die Modellierung in ähnlichem Verhältnis wie in der Nordszene ein. Mit Ausnahme des mittleren Infrarots (ETM⁺-Band 5) waren wiederum alle Sensorenkanäle signifikant für die Klassifikation.

Auch hier führte die Anwendung des Modells auf die Eingangsdaten zu einer Gesamtgenauigkeit von 98,6 %, die sich angesichts des Vergleiches mit den BWI-Daten deutlich relativierte. Wie in der Nordszene reihte sich die erzielte Genauigkeit der Karte zwischen das BWI-Modell mit Kanal 2 und dem Vollmodell ein. 90,7 % der Nichtwaldecken wurden richtig eingeordnet, dasselbe galt für 93,6 % der Waldecken. Als Gesamtgenauigkeit resultierte ein Wert von 92,2 %.

Dem Verfahren kann also durchaus eine Eignung zur Kartierung von Waldvorkommen zugesprochen werden. Dabei gelten jedoch die gleichen Einschränkungen wie bei dem oben vorgestellten Verfahren. Für eine Stratifizierung ist die erreichbare Genauigkeit akzeptabel; die genaue Kartierung von Waldflächen oder gar Flächenveränderungen entzieht sich diesem Verfahren mangels höherer Genauigkeit.

7.1.4 QuickBird

7.1.4.1 Regressionsmodelle

Ursprünglich beinhaltete der BWI-II-Datenbankauszug auf dem Gebiet der QuickBird-Szene 1008 Datensätze. Die Aufteilung in Nichtwald- und Waldecken lag in einem Verhältnis von fast 1 : 10 vor, was eine einfache Verarbeitung der Daten wenig sinnvoll erscheinen ließ. Tatsächlich führte die große Überzahl der Waldecken bei der Modellanpassung dazu, dass die 94 Nichtwaldecken kaum mehr ins Gewicht fielen und entsprechend häufig fehlklassifiziert wurden. Die Modelle wurden quasi auf die alleinige Erkennung von Waldecken „optimiert“. Zur Vermeidung dieser Imbalance wurde aus der Menge der 914 Waldecken eine Zufallsstichprobe von 94 Datensätzen gezogen. Danach wurden alle Untersuchungen mit dem 2 × 94 Fälle umfassenden Datenkollektiv durchgeführt.

Die Nutzung eines einzelnen Pixels in der nativen Auflösung einer multispektralen QuickBird-Szene—280 cm (!)—zur Extraktion von Trainingsinformationen erscheint wenig praxistauglich, denn abgesehen von den schwerlich erfüllbaren Anforderungen an die räumliche Übereinstimmung von BWI-Plot und FE-Informationen integriert ein solches Pixel die spektralen Informationen von weniger als einem (erwachsenen) Baum. Selbst die oft praktizierte Methode der Mittelung von 3×3 Bildelementen dürfte in diesem Fall zu kurz greifen. Die höchsten Klassifizierungsgenauigkeiten waren unter Einbeziehung aller 4 Kanäle und des NDVI als Prädiktoren erreichbar; es konnten 71 der 94 Nichtwaldpixel und 85 der 94 Waldpixel richtig identifiziert werden. Diese optimistische Schätzung ergibt eine Gesamtgenauigkeit von 83 %.

Zur geringfügig besseren Klassifizierungen kam die künstlich auf 5 m verringerte Auflösung. Interessanterweise zeigte sich im Vergleich mit den anderen Auflösungsstufen kein einheitlicher Trend bezüglich der Interpolationsverfahren. Für die 5-m-Auflösung führte Bilinear Interpolation Resampling unter Verwendung aller Kanäle inklusive NDVI mit 85,1 % Gesamtgenauigkeit zur besten Klassifikation. Würde man nur Kanal 1 (Blau) verwenden, wäre bereits eine fast vergleichbare Klassifikation mit 83 % (gleiches Resamplingverfahren) erzielbar.

Bei 15 m Auflösung lieferte das am wenigstens rechenaufwendige Verfahren Nearest Neighbor Resampling die beste Gesamtgenauigkeit von ca. 81 % unter Verwendung der Kanäle Blau, Infrarot und NDVI. Geringfügig niedrigere Genauigkeiten waren mit der Bilinear-Interpolation-Neuberechnung bei gleicher Bandkombination möglich.

Die 25-m-Variante der QuickBird-Daten reichte nicht ganz an das Niveau der Landsat-Modelle heran. Wiederum was es das Bilinear Interpolation Resampling, das die höchste Klassifizierungsgenauigkeit von insgesamt 85,1 % (alle Kanäle + NDVI) erreichte. Konsequenz auch die univariate Variante: Der Bestwert von 81,4 % wurde über das gleiche Resamplingverfahren und Kanal 1 erreicht.

Die geringeren Klassifikationsgenauigkeiten der QuickBird-Daten im Vergleich zu Landsat-Daten entsprechen nicht den Erwartungen. Offenbar ist weder die höhere räumliche noch die bessere spektrale Auflösung ausschlaggebend für eine bessere Eignung. Auffallend war der Salz- und Pfeffer-Effekt, den die in der Auflösung veränderten QuickBird-Szenen hinterließen. Gerade die Variante mit 25 Metern Pixelgröße war in keiner Beziehung mit dem entsprechenden Ausschnitt der Landsat-Szene vergleichbar. Vermutlich führt die drastische Pixelgrößenänderung zu Problemen bei der Grauwertmittelung. Im Erdas Field Guide (LEICA GEOSYSTEMS, 2003) sind zu dieser Problematik keine weiterführenden Angaben zu finden; es wird lediglich empfohlen, bei deutlicher Veränderung der Zellgröße das Bicubic-Spline-Interpolation-Verfahren zu verwenden. Dessen Ergebnisse zeigen in der Praxis jedoch keine Vorteile gegenüber den anderen Verfahren, wie Tabelle 7.5 auf der nächsten Seite zeigt. Es waren stets die beiden einfacheren Resamplingverfahren Nearest Neighbor und Bilinear Interpolation, welche die höchsten Genauigkeiten bei der Modellanwendung ermöglichten. Auch andere Autoren (CRACKNELL, 1998; HUSEBY ET AL., 2005) verweisen auf Resamplingverfahren nur im Zusammenhang mit geometrischen Anpassungen, nicht jedoch auf deren Nutzung zur Veränderung der Pixelgröße. Im Bereich der Signalverarbeitung und Photobearbeitung lassen sich zwar Hinweise auf Probleme³ beim „Downsampling“, d. h. der Verringerung der Auflösung, finden, jedoch nicht im Bezug auf Fragestellungen aus Fernerkundung und GIS. MIKHAIL ET AL. (2001) erwähnen die Notwendigkeit einer Bildglättung vor dem Resampling

³ Entstehung so genannter Aliasing-Artefakte (JÄHNE, 2002)

7 Ergebnisse

Tabelle 7.5: Vergleich der Modellklassifizierungsergebnisse für Merkmal *Waldstatus* (QuickBird-Daten)

Auflösung in m	Resamplingverfahren ^a	Bandkombination ^b	Erstellergenauigkeit in % (NW ▷ W ▷ G) ^c
3	n/v ^d	B-G-R-IR-NDVI	75,5 ▷ 90,4 ▷ 83,0
5	BLI	B-G-R-IR-NDVI	79,8 ▷ 90,4 ▷ 85,1
15	NN	B-IR-NDVI	75,5 ▷ 88,3 ▷ 81,9
25	BLI	B-G-R-IR-NDVI	78,8 ▷ 91,5 ▷ 85,1

^a NN = Nearest Neighbor, BLI = Bilinear, CC = Cubic Convolution, BCSI = Bicubic Spline

^b B = Blau, G = Grün, R = Rot, IR = Infrarot, NDVI = Normalized Difference Vegetation Index

^c NW = Nichtwald, W = Wald, G = Gesamt

^d unveränderte Originalauflösung

zur Vermeidung von Aliasing-Effekten, gehen darauf jedoch nicht näher ein.

Aufgrund der besseren Ergebnisse von Landsat und den erheblichen Mehrkosten von Quickbird-Daten wurden diese Analysen nicht weitergeführt und keine Karten erstellt. Lediglich eine Analyse zum Einfluss der Waldeckenauswahl wurde noch unternommen, um ausschließen zu können, dass die vergleichsweise schlechteren Ergebnisse nicht das Ergebnis einer verzerrten Zufallsstichprobe waren. Zu diesem Zweck wurde jene 5-m-Variante herangezogen, die mit Bilinear Interpolation Resampling erzeugt wurde und die beste Klassifikation zuließ. Zusätzlich zu der ersten Stichprobe wurden neun weitere vom Umfang $N = 94$ aus der Population der 914 Waldecken gezogen und für Modellanpassungen verwendet. Es konnte gezeigt werden, dass Nichtwaldecken mit einer Erstlergenauigkeit im Bereich von 76 bis 81 % richtig klassifiziert wurden, während sich die Ergebnisse bei den Waldecken zwischen 86 und 93 % bewegten. Als Bereich der Gesamterstellergenauigkeit ergaben sich Werte zwischen 81 und 86 %. Daraus ist zu schließen, dass die verwendete Stichprobe von 94 Waldecken keine ungewöhnlich schlechte Repräsentation des Gesamtdatenpools darstellte.

7.1.5 Einfluss der Testgebietswahl

Durch die Verkleinerung der QuickBird-Szene von ursprünglich ~ 590 auf 3 km^2 und eine erhöhte Anzahl von Trainingspixeln ließ sich die Hypothese bestätigen, dass die Wahl der Lage und Größe eines Untersuchungsgebietes erheblich am Zustandekommen hoher Klassifikationsgenauigkeiten beteiligt ist.

Das Regressionsmodell für die ursprüngliche QuickBird-Szene ordnete insgesamt 83 % der 2×94 Trainingspixel den richtigen Klassen zu. Setzt man diese 188 Bildelemente ins Verhältnis zur Gesamtpixelzahl des Bildes—ca. 64.000.000—resultiert eine Stichprobenrate von 0,0029 ‰. Bei dem Kleinstgebietsversuch wurde mit zwei Stichprobenraten gearbeitet: ~ 75 und ~ 4 %. In Verbindung mit der überwiegend kompakten Waldform ließen sich mit den beiden Regressionsmodellen Genauigkeiten von 99 bzw. 98,8 % erreichen. Daraus lässt sich einerseits ableiten, dass durch Delimitierung gewonnene Trainingsdaten eine angemessenere Parametrisierung möglich ist, als durch die wenigen, punktuellen Informationen der BWI-Datenbank. Dies spiegelt sich auch in der sichtbaren Steigerung der Vorhersageeffizienz von $\tau_p = 0,660$ auf 0,977 (für beide Stichprobenraten) wider.

Tabelle 7.6: Validierungsergebnisse der Kleingebiets-Waldkartierung(QuickBird-Daten)

Klassifikation ^b	Referenz ^a			% korr. (N)	Klassifikation ^b	Referenz ^a			% korr. (N)
	Waldstatus					Waldstatus			
	0	1			0	1			
723.594 Trainingspixel				42.371 Trainingspixel					
<i>Flächenbezogene Auswertung</i>									
Waldstatus	0	2,80	0,21	93,1	Waldstatus	0	2,81	0,22	92,7
	1	0,32	5,66	94,6		1	0,31	5,65	94,7
% korr. (E)		89,7	96,4	(G) 94,1	% korr. (E)		90,0	96,2	(G) 94,1
<i>Punktbezogene Auswertung</i>									
Waldstatus	0	124	8	93,9					
	1	14	254	94,8					
% korr. (E)		89,9	96,9	(G) 94,5					

^a laut panchromatischem QuickBird-Kanal

^b laut Regressionsmodell

% korr. = korrekte Klassifikationen in Prozent,
gegliedert in E = Ersteller-, N = Nutzer-, G = Gesamtgenauigkeit

Zum anderen wird deutlich, dass der Verzicht auf rund 682.000 Trainingsbildelemente durch die Reduktion der Stichprobenrate auf 4 % keine signifikante Verschlechterung der Modellvorhersagen bedeutet.

Es wurde bereits angedeutet, dass die resultierenden Waldmasken einer Validierung bzw. Genauigkeitsanalyse unterzogen wurden. Das Ergebnis wird in zwei Tabellen auf unterschiedliche Weise dargestellt. In Tabelle 7.6 erfolgt die Gegenüberstellung der Waldmaskengenauigkeiten für *beide* Stichprobenraten. Für diese Tabelle wurde ein Vergleich der Waldflächen aus Validierungspolygonen und Waldmasken vorgenommen. Mit einer Gesamtgenauigkeit von 94,1 % unabhängig von der Stichprobenrate liegt der Anteil richtig klassifizierter Pixel in der Waldmaskenstichprobe zwar rund 5 Prozent unter dem durch die Modellauswertung vorhergesagten Wert, jedoch auf einem hohen Niveau für einen pixelbasierten Ansatz.

In der gleichen Tabelle wird auch das Ergebnis eines anderen Auswertungsweges für die auf 723.591 Trainingspixeln beruhende Waldmaske dargestellt. Das Augenmerk lag dort auf einem Anteilsvergleich, um den Einfluss von Streupixeln zu mindern. Dazu wurde eine einfache Gegenüberstellung der dominierenden Klasse in Prüfpolygon und Waldmaske vorgenommen. Hatte bspw. ein Prüfpolygon eine Waldflächensumme $> 112,5 \text{ m}^2$ (die Hälfte der Polygonfläche von 225 m^2), wurde diesem Polygon als Ganzem der Status *Wald* zugewiesen. Lag die Waldfläche der Waldmaske auf dem Gebiet des Prüfpolygons aber unter $112,5 \text{ m}^2$, so erhielt es den Status *Nichtwald* und wurde als fehlklassifiziert eingestuft. Interessanterweise ergeben sich durch diese Methodik keine nennenswerten Unterschiede, denn auch in diesem Fall liegen die Genauigkeiten auf ähnlichem Niveau. Wieder werden etwas mehr als 94 % richtige Klassifizierungen erreicht.

7.1.6 Diskussion

Bei der Gegenüberstellung der Ergebnisse der Waldkartierung mittels logistischer Regression mit Resultaten anderer Forschungsbemühungen ergibt sich ein differenziertes Bild. SCHARDT (1998) gibt als maximale Klassifizierungsgenauigkeit über 97 % für Landsat-Daten an, zeigt als Beispiel in dem erwähnten Artikel allerdings einen Szenenausschnitt des mit 10 m Pixelgröße deutlich höher auflösenden SPOT-Systems. Auch schreibt er in einer früheren Publikation (SCHARDT, 1990), dass bei unitemporaler Verwendung von Sommerszenen keine zufrieden stellende, automatisierte Trennung der Klassen *Wald* und *Nichtwald* zu erreichen ist.

RACK (2001) untersuchte in einem ganz ähnlichen Kontext⁴ wie das WEI-BWI-Projekt die Eignung von ETM⁺-Daten zur Waldkartierung als erster Phase in einer Stratifikation. Er nutzte ein mehrstufiges Verfahren, welches sowohl die Festlegung von Trainingsgebieten, unüberwachte Klassifizierungen zur Signaturgewinnung als auch die Verwendung von Orthophotos zur Verbesserung des Klassifizierungsregelwerkes umfasste. Damit war er in der Lage, Gesamtgenauigkeiten zwischen 85 und 88 % zu erzielen. Interessant an dem Verfahren ist die ausgeprägte Tendenz, stets die Klasse *Wald* mit besserer Genauigkeit zu erfassen.

Mit 92 bzw. 94 % Gesamtklassifizierungsgenauigkeit liegen die von KOCH ET AL. (2003) erzielten Werte etwa auf dem Niveau der in dieser Arbeit präsentierten Werte. Als Grundlage der Waldkartierung diente ein hierarchisches Klassifizierungsverfahren im Zusammenspiel mit einem DHM und Bildfusionen aus IRS 1-D- und ETM⁺-Szenen, die somit eine wesentlich höhere Auflösung aufwiesen. Der Artikel enthält keine expliziten Angaben zur Größe der beiden in der Schweiz gelegenen Testgebiete, ein Ergebnisdiagramm lässt aber auf jeweils rund 1.800 ha schließen. Mit vergleichbaren Verfahren und Daten arbeitete IVITS-WASSER (2004), unterteilte jedoch in insgesamt vier Klassen (Wald, Grünland, Versiegelt, Offene Fläche) und erreichte eine Erstellergenauigkeit von über 97 % für die Klasse *Wald*. In dieser Untersuchung wurden sechs Testgebiete mit jeweils 1 km² Fläche bearbeitet.

Ebenso wie RACK (2001) untersuchte BLINN (2005) Landsat-Daten im Kontext des Forest-Inventory-and-Analysis-Programms des US-Landwirtschaftsministeriums. Die höchsten erzielbaren Gesamtgenauigkeiten für die Kartierung von Wald lagen zwischen 83 und 88 %—Werte die nur in einigen Fällen in dem für das Programm erforderlichen Präzisionsrahmen lagen.

Die ebenfalls im Rahmen des WEI-BWI-Projektes durchgeführten Untersuchungen von OEHMICHEN (2006) zeigen mit einer anderen Methode vergleichbare bzw. leicht bessere Genauigkeiten. Angewendet wurde ein mehrstufiges Verfahren auf Grundlage von Schwellenwerten, das über umfangreiche Signaturanalysen und eine abschließende Adaption der Rohkarten an die BWI-Walddefinition insgesamt eine etwas bessere Kartierungs- bzw. Klassifizierungsgrundlage bildet als die logistische Regression.

Für QuickBird lassen sich Anwendungsbeispiele finden, die hinsichtlich einer Waldklassifizierung erheblich bessere Werte erzielen. IVITS und KOCH (2003) untersuchten neben Fusionsprodukten aus IRS- und Landsat-Daten (siehe oben) auch QuickBird-Daten und Orthophotos. Die Klassifikation der QuickBird-Szenen erfolgte für die Klasse Wald mit 100%iger (!) Nutzer- und Erstellergenauigkeit, die nur mit einem objektbasierten Verfahren erreichbar ist. Verwendet wurde ein kombinierter Ansatz mit Diskriminanzanalysen und objektorientierter Klassifizierung. Bezogen

⁴ Phase One Stratification for Forest Inventory and Analysis des United States Department of Agriculture Forest Service

auf einen hier angestrebten bundesweiten Einsatz sind jedoch die bereits erwähnten Testgebietsgrößen zu beachten!

Ebenfalls bessere Ergebnisse erreichte OEHMICHEN (2006) mit einem pixelbasierten Verfahren: über 94 % Gesamtgenauigkeit wurden erreicht. Über Adaptionen an die BWI-Definition⁵ war eine Steigerung um weitere 3 Prozent bei Verwendung der BWI-Daten als Referenz möglich.

Der Kleingebietsversuch innerhalb der QuickBird-Szene hat deutlich gemacht, welchen Einfluss veränderte Trainingsdatengewinnung, Waldfragmentierung und Gebietsgröße haben. Ergebnisse aus Pilotstudien oder Klassifizierungen kleinster Untersuchungsgebiete sind also nur mit großer Vorsicht zu verallgemeinern. Der pixelbasierte Ansatz in Verbindung mit logistischer Regression erscheint daher nicht als geeignetes Verfahren zur großflächigen Anwendung, wenn man neben den erreichten Genauigkeiten die immensen Kosten pro Flächeneinheit betrachtet. Denkbar ist die Verwendung als ein Schnelltestverfahren mit moderatem Aufwand für kleinere Gebiete oder im Kontext einer mehrstufigen Analyse. Zur Kombination mit objektorientierten Ansätzen bedarf es weiterführender Forschung.

Für die direkte Kartierung und Ermittlung der Waldfläche im Rahmen einer dritten Bundeswaldinventur ist das Regressionsverfahren unabhängig vom Satellitensensor ungeeignet, da sich die komplexe Walddefinition gemäß BWI nicht vollständig reproduzieren lässt und die Genauigkeiten unter den Anforderungen der BWI liegen. RIEDEL (2006) konnte jedoch zeigen, dass die resultierenden Waldkarten ausreichend genau für Stratifizierungen sind, die in einem mehrstufigen Verfahren Voraussetzung zur Steigerung der Genauigkeit oder alternativ zur Senkung von Kosten wären. Eine weitere Anwendungsmöglichkeit besteht in der Klassifizierung kleinerer Gebiete, in denen durch eine individuelle Anpassung an gegebene Verhältnisse und Einbindung von Hilfsdaten Steigerungen der Genauigkeit über die hier ermittelten Werte hinaus möglich sind.

7.2 Merkmal *Mischung*

Die im Methodenteil erwähnten Vorschläge von ITTEN ET AL. (1992) legten den Schluss nahe, dass es ähnlich wie beim Waldstatus—abgesehen von geringen Überlappungsbereichen—zumindest im nahen Infrarot (ETM⁺-Band 4) sichtbar abgesetzte Merkmalsbereiche für jede der drei Klassen *Nadelwald*, *Laubwald* und *Mischwald* geben müsste. Das konnte in dieser Untersuchung nicht gezeigt werden. Selbst dreidimensionale Streudiagramme mit drei Kanälen ließen keine zuverlässig trennbaren Merkmalsräume erkennen. Die z. T. weiten Überschneidungen werden in der Abbildung auf der nächsten Seite und 7.7 auf der nächsten Seite verdeutlicht. Band 1 bspw. gestattet keine Klassifizierung, aber selbst Band 4 lässt mit seiner deutlich größeren Dynamik nur eine Trennbarkeit der Klassen *Nadelwald* und *Laubwald* erkennen, während sich *Mischwald* unter den beiden anderen Klassen verbirgt.

Dennoch wurde versucht, durch die umfassende Einbindung aller trennungsrelevanten Kanäle ein nutzbares Resultat zu erhalten. Tabelle 7.7 auf Seite 76 enthält die Ergebnisse der Modellanpassung für die nördliche Szene des Testgebiets 1. Es ist ersichtlich, dass der komplette zur Verfügung stehende Prädiktorensatz Verwendung fand. Nachteilig an der Einbindung aller Kanäle sind auftretende Redundanzen bzw. insignifikante Koeffizienten (siehe Spalte „Sig.“). Durch den

⁵ Eliminierung von Kleinstholzbodenflächen per „Clump & Sieve“ und Einbindung von ATKIS-Daten, für Erläuterungen siehe OEHMICHEN (2006)

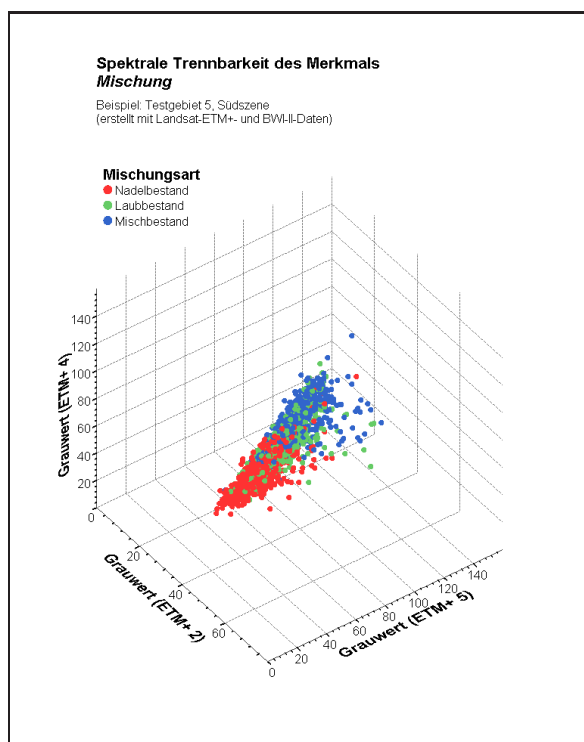


Abbildung 7.6
 3D-Streudiagramm des Merkmals *Mischung* (Pfälzerwald, Südszene).

Das Streudiagramm zeigt, wie wenig der Mischwald von den beiden anderen Klassen trennbar ist. In weiten Bereichen des Grauertraumes gibt es eine Überlappung.

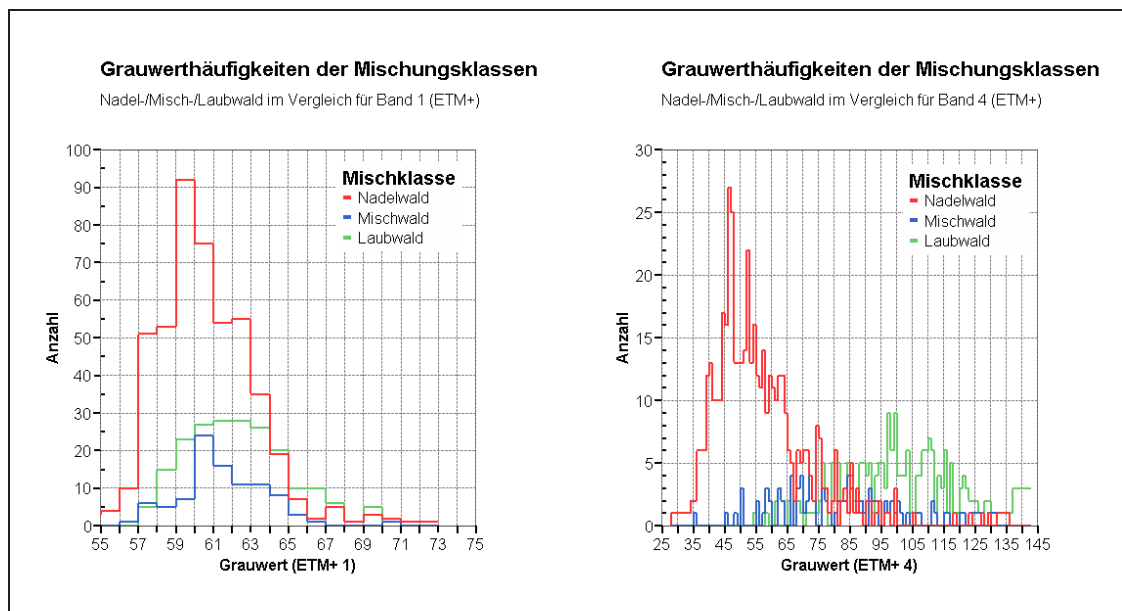


Abbildung 7.7: Unterschiedliche Trennbarkeit der Mischungsklassen in den ETM⁺-Bändern 1 (links) und 4 (rechts)

Ausschluss der laut Statistik nicht signifikant von Null verschiedenen Prädiktoren verschlechterte sich im Gegenzug die Klassifizierbarkeit einzelner Klassen. In den häufigsten Fällen wurden dann weniger Mischwaldpixel richtig zugeordnet.

Die Umsetzung der in Tabelle 7.7 auf der nächsten Seite enthaltenen Informationen gestaltet sich etwas anders als bei dem Merkmal *Waldstatus*. Mit Hilfe von drei Gleichungen werden getrennt die Wahrscheinlichkeiten für alle drei Mischungsklassen ermittelt (HOSMER und LEMESHOW, 1989):

$$P(Y = \text{Nadelwald}|\mathbf{X}) = \frac{1}{1 + e^{\beta_{\text{MW}}\mathbf{X}} + e^{\beta_{\text{LBW}}\mathbf{X}}} \quad (7.2)$$

und

$$P(Y = \text{Mischwald}|\mathbf{X}) = \frac{e^{\beta_{\text{MW}}\mathbf{X}}}{1 + e^{\beta_{\text{MW}}\mathbf{X}} + e^{\beta_{\text{LBW}}\mathbf{X}}} \quad (7.3)$$

sowie

$$P(Y = \text{Laubwald}|\mathbf{X}) = \frac{e^{\beta_{\text{LBW}}\mathbf{X}}}{1 + e^{\beta_{\text{MW}}\mathbf{X}} + e^{\beta_{\text{LBW}}\mathbf{X}}}. \quad (7.4)$$

$\beta_{\text{MW}}\mathbf{X}$ und $\beta_{\text{LBW}}\mathbf{X}$ sind die linear verknüpften Prädiktoren mit ihren Koeffizienten. Setzt man beispielsweise die per Regression geschätzten Koeffizienten aus der Tabelle für Mischwald (MW) ein, erhält man:

$$\begin{aligned} \beta_{\text{MW}}\mathbf{X} = & 3,985 + 0,199k_1 - 0,129k_2 - 0,096k_3 - 0,072k_4 \\ & + 0,175k_5 - 0,186k_6 - 0,091k_7 + 0,091ndvi \end{aligned} \quad (7.5)$$

Ein zu klassifizierendes Bildelement wird jener Klasse zugeordnet, deren Wahrscheinlichkeit am größten ist. Auch die SPSS-Routine NOMREG produziert auf Wunsch eine Klassifikationstabelle auf Basis der Eingangsdaten und geschätzten Gleichungen (Tabelle 7.8 auf der nächsten Seite). Die für die westliche Mittelgebirgsschwelle (TG1) gezeigten Werte können als Stellvertreter für die anderen Testgebiete gelten: Die Gesamtgenauigkeit bewegt sich zwischen 65,6 und 70,5 %—den beiden Extrema, die in Süd- bzw. Nordszene errechnet wurden. Am schlechtesten schnitt stets und vorhersehbar die Klasse *Mischwald* mit Erstellergenauigkeiten zwischen reichlich 39 und knapp 60% ab. Die letzte Zeile „Gesamt-%“ zeigt den Anteil der insgesamt pro Klasse eingeordneten Fälle an der Summe aller betrachteten Fälle.

Im Kontext einer möglichen nächsten BWI sind Klassifizierungsgenauigkeiten auf dieser Ebene als unbrauchbar zu beurteilen. Betrachtet man das untere Ende des Intervalls, nämlich die Südszene des ersten Testgebietes, fällt auf, dass die erwähnten 39,5 % Genauigkeit für MW nur mäßig besser als eine Zufallsklassifizierung (bei drei Klassen 33 %) sind—und dies nach einer systembedingt optimistischen Einschätzung durch SPSS (siehe Seite 43)! Auch das Maß für die Vorhersageeffizienz, τ_p , bestätigt dies mit einem vergleichsweise niedrigen Wert von 0,48. Selbst die Nordszene mit dem besten Ergebnis weist nur einen Wert von $\tau_p = 0,56$ auf.

Bei zweistufiger Annäherung an die Fragestellung ergeben sich bessere Werte, wie Tabelle 7.9 auf Seite 77 zeigt. Hier wurde zunächst eine Unterteilung in Nadelwald- und andere Bestände⁶ vorgenommen, im Anschluss erfolgte die Trennung von Laub- und Mischwald. Auch dabei wurden wieder in der Nordszene des ersten Testgebietes anteilmäßig die meisten Pixel den richtigen Klassen zugeordnet. Lediglich in einer anderen Szene (im Westen der Alpen) wurde in der zweiten

⁶ eine Zusammenfassung aller Laub- und Mischwaldbestände

7 Ergebnisse

Tabelle 7.7: Koeffizienten für LogReg-Modell zur Mischungskartierung im Norden des ersten Testgebiets (ETM⁺-Daten)

Mischung ^a	Prädiktor	Koeffizient	St'dfehler	Wald ^b	FG	Sig.
Mischwald	Intercept	3,985	12,727	0,098	1	0,754
	k_1	0,199	0,115	3,027	1	0,082
	k_2	-0,129	0,125	1,061	1	0,303
	k_3	-0,096	0,126	0,586	1	0,444
	k_4	-0,072	0,052	1,903	1	0,168
	k_5	0,175	0,059	8,745	1	0,003
	k_6	-0,186	0,072	6,626	1	0,010
	k_7	-0,091	0,075	1,481	1	0,224
	ndvi	0,091	0,053	2,916	1	0,088
Laubwald	Intercept	7,969	17,952	0,197	1	0,657
	k_1	0,328	0,139	5,584	1	0,018
	k_2	-0,219	0,148	2,180	1	0,140
	k_3	-0,190	0,165	1,333	1	0,248
	k_4	-0,082	0,061	1,820	1	0,177
	k_5	0,269	0,073	13,490	1	0,000
	k_6	-0,303	0,093	10,530	1	0,001
	k_7	-0,116	0,102	1,297	1	0,255
	ndvi	0,124	0,075	2,723	1	0,099

^a Die Referenzklasse lautet: Nadelwald.

^b Gemeint ist die Wald-Statistik, d. h. eine Prüfgröße für LogReg-Verfahren.

St'dfehler = Standardfehler, FG = Freiheitsgrade, Sig. = Signifikanz

Tabelle 7.8: Klassifikationsergebnisse der einphasigen Modelle für das Merkmal *Mischung* im ersten Testgebiet (ETM⁺-Daten)

Nordscene					Südscene				
Referenz	Vorhersage				Referenz	Vorhersage			
	Ndw	Mw	Lbw	% richtig (E)		Ndw	Mw	Lbw	% richtig (E)
Ndw	78	13	4	82,1	Ndw	103	20	1	83,1
Mw	18	51	26	53,7	Mw	27	49	48	39,5
Lbw	4	19	72	75,8	Lbw	6	26	92	74,2
Gesamt-%	35,1	29,1	35,8	(G) 70,5	Gesamt-%	36,6	25,5	37,9	(G) 65,6

Tabelle 7.9: Klassifikationsergebnis des zweiphasigen Modells für das Merkmal *Mischung* im Norden des ersten Testgebietes (ETM⁺-Daten)

Referenz	nNdw	Vorhersage	
		Ndw	% richtig (E)
nNdW	164	26	86,3
NdW	18	172	90,5
(G) 88,4			
Referenz	Lbw	Mw	% richtig (E)
Lbw	73	22	76,8
Mw	25	70	73,7
(G) 75,3			

Phase ein geringfügig höherer Wert von fast 79 % erreicht.

Oberflächlich betrachtet gestattet dieses Verfahren die besseren Klassifizierungen—vor allem der Klasse *Mischwald*. Tatsächlich aber spiegeln diese Zahlen keine praktisch erzielbaren Werte wider, da mit dem a priori vorhandenen Wissen über die Klassenzugehörigkeiten gearbeitet wurde. In Stufe eins wurde dem Modell als abhängige Variable die Mischung mit den Ausprägungen *Nadelwald* und *Andere* präsentiert. Im zweiten Schritt nahm ein weiteres Modell die Mischung mit den (vorher bekannten) Ausprägungen *Mischwald/Laubwald* auf. Dabei blieben die Nadelwaldpunkte bzw. das Ergebnis der ersten Regression unberücksichtigt; die zweite Anpassung erfolgte also unabhängig von der ersten. So übernahm sie die Fehlklassifikationen aus der ersten Stufe nicht. Damit lässt sich eher das spektrale Potenzial auf Landsat-Basis als die praktische Anwendbarkeit abschätzen.

In der Praxis hätte der Ansatz das Problem zu bewältigen, auf der ersten Klassifizierung *aufbauen* zu müssen. Das hieße, dass von den laut Modell insgesamt $164 + 18 = 182$ vorhergesagten Nichtnadelwaldecken zwangsläufig mindestens die 18 „fehlsortierten“ in der zweiten Phase als Grundfehler auftauchen würden. Rechnet man die Fehler aus Fehlklassifikationen des zweiten Schritts hinzu, erscheint das Verfahren genauso unzulänglich wie jenes mit einer multinomialen Regression. Der beschriebene Effekt wurde zur Illustration für die Nordszene des ersten Testgebietes quantifiziert. Das Ergebnis der Berechnung ist in Tabelle 7.10 auf der nächsten Seite dargestellt. Zwar steigt die Gesamtgenauigkeit der Klassifizierung um ca. 5 % gegenüber dem NOMREG-Modell, der ohnehin schlecht klassifizierbare *Mischwald* geht aber noch stärker in den beiden anderen Klassen unter. τ_p wächst wie die Gesamtgenauigkeit leicht an und erreicht den Wert 0,60. Im Grunde bestätigen diese Ergebnisse die eingangs formulierte Hypothese, dass *Mischwald*—schon begrifflich eine Mischung aus Laub- und Nadelwald—spektral schlecht trennbar ist.

Nach MENARD (2001) ist die einstufige Variante in Form der multinomialen logistischen Regression einem sequentiellen Verfahren mit zwei binomialen Regressionen vorzuziehen. Diese Aussage lässt sich im gezeigten Fall nicht bedingungslos bestätigen; letztendlich entscheidet die Problemstellung über die Wahl eines geeigneten Verfahrens. Führt man sich vor Augen, dass seit Beginn der neunziger Jahre vor allem in den neuen Bundesländern, aber auch im alten Bundesgebiet mas-

Tabelle 7.10: Klassifikationsergebnis des zweiphasigen, realistischen Modells für das Merkmal *Mischung* im Norden des ersten Testgebietes (ETM+-Daten)

Referenz	Vorhersage			% richtig (E)
	Ndw	Mw	Lbw	
Ndw	172	5	13	90,5
Mw	23	31	41	32,6
Lbw	3	9	83	87,4
% richtig (N)	86,9	68,9	60,6	(G) 75,3

sive Bemühungen zum Waldumbau⁷ unternommen werden, und dass diese Anstrengungen in einer erheblichen flächenmäßigen Ausweitung von Mischbeständen resultieren, kann das beschriebene Verfahren nicht als geeignet eingestuft werden. Weiterer Forschungsbedarf besteht wiederum in Hinblick auf die mögliche Kombination der Regression mit objektbasierten Verfahren und spektralen Entmischungsverfahren zur Schätzung von Subpixel-Anteilen der betrachteten Klassen (z. B. JU ET AL. (2002)).

WERNER (2002) unterzog CIR-Luftbilder und IRS-1C/D-Daten einem visuellen Vergleich. Während in einem CIR-Luftbild Mischwälder deutlich und differenziert erkennbar sind, bestimmen im Satellitenbild die beteiligten Baumarten und Mischungsanteile, ob eine Unterscheidungsmöglichkeit gegeben ist. Da visuelle Interpretationsverfahren bezüglich der Differenzierbarkeit von Landbedeckungsklassen einer automatischen Klassifizierung (zurzeit noch) überlegen sind, gibt diese Arbeit einen Hinweis auf die Herausforderung bei der Kartierung von Mischwald.

In der bereits erwähnten Studie von IVITS und KOCH (2003) wurden mit objektbasierten Algorithmen auch weiterführende Klassifikationen vorgenommen, die u. a. die Klasse *Mischwald* enthielten. Es konnten je nach Datenquelle zwischen 68 % (QuickBird), 89 % (Fusion Landsat-IRS) und 99 % (Luftbild) Erstellergenauigkeit erzielt werden, was für eine deutliche Überlegenheit objektbasierter Ansätze spricht, sieht man von der Testgebietsgröße an dieser Stelle ab. Pixelbasierte Verfahren nutzen zur Klassifizierung von Mischwäldern, also Gebieten, die nahezu vollständig aus Mischpixeln bestehen, ausschließlich Grauwerte, während sich an Objekten orientierende Verfahren zusätzliche Informationen wie Textur und Nachbarschaftsbeziehungen nutzen können.

In der Schweiz wurde eine Waldmischungsgrad-Karte unter Verwendung von Landsat-TM-Daten erstellt (BUNDESAMT FÜR STATISTIK, SEKTION GEOINFORMATION, 2004). Im ersten Schritt erfolgte die Klassifikation der Szenen in Wald und Nichtwald, was mit einer Gesamtgenauigkeit von fast 92 % gelang. Dieser hohe Wert ist bemerkenswert, insbesondere, wenn man das bewegte Relief und die Größe der Schweiz bedenkt. Für die Abgrenzung von Waldmischungen wurde eine weitere Unterteilung des Waldes in Nadel-, Nadelmisch-, Laubmisch- und Laubwald vorgenommen. Dabei zeigten sich vergleichbare Schwierigkeiten wie in dieser Arbeit. Die beiden Mischwaldklassen

⁷ vgl.

LANDESFORSTPRÄSIDIUM SACHSEN (2003)

LANDESFORSTANSTALT EBERSWALDE, FACHBEREICH WALDENTWICKLUNG/MONITORING (2004)

BAYERISCHE LANDESANSTALT FÜR WALD UND FORSTWIRTSCHAFT (2005)

LANDESFORSTPRÄSIDIUM SACHSEN (2005)

FRITZ (2006)

wurden zu lediglich 38 % Herstellergenauigkeit erkannt, da in großen Anteilen den jeweiligen Reinbestandesklassen zugeordnet. Die Gesamtgenauigkeit der Karte bei Vergleich mit dem ersten Schweizerischen Landesforstinventar lag bei rund 60 %.

7.3 Merkmal *Totholzvorkommen*

7.3.1 Regressionsmodelle

Die Klassifizierung von Totholzvorkommen auf den BWI-Plots war mit sehr unterschiedlichen Genauigkeiten möglich (Tabelle 7.11 auf der nächsten Seite). Die Nutzergenauigkeiten für die Klasse *Totholz vorhanden* liegen zwischen 67 % im Norden der westlichen (TG1) und 84 % im Norden der östlichen Mittelgebirgsschwelle (TG3). Das Intervall der Klassifizierungsgenauigkeiten für Plots *ohne* Totholz reicht von 73 % bis 97 % (jeweils die gleichen Szenen wie bei vorhandenem Totholz). Gemessen an der Gesamtgenauigkeit war Testgebiet 3 mit den kleinsten Parametrisierungskollektiven ($n = 43$ im Norden und $n = 42$ im Süden) das am genauesten klassifizierte. Die Werte für die Vorhersageeffizienz bewegen sich in einem vergleichbar großen Rahmen, was eine einheitliche Wertung der Ergebnisse erschwert.

Nennenswert ist auch der Umstand, dass die überwiegende Mehrheit der Totholzfunde liegendes Material war. Davon ausgehend wäre ähnlich wie in der Literatur eine sehr schlechte Klassifizierbarkeit zu erwarten gewesen. Nichts desto trotz wurden im Durchschnitt rund 70–75 % des abgestorbenen Holzes im Landsat-Bild erkannt. Zur Klärung der Ursachen wurden Luftbilder, QuickBird-Daten, die entsprechenden Landsat-Szenen und schließlich die Punktinformationen aus der BWI-Datenbank in einem GIS verglichen. Die visuelle Interpretation ließ keine ganzheitlichen Schlüsse zu. So waren an manchen Stellen durchaus Blößen im Luftbild erkennbar, die ausreichend groß waren, um sich auch in einem 900 m² großen Landsat-Pixel zu manifestieren, an anderen Stellen ließ das Luftbild in keiner Weise den Schluss auf Totholz zu, und trotzdem wies das zugehörige Landsat-Bildelement minimal andere Grauwerte als die es umgebenden Pixel auf. Auch ungewöhnlich war der große Beitrag der sichtbaren Kanäle Grün und Rot für die Klassifizierung, die neben dem thermischen Infrarot (!) deutlich mehr Einfluss auf das Ergebnis ausübten als das nahe Infrarot. Dies steht in markantem Gegensatz zu den Ergebnissen anderer Untersuchungen (s. o.), in denen der NIR-Kanal ein wichtiges Element zur Totholzerfassung ist.

7.3.2 Validierung der Regressionsmodelle

Bedingt durch die geringen Stichprobenumfänge hätte eine szenenbasierte Stabilitätsprüfung per Kreuzvalidierung wenig aussagekräftige Ergebnisse erzielt. Im Testgebiet 3 wäre die Größe des Validierungsdatensatzes auf 9 geschrumpft (10 % von ca. 90 Punkten). Wie in der Methodik erwähnt erfolgte unter diesen Bedingungen nur eine gleichzeitige Betrachtung aller zur Verfügung stehenden Totholzinformationen. Die Ergebnisse geben dem Verfahren hinsichtlich der numerischen Stabilität Recht. In einem sehr engen Rahmen bewegen sich die Gesamtgenauigkeiten der Parametrisierungsdatensätze: Die untere Grenze wurde bei 77,4 % erreicht, die obere bei 78,5 %. Für die Validierungsdatensätze sieht das Bild weniger einheitlich aus, dennoch ist ein deutlicher Trend erkennbar. Die Erststellergenauigkeiten für Traktecken ohne Totholz sind stets niedriger als jene der Ecken mit Totholzfunden und umfassen einen Bereich von 62,3–78,1 %. Ecken mit Totholz

7 Ergebnisse

Tabelle 7.11: Klassifikationsergebnisse der Modelle für das Merkmal *Totholzvorkommen* (ETM⁺-Daten)

Beobachtung ^b	Vorhersage ^a			% korr. (E)	Beobachtung ^b	Vorhersage ^a			% korr. (E)
	TH-Vork.	0	1			TH-Vork.	0	1	
Testgebiet 1 Nord ($\tau_p = 0,394$)					Testgebiet 1 Süd ($\tau_p = 0,498$)				
TH-Vork.	0	102	63	61,8	TH-Vork.	0	178	103	63,3
	1	37	128	77,6		1	38	243	86,5
% korr. (N)		73,4	67,0	(G) 69,7	% korr. (N)		82,4	70,2	(G) 74,9
Testgebiet 2 Ost ($\tau_p = 0,626$)					Testgebiet 2 West ($\tau_p = 0,647$)				
TH-Vork.	0	188	69	73,2	TH-Vork.	0	148	53	73,6
	1	27	230	89,5		1	18	183	91,0
% korr. (N)		87,4	76,9	(G) 81,3	% korr. (N)		89,2	77,5	(G) 82,3
Testgebiet 3 Nord ($\tau_p = 0,791$)					Testgebiet 3 Süd ($\tau_p = 0,714$)				
TH-Vork.	0	35	8	81,4	TH-Vork.	0	35	7	83,3
	1	1	42	97,7		1	5	37	88,1
% korr. (N)		97,2	84,0	(G) 89,5	% korr. (N)		87,5	84,1	(G) 85,7
Testgebiet 4 Ost ($\tau_p = 0,448$)					Testgebiet 4 West ($\tau_p = 0,537$)				
TH-Vork.	0	134	89	60,1	TH-Vork.	0	107	57	65,2
	1	34	189	84,8		1	19	145	88,4
% korr. (N)		79,8	68,0	(G) 72,4	% korr. (N)		84,9	71,8	(G) 76,8
Testgebiet 5 Nord ($\tau_p = 0,704$)					Testgebiet 5 Süd ($\tau_p = 0,553$)				
TH-Vork.	0	141	45	75,8	TH-Vork.	0	452	259	63,6
	1	10	176	94,6		1	59	652	91,7
% korr. (N)		93,4	79,6	(G) 85,2	% korr. (N)		88,5	71,6	(G) 77,6

^a laut Regressionsmodell

^b laut BWI-Datenbank

TH-Vork. = Totholzvorkommen: 0 = keine Funde, 1 = Totholzfunde (stehend oder liegend)

% korr. = korrekte Klassifikationen in Prozent: E = Ersteller-, N = Nutzer-, G = Gesamtgenauigkeit

wurden mit Genauigkeiten zwischen 84,3 und 90,8 % richtig erkannt. Insgesamt ergaben sich aus diesen Werten Gesamtgenauigkeiten von 75,3–82,4 %.

Die Auswirkung der unterschiedlichen geschätzten Koeffizienten je Parametrisierungsdatensatz lässt sich für *Totholz* nicht so leicht interpretierbar visualisieren wie beim Merkmal *Waldstatus*. Erschwerend wirkt hierbei die Verwendung von insgesamt acht Prädiktoren, die es unmöglich machen, einzelne Grauwertschwellen zu identifizieren, oberhalb derer einem Pixel der Status *Totholz vorhanden* zugewiesen wird.

In Abbildung 7.8 auf der nächsten Seite ist das unterschiedliche Verhalten der zehn Modellgleichungen dargestellt. Zunächst wurde basierend auf den Grauwertkombinationen aller Traktecken

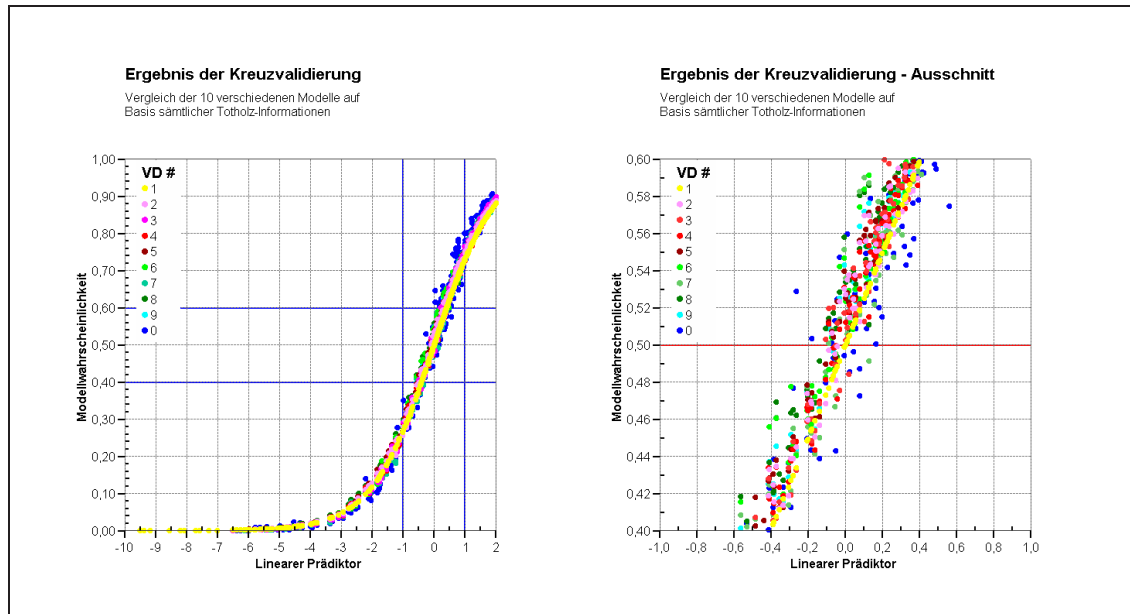


Abbildung 7.8: Kreuzvalidierung des Totholzgesamtmodells

des Validierungsdatensatzes #1 mit den zugehörigen Modellkoeffizienten⁸ der Wert der linearen Prädiktorenkombination errechnet, die den Exponenten der Euler'schen Zahl in Gleichung 3.1 auf Seite 58 bildet. Dieser Schritt sollte später einen Vergleich der Modelle ermöglichen.

Weiter wurde die Wahrscheinlichkeit für das Auftreten von Totholz abhängig von den Grauwertkombinationen der Traktecken mit den zehn Modellgleichungen ermittelt. Als Resultat ergaben sich somit für jede Traktecke zehn leicht differierende Wahrscheinlichkeiten.

Abschließend wurden die zehn Wahrscheinlichkeiten pro Traktecke über dem Wert des linearen Prädiktors—der Summe aller Koeffizient/Prädiktor-Paare plus Konstante—aufgetragen. Da diese Summen mit den Koeffizienten und Grauwerten des Modells #1 ermittelt wurden, folgt aus Gleichung 3.1 ein „perfekter“ sigmoider Kurvenverlauf für Modell #1. Es bildet gewissermaßen die Referenz für die anderen neun Modelle.

Im linken Teil der Abbildung wird der Wertebereich umfassend dargestellt; es ist gut zu erkennen, dass sich die Wahrscheinlichkeiten der Modelle #2–10 eng an den Verlauf des ersten Modells anpassen. Ein und dieselbe Grauwertkombination in allen sieben Kanälen (und damit zwangsläufig gleichem NDVI) bewirkt also auch bei unterschiedlicher Datengrundlage sehr ähnlich klassifizierende Modelle. Um die relevanten Unterschiede im Bereich um die Schwellenwahrscheinlichkeit $P = 0,5$ besser abschätzen zu können, wurde im rechten Teil der Abbildung eine Ausschnittsvergrößerung vorgenommen. Eine ausgeprägtere Streuung wird deutlich; allerdings sollte man dabei die Achsenskalierung beachten!

Obwohl die Ergebnisse zeigen, dass z. T. eine gute Klassifizierung von Totholz mit dem vorliegenden Verfahren möglich ist, gibt die Streuung der Ergebnisse Anlass zum Verwerfen des Ansatzes im Zusammenhang mit dem angedachten Verwendungszweck. Für Stratifizierungen sind Fehlklassifikationen im Bereich von 40 % für den schlechtesten Fall nicht zielführend, da die Effizienz des Inventurdesigns zu weit absinkt. Damit einher geht die Steigerung des Stichprobenfehlers. Bringt

⁸ die in diesem Fall auf den Teildatensätzen 2–10 beruhten

die Einbindung der FE-Daten aber keine Vorteile mehr (Kostensenkung oder Genauigkeitssteigerung), ist die Anschaffung der Daten eine Fehlinvestition.

Im Kriterienkatalog der Ministerkonferenz zum Schutze der Wälder in Europa (MCPFE) kommt Totholz als fünfter Indikator für das Kriterium 4 „Maintenance, Conservation and Appropriate Enhancement of Biological Diversity in Forest Ecosystems“ zum Einsatz (MINISTERIAL CONFERENCE ON THE PROTECTION OF FORESTS IN EUROPE, LIAISON UNIT VIENNA, 2002). Bei der Umsetzung von FFH-Richtlinien kann es ebenso notwendig sein, schnell und preiswert Informationen über Totholz zu sammeln. In diesen Bereichen des Naturschutzes mögen die erreichten Genauigkeiten eventuell hinreichend sein, aber selbst dann ist die schwierige Validierbarkeit der Ergebnisse hinderlich. Es sei daran erinnert, dass viele der Totholzvorkommen weder im Luftbild noch in einer hoch aufgelösten Satellitenszene auffindbar waren. Eine Überprüfung der Totholzkartierung wäre dann wieder auf terrestrische Überprüfungen angewiesen, die man mit dem Verfahren erübrigen wollte. Darüber hinaus erfordert die Parametrisierung der Modelle a priori Wissen über Totholzvorkommen, die ebenso auf terrestrische Erhebungen oder extrem hoch aufgelöste Orthophotos aufbauen müssten (FREI ET AL., 2003). Hier bedürfte es also eines kombinierten Ansatzes.

Kapitel 8

Zusammenfassung

In dieser Arbeit wurde die Eignung logistischer Regressionsmodelle zur fernerkundungsbasierten Klassifizierung bzw. Kartierung der folgenden forstlichen Parameter untersucht:

- Waldfläche/Waldvorkommen,
- Waldmischung und
- Totholzvorkommen.

Fernerkundungsdaten der optischen Systeme Landsat (Thematic Mapper und Enhanced Thematic Mapper Plus) und QuickBird (Ball Aerospace High-resolution Camera 60) wurden als Klassifizierungsgrundlage in fünf Testgebieten eingesetzt. Diese Testgebiete beinhalten 11 % der deutschen Waldfläche und wurden so gewählt, dass möglichst unterschiedliche Verhältnisse und Fragestellungen abgedeckt werden konnten. Dazu zählten Fragmentierungsgrad, Topographieeffekte und unterschiedliche Baumarten.

8.1 Allgemeine Eigenschaften des Verfahrens

Unabhängig vom betrachteten Merkmal wurden einige grundlegende Besonderheiten bei der Verwendung logistischer Regressionen deutlich. Das Verfahren ist anfällig für sehr unterschiedliche Trainingsdatenumfänge pro untersuchter Merkmalsklasse. Die Art der Anpassung der Modellparameter bewirkt in solchen Fällen eine unausgewogene Ausrichtung der Koeffizientenschätzer zu der/den stärker besetzten Klasse(n) hin, wodurch eine schlechtere Vorhersagegenauigkeit für die weniger zahlreich vertretene(n) Klasse(n) resultiert.

Die Verwendung von Informationen aus der Bundeswaldinventur-Datenbank zur Gewinnung von Trainingsdaten stellt hohe Anforderungen an die Lagegenauigkeit sowohl der BWI-Trakte selbst als auch der Satellitenszenen. Bei der Nutzung von Landsat-Daten sind kleinere Ungenauigkeiten unkritisch. Die Vorteile hoch auflösender Systeme treten bei Lageabweichungen jedoch in den Hintergrund, da falsche Verknüpfungen zwischen Datenbankinformationen und Bildgrauwerten erstellt werden können. In diesem Falle ist die Gewinnung von Trainingsdaten über visuelle Interpretation besser geeignet.

Wie alle pixelbasierten Verfahren ist die beschriebene Methodik anfällig für Bildrauschen, welches sich in der klassifizierten Szene durch Streupixel äußert. Die Anwendung von statistischen

Filtern verbessert zwar den optischen Eindruck der Karte, führt aber unter Umständen mit steigender Fragmentierung des klassifizierten Gebietes zu Verzerrungen, da richtig erkannte Kleinparzellen einer Klasse nach erfolgter Filterung fälschlicherweise der umgebenden Klasse zugeordnet werden.

8.2 Merkmal *Waldstatus*

Das Merkmal *Waldstatus* bzw. *Waldvorkommen* ließ sich mittels logistischer Regression auf Grundlage von Landsat-Daten mit Genauigkeit auf dem aktuellen Stand der Technik abbilden. In Abhängigkeit von Fragmentierungsgrad, Topographie und Aufnahmezeitpunkt der akquirierten Szene waren Modellgesamtgenauigkeiten zwischen 85,1 und 95,4 % erzielbar. Die Klassifizierungsgenauigkeiten der Waldmasken folgten dem Trend der Modellergebnisse, lagen aber zumeist auf etwas niedrigerem Niveau. Über Validierungsstichproben wurde ein Bereich von 84,7–94,6 % Übereinstimmung zwischen Klassifikation und Referenz ermittelt.

Die Verwendung von QuickBird-Daten erbrachte mit 83 % Modellgesamtgenauigkeit kein den Untersuchungen anderer Autoren vergleichbares Ergebnis. Mögliche Ursachen sind die eingangs erwähnten Lagedifferenzen und ein in Relation zur Gesamtpixelzahl sehr kleiner Trainingsdatensatz. Zusätzliche Untersuchungen an einem 3×3 km großen QuickBird-Testgebiet zeigten, dass durch im Bild gewonnene Trainingsdaten mit größerem Umfang (Stichprobenraten 4 bzw. ~ 75 %) erheblich bessere Modellgenauigkeiten von fast 99 % realisierbar waren. Die Überprüfung der Karte ergab eine Gesamtgenauigkeit von 94,1 %.

In einer möglichen dritten Bundeswaldinventur erscheint das Verfahren ungeeignet zur direkten Schätzung der Waldfläche, da Höhe und Variabilität der Klassifizierungsgenauigkeiten nicht innerhalb der strikten Fehlertoleranzen der BWI liegen. Ebenso wenig lässt sich eine Waldflächenveränderung statistisch abgesichert bilanzieren. In Deutschland liegt die Veränderungsrate bei unter einem Prozent der Waldfläche zwischen BWI I und BWI², was unerreichbare Anforderungen an die Genauigkeit der Karten beider Zeitpunkte stellt (FULLER ET AL., 2003). Größeres Potenzial ist der Methode als Stratifizierungsgrundlage oder als Schnelltestverfahren für kleinere Gebiete zuzuschreiben. Die (geringe) Zahl der Fehlklassifikationen lässt eine ausreichend hohe Stichprobeneffizienz zu, um den Einsatz der Waldmasken für ein stratifiziertes Stichprobenverfahren lohnenswert erscheinen zu lassen (RIEDEL, 2006).

8.3 Merkmal *Mischung*

Die Klassifizierbarkeit der drei Mischungsformen Nadel-, Misch- und Laubwald wurde anhand zweier Methoden mit ETM⁺-Daten untersucht. Weder das einstufige Verfahren über eine multinomiale logistische Regression noch der zweistufige Ansatz über zwei binomiale logistische Regressionen gestatteten eine zuverlässige Trennung aller drei Klassen. Gesamtgenauigkeiten zwischen rund 64 und 75 % täuschen über die niedrigen Erstellergenauigkeiten um 35 % für die Klasse *Mischwald* hinweg. Die Modelle ermöglichen demnach für diese Klasse eine dem Zufall nur geringfügig überlegene Genauigkeit.

Eine mögliche Modifikation des Verfahrens bestünde darin, die Mischungsunterteilung nur in zwei Klassen, nämlich Nadel- und Laub(misch)wald, vorzunehmen und anschließend anhand der Modell-Wahrscheinlichkeit Werte zwischen 0,3 und 0,7 als Mischwald zu führen. Dies bedarf aber

weiterer Untersuchungen.

Im Hinblick auf die forcierten Waldumbaubemühungen mit dem Ziel größerer Mischwaldanteile kann keinem der beiden untersuchten Regressionsansätze eine Eignung zur zuverlässigen Kartierung von Waldmischungsformen zugesprochen werden.

8.4 Merkmal *Totholzvorkommen*

Das auf den BWI-Trakten erfasste Totholz konnte über logistische Regressionsmodelle in den Landsat-ETM⁺-Szenen mit sehr unterschiedlichen Genauigkeiten erfasst werden. Der niedrigste Anteil übereinstimmender Klassifikationen wurde mit 69,7 % erreicht, der höchste mit 89,5 %. Die Modelle verhalten sich einzeln betrachtet numerisch stabil, so dass auch bei veränderter Datengrundlage konstante Ergebnisse zu erwarten sind. Bedenklich für einen operationellen Einsatz ist das große Intervall von ~ 20 %, in dem sich die Ergebnisse beim Vergleich zwischen den Modellen bewegen. Dadurch wird eine Fehlerquantifizierung erheblich erschwert. Validierungen erwiesen sich als undurchführbar, da stichprobenartige Tests zeigten, dass die Mehrzahl der Totholzfunde im Orthophoto oder QuickBird-Bild nicht erkennbar waren.

Geht man davon aus, dass an die Kartierung von Totholz als Indikator für Biodiversität weniger hohe Genauigkeitsforderungen gestellt werden, ist der vorgestellte Ansatz als schnelles Klassifizierungsinstrument denkbar, wenn Kenntnisse über Totholz zur Kalibrierung der Modelle vorliegen. Andernfalls müsste eine Kombination mit terrestrischen Erhebungen erfolgen, deren Kosten und Zeitaufwand kritisch zu prüfen wären.

8.5 Wahl der Testgebiete

Die Wahl geeigneter Testgebiete ist von entscheidender Bedeutung für die Aussage und Übertragbarkeit einer Machbarkeitsstudie. Es konnte gezeigt werden, dass mit einem 9 km² großen Testgebiet hohe Genauigkeiten sowohl bei der Modellierung als auch bei der Kartierung von Waldvorkommen erreichbar sind. Der dafür notwendige Aufwand bei Trainingsdatengewinnung, Modellierung und Validierung ist jedoch ungeeignet zur operationellen Umsetzung in einem bundesweiten Vorhaben. Während also ein sehr kleines Testgebiet zu einer Empfehlung der untersuchten Fernerkundungsmethode führt, zeigt sich ein anderes Bild, wenn man größere Gebiete mit variierenden Bestockungsverhältnissen und Waldstrukturen berücksichtigt.

Literaturverzeichnis

- AGRESTI, A. (1990): *Categorical Data Analysis*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 558 Seiten.
- ALBERTZ, J. (2001): *Einführung in die Fernerkundung—Grundlagen der Interpretation von Luft- und Satellitenbildern*. Wissenschaftliche Buchgesellschaft, Darmstadt, 2. überarbeitete und erweiterte Auflage, 249 Seiten.
- BACKHAUS, R. (2004): Anforderungsanalyse der Nutzung von satellitenbasierten Erdbeobachtungssystemen für die Umweltpolitik (SATUM). *Forschungsbericht 202 91 601, UBA-FB 000693*, Umweltbundesamt, Postfach 33 00 22, 14191 Berlin.
- BALL AEROSPACE & TECHNOLOGIES CORP. (a): Ball Global Imaging System 2000. Pressemitteilung im Internet, Stand: 6. April 2005.
URL: <http://www.ballaerospace.com/pdf/bgis.pdf>.
- BALL AEROSPACE & TECHNOLOGIES CORP. (b): Ball Commercial Platform 2000. Pressemitteilung im Internet, Stand 6. April 2005.
URL: <http://www.ballaerospace.com/pdf/bcp2.pdf>.
- BALL AEROSPACE & TECHNOLOGIES CORP. (c): Ball High Resolution Camera 60. Pressemitteilung im Internet, Stand: 6. April 2005.
URL: <http://www.ballaerospace.com/pdf/bhrc.pdf>.
- BALTES-GÖTZ, B. (2004): *Logistische Regressionsanalyse mit SPSS*. Universitätsrechenzentrum Trier.
- BAYERISCHE LANDESANSTALT FÜR WALD UND FORSTWIRTSCHAFT (2005): *Waldzustandsbericht 2005*. Bayerisches Staatsministerium für Landwirtschaft und Forsten, Ludwigstraße 2, 80539 München, 45 Seiten.
- BEAUBIEN, J., LATIFOVIC, R., CIHLAR, J. und SIMARD, G. (2001): BOREAS Follow-On DSP-01 Landsat TM Land Cover Mosaic of the BOREAS Transect. Data set. Available on-line [<http://www.daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A; Stand: 7. April 2006.
URL: www-eosdis.ornl.gov/BOREAS/FollowOn/guides/dsp01_tm_landcover_doc.html.
- BENGIO, Y. und GRANDVALET, Y. (2004): No Unbiased Estimator of the Variance of K-Fold Cross-Validation. *Journal of Machine Learning Research* **5**: 1089–1105.

- BIRK, R. J., STANLEY, T., SNYDER, G. I., HENNIG, T. A., FLADELAND, M. M. und POLICELLI, F. (2003): Government programs for research and operational uses of commercial remote sensing data. *Remote Sensing of Environment* **88**: 3–16.
- BLINN, C. E. (2005): *Increasing the Precision of Forest Area Estimates through Improved Sampling for Nearest Neighbor Satellite Image Classification*. Dissertation, Faculty of the Virginia Polytechnic Institute and State University.
- BORTZ, J. (2005): *Statistik für Human- und Sozialwissenschaftler*. Springer Medizin Verlag Heidelberg, 6., vollständig überarbeitete und aktualisierte Auflage, 882 Seiten.
- BRICKLEMYER, R. S., LAWRENCE, R. L. und MILLER, P. R. (2002): Documenting no-till and conventional till practices using Landsat ETM⁺ imagery and logistic regression. *Journal of Soil and Water Conservation* **57**(5): 267–271.
- BUNDESMINISTERIUM FÜR ERNÄHRUNG, LANDWIRTSCHAFT UND FORSTEN (1998a): Bundeswaldgesetz.
- BUNDESMINISTERIUM FÜR ERNÄHRUNG, LANDWIRTSCHAFT UND FORSTEN (1998b): Verordnung über die Durchführung einer zweiten Bundeswaldinventur (Zweite BWI-Verordnung). Bundesgesetzblatt Jahrgang 1998 Teil I Nr. 31.
- BUNDESAMT FÜR STATISTIK, SEKTION GEOINFORMATION (2004): Waldmischungsgrad der Schweiz. Dokumentation zur Waldmischungsgradkarte der Schweiz.
- CANTY, M. J., NIELSEN, A. A. und SCHMIDT, M. (2004): Automatic Radiometric Normalization of Multitemporal Satellite Imagery. *Remote Sensing of Environment* **91**(3–4): 441–451.
- CLARK, D. B., CASTRO, C. S., ALVARADO, L. D. A. und READ, J. M. (2004): Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data. *Ecology Letters* **7**: 52–59.
- COCHRAN, W. G. (1977): *Sampling Techniques*. John Wiley & Sons, New York, 428 Seiten.
- COENRADIE, B. (1991): Waldklassifizierung und Tothholzkartierung im Nationalpark Bayerischer Wald unter Verwendung von LANDSAT-TM- und digitalen Zusatzdaten. In *7. Nutzerseminar des Deutschen Fernerkundungsdatenzentrums der DLR - Tagungsband*, Nr. 91-09 in DLR-Mitteilungen, Seiten 46–49.
- CRACKNELL, A. P. (1998): Review Article—Synergy in remote sensing—what’s in a pixel? *International Journal of Remote Sensing* **19**(11): 2025–2047.
- DEVENTER, A. P. V., WARD, A. D., GOWDA, P. H. und LYON, J. G. (1997): Using Thematic Mapper Data to Identify Contrasting Plains and Tillage Practices. *Photogrammetric Engineering & Remote Sensing* **63**(1): 87–93.
- DIGITALGLOBE™ (a): Quickbird Specifications. Internetpräsenz der DigitalGlobe Incorporated, Stand 6. April 2005.
URL: <http://www.digitalglobe.com/about/quickbird.html>.

- DIGITALGLOBE™ (b): Quickbird Photos. Internetpräsenz der DigitalGlobe Incorporated, Stand 6. April 2005.
URL: <http://www.digitalglobe.com/press/Spacecraftimages.html>.
- DOBSON, A. J. (2001): *An Introduction to Generalized Linear Models*. Chapman and Hall, 2. Auflage, 245 Seiten.
- DONOGHUE, D. N. M., WATT, P. J., COX, N. J., DUNFORD, R. W., WILSON, J., STABLES, S. und SMITH, S. (2004): An evaluation of the use of satellite data for monitoring early development of young Sitka spruce plantation forest growth. *Forestry—An International Journal of Forest Research* **77**(5): 383–396.
- DRAPER, N. R. und SMITH, H. (1998): *Applied Regression Analysis*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 706 Seiten.
- EURIMAGE (2005): Eurimage Price List (May, 2005). Internetpräsenz von Eurimage, Stand 13. Mai 2005.
URL: <http://www.eurimage.com>.
- EUROMAP (2005): Price List (March 2005). Internetpräsenz von Euromap, Stand: 9 April 2005.
URL: http://www.euromap.de/products/prod_001.html.
- FITZPATRICK-LINS, K. (1981): Comparison of Sampling Procedures and Data Analysis for a Land-use and Land-cover Map. *Photogrammetric Engineering & Remote Sensing* **47**(3): 343–351.
- FOODY, G. M. (2005): Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing* **26**(6): 1217–1228.
- FRASER, R. H., FERNANDES, R. und LATIFOVIC, R. (2003): Multi-temporal mapping of burned forest over Canada using satellite-based change metrics. *Geocarto International (Hong Kong)* **18**(2): 37–47.
- FREI, U., KELLENBERGER, T., MEURET, P., LENCINAS, J., GINZLER, C. und SCHWARZ, M. (2003): Pilotprojekt LB-SNP 2000, Teilprojekt B: Digitale, automatisierte Auswerteverfahren. *Schlussbericht*, Remote Sensing Laboratories (RSL) & Eidgenössische Forschungsanstalt WSL.
- FRITZ, P. (Herausgeber) (2006): *Ökologischer Waldumbau in Deutschland—Fragen, Antworten, Perspektiven*. oekom verlag, Gesellschaft für ökologische Kommunikation mbH, 351 Seiten.
- FUCHS, H.-J. (2003): *Methodische Ansätze zur Erfassung von Waldbäumen mittels digitaler Luftbilddauswertung*. Dissertation, Georg-August-Universität Göttingen; Fakultät für Forstwissenschaften und Waldökologie.
- FULLER, R., SMITH, G. und DEVEREUX, B. (2003): The characterisation and measurement of land cover change through remote sensing: problems in operational applications? *International Journal of Applied Earth Observation and Geoinformation* **4**: 243–253.

- GALLEGO, F. J. (2004): Remote sensing and land cover area estimation. *International Journal of Remote Sensing* **25**(15): 3019–3047.
- GAUER, J. und ALDINGER, E. (Herausgeber) (2005): *Waldökologische Naturräume Deutschlands—Forstliche Wuchsgebiete und Wuchsbezirke—mit Karte 1 : 1.000.000*. Nr. 43 in Mitteilungen des Vereins für forstliche Standortskunde und Forstpflanzenzüchtung, Uhlman Verlag Stuttgart, 324 Seiten.
- GHEBREMICAEL, S., SMITH, C. und AHMED, F. (2004): Estimating the leaf area index (LAI) of black wattle from Landsat ETM+ satellite imagery. *Southern African Forestry Journal* **201**: 3–12.
- GOWDA, P. H., DALZELL, B. J., MULLA, D. J. und KOLLMAN, F. (2001): Mapping Tillage Practices with Landsat Thematic Mapper Based Logistic Regression Models. *Journal of Soil and Water Conservation* **56**(2): 14–18.
- HALL, F. G., STREBEL, D. E., NICKESON, J. E. und GOETZ, J. S. (1991): Radiometric Rectification: Toward a Common Radiometric Response Among Multidate, Multisensor Images. *Remote Sensing of Environment* **35**: 11–27.
- HARTUNG, J. und ELPELT, B. (1999): *Multivariate Statistik (Lehr- und Handbuch der angewandten Statistik)*. R. Oldenbourg Verlag München Wien, 6. unwesentlich veränderte Auflage, 815 Seiten.
- HILDEBRANDT, G. (1996): *Fernerkundung und Luftbildmessung für Forstwirtschaft, Vegetationskartierung und Landschaftsökologie*. Herbert Wichmann Verlag, Hüthig GmbH, Heidelberg, 680 Seiten.
- HOSMER, D. W., JR. und LEMESHOW, S. (1989): *Applied Logistic Regression*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 307 Seiten.
- HURD, J. D., CIVCO, D. L., GILMORE, M. S., PRISLOE, S. und WILSON, E. H. (2005): Coastal marsh characterization using satellite remote sensing and in situ radiometry data: Preliminary Results. In *Proceedings of 2005 ASPRS Annual Convention, Baltimore (Maryland)*.
- HUSEBY, R. B., HALCK, O. M. und SOLBERG, R. (2005): A model-based approach for geometrical correction of optical satellite images. *International Journal of Remote Sensing* **26**(15): 3205–3223.
- ITTEN, K. I., MEYER, P., KELLENBERGER, T., LEU, R., SANDMEYER, S., BITTER, P. und SEIDEL, K. (1992): *Correction of the Impact of Topography on Landsat-TM Forest Mapping of Alpine Regions, Remote Sensing Series*, Band 18. 48 Seiten.
- IVITS, E. und KOCH, B. (2003): Optimierung der Erfassung der Landschaftsdiversität auf der Basis von Satelliten- und Luftbildern: ein Europäisches Landschaftskonzept. *IÖR Schriften Band 40*.
- IVITS-WASSER, E. (2004): *Potential of Remote Sensing and GIS as Landscape Structure and Biodiversity Indicators—Methodological Study Relating Field Data to Visually Interpreted*

- and Segmented Landscape Objects and Image Grey Values*. Dissertation, Albert-Ludwigs-Universität Freiburg im Breisgau; Fakultät für Forst- und Umweltwissenschaften.
- JENSEN, J. R. (2005): *Introductory Digital Image Processing—A Remote Sensing Perspective*. Prentice Hall series in geographic information science, Pearson Prentice Hall, Upper Saddle River, NJ, 3. Auflage, 526 Seiten.
- JÄHNE, B. (2002): *Digitale Bildverarbeitung*. Springer-Verlag Berlin Heidelberg New York, 5. überarbeitete und erweiterte Auflage.
- JU, J., KOLACZYK, E. D. und GOPAL, S. (2002): Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sensing of Environment* **84**: 550–560.
- KEIL, M., MOHAUPT-JAHR, B., KIEFL, R. und STRUNZ, G. (2003): Update of the CORINE Land Cover Data Base in Germany. In JÜRGENS, C. (Herausgeber), *Proceedings of the ISPRS WG VII/4 Symposium on Remote Sensing of Urban Areas, Regensburg*, Seiten 72–77.
- KÖHL, M. (1994): *Statistisches Design für das zweite Schweizerische Landesforstinventar: Ein Folgeinventurkonzept unter Verwendung von Luftbildern und terrestrischen Aufnahmen*, Mitteilungen der Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft, Band 69, Heft 1. Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft Birmensdorf, 141 Seiten.
- KLEINBAUM, D. G. und KLEIN, M. (2002): *Logistic Regression—A Self-Learning Text*. Statistics for Biology and Health, Springer-Verlag Inc., New York, 2. Auflage, 513 Seiten.
- KOCH, B., JOCHUM, M., IVITS, E. und DEES, M. (2003): Pixelbasierte Klassifizierung im Vergleich und zur Ergänzung zum objektbasierten Verfahren. *Photogrammetrie Fernerkundung Geoinformation* **3**: 195–204.
- LALIBERTE, A. S., RANGO, A., HAVSTAD, K. M., PARIS, J. F., BECK, R. F., MCNEELY, R. und GONZALEZ, A. L. (2004): Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* **93**: 198–210.
- LANDESFORSTANSTALT EBERSWALDE, FACHBEREICH WALDENTWICKLUNG/MONITORING (2004): *Waldzustandsbericht 2004 der Länder Brandenburg und Berlin*. Ministerium für Ländliche Entwicklung, Umwelt und Verbraucherschutz, Abteilung Forst, Referat 42, 101 Seiten.
- LANDESFORSTPRÄSIDIUM SACHSEN (2003): *Forstbericht der Sächsischen Staatsregierung, Berichtszeitraum 1. Januar 1998 bis 31. Dezember 2002*. Sächsisches Staatsministerium für Umwelt und Landwirtschaft, Postfach 10 05 10, 01075 Dresden, 112 Seiten.
- LANDESFORSTPRÄSIDIUM SACHSEN (2005): *Waldzustandsbericht 2005*. Sächsisches Staatsministerium für Umwelt und Landwirtschaft (SMUL), Postfach 10 05 10, 01075 Dresden, 38 Seiten.

- LANE, P. W. (2002): Generalized linear models in soil science. *European Journal of Soil Science* **53**(2): 241–251.
- LAWRENCE, R., BUNN, A., POWELL, S. und ZAMBON, M. (2004): Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment* **90**: 331–336.
- LEICA GEOSYSTEMS (2003): *Erdas Field Guide*TM. GIS and Mapping, LLC, Atlanta, Georgia, 7. Auflage.
- LILLESAND, T. M., KIEFER, R. W. und CHIPMAN, J. W. (2004): *Remote Sensing and Image Interpretation*. John Wiley & Sons, Inc., 5. Auflage, 784 Seiten.
- LLOYD, C. J. (1999): *Statistical Analysis of Categorical Data*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 468 Seiten.
- MAGNUSSEN, S., BOUDEWYN, P. und WULDER, M. (2004): Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing* **25**(12): 2421–2440.
- MARBURGER, J. H. I. (2004): Landsat Data Continuity Strategy. Memorandum des Executive Office of the President, Office of Science and Technology Policy, Washington, D.C. (20502).
- MCCULLAGH, P. und NELDER, J. A. (1989): *Generalized Linear Models, Monographs on Statistics and Applied Probability*, Band 37. Chapman and Hall, 2. Auflage, 511 Seiten.
- MENARD, S. (2001): *Applied Logistic Regression Analysis*. Nr. 07-106 in Sage University Papers Series on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage, 111 Seiten.
- MIKHAIL, E. M., BETHEL, J. S. und MCGLONE, J. C. (2001): *Introduction to Modern Photogrammetry*. John Wiley & Sons, Inc.
- MINISTERIAL CONFERENCE ON THE PROTECTION OF FORESTS IN EUROPE, LIAISON UNIT VIENNA (2002): Improved Pan-European Indicators for Sustainable Forest Management as adopted by the MCPFE Expert Level Meeting 7-8 October 2002, Vienna, Austria. Herunterzuladen unter <http://www.mcpfe.org>. Stand: 20. Juni 2006.
- MÄKELÄ, H. und PEKKARINEN, A. (2004): Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. *Forest Ecology and Management* **196**: 245–255.
- NASA (1998): Landsat 7 Science Data Users Handbook. Online-Dokumentation im Internet, Stand 9. April 2005.
URL: http://ltpwww.gsfc.nasa.gov/IAS/handbook/handbook_toc.html.
- NASA (1999): Landsat 7. Bilderarchiv im Internet, Stand 6. April 2005.
URL: <http://landsat.gsfc.nasa.gov/images/hardware.html>.
- NELDER, J. A. und WEDDERBURN, R. W. M. (1972): Generalized Linear Models. *Journal of the Royal Statistical Society A* **135**: 370–384.

- OEHMICHEN, K. (2006): *Erstellung von Waldmasken mittels eines Schwellenwertverfahrens*. Dissertation, Universität Hamburg. (in Vorbereitung).
- OVER, M., SCHÖTTKER, B., BRAUN, M. und MENZ, G. (2003): Relative Radiometric Normalisation of Multitemporal Landsat Data—A Comparison of Different Approaches. In *Proceedings of the IGARRS 2003 Symposium. 21-25 July. Toulouse (France)*.
- PODRUZYNY, S. R., SCHWARTZ, C. C., LAWRENCE, R. L. und LANDENBURGER, L. (2004): Mapping Habitat for a Threatened & Endangered Species: Grizzly Bears in the Greater Yellowstone Area. Pilot Project Activity Status Summary im Internet, Stand: 6. April 2005. URL: <http://remotesensing.montana.edu/ProgressReportofGrizzlyBearPilotStudy.htm>.
- POLLEY, H., HENNIG, P. und SCHWITZGEBEL, F. (2004): Ergebnisse und Methoden der zweiten Bundeswaldinventur: Holzvorrat, Holzzuwachs und Holznutzung. Seminar zur Bundeswaldinventur und zur Waldentwicklungs- und Holzaufkommensmodellierung, 16./17.11.2004, Göttingen: Foliensammlung.
- POLLEY, H., KÖHL, M., IOST, A., KROIHER, F., OEMICHEN, K. und RIEDEL, T. (2006): Bundeswaldinventur: Weiterentwicklung des Inventurverfahrens durch Integration von Erdbeobachtungsdaten (WEI-BWI). *Schlussbericht*, Bundesforschungsanstalt für Forst- und Holzwirtschaft und Universität Hamburg.
- PUISSANT, A., HIRSCH, J. und WEBER, C. (2005): The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing* **26**(4): 733–745.
- RACK, J. (2001): Forest/Nonforest Classification of Landsat TM Data For Annual Inventory Phase One Stratification. In REAMS, G. A., MCROBERTS, R. E. und VAN DEUSEN, P. C. (Herausgeber), *Proceedings of the second annual Forest Inventory and Analysis symposium; 2000 October 17-18; Salt Lake City, UT*, SRS-47, U.S. Department of Agriculture, Forest Service, Southern Research Station, Seiten 8–10.
- RAT DER EUROPÄISCHEN GEMEINSCHAFTEN (1992): Richtlinie 92/43/EWG (Zur Erhaltung der natürlichen Lebensräume sowie der wildlebenden Tiere und Pflanzen).
- RICHARDS, T., GALLEG0, J. und ACHARD, F. (2000): Sampling of forest cover change assessment at the pan-tropical scale. *International Journal of Remote Sensing* **21**(6-7): 1473–1490.
- RIEDEL, T. (2006): *Evaluierung alternativer Stichprobendesigns für die Bundeswaldinventur*. Dissertation, Universität Hamburg, Zentrum Holzwirtschaft. (in Vorbereitung).
- ROTHMAN, K. J. (1998): *Modern Epidemiology*. Lippincott Williams and Wilkins, 2. Auflage, 752 Seiten.
- SACHS, L. (2002): *Angewandte Statistik (Anwendung statistischer Methoden)*. Springer-Verlag Berlin Heidelberg New York, 10. überarbeitete und aktualisierte Auflage, 889 Seiten.
- SAGISCHEWSKI, H., KRÜGER, J., KOCH, B., CHMARA, S., NEUMANN, K., DONNARUMMA, F. und HIRSCHFELD, H. (2003): *Verfahrensentwicklung zur fernerkundungsbasierten Erstellung*

- des Waldverzeichnisses Thüringens — Abschlussbericht* —. Nr. 22/2003 in Thüringer Landesanstalt für Wald, Jagd und Fischerei, Mitteilungen, Thüringer Landesanstalt für Wald, Jagd und Fischerei, 120 Seiten.
- SAWAYA, K. E., OLMANSON, L. G., HEINERT, N., BREZONIK, P. L. und BAUER, M. E. (2003): Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. *Remote Sensing of Environment* **88**: 144–156.
- SCHARDT, M. (1990): *Verwendbarkeit von Thematic Mapper-Daten zur Klassifizierung von Baumarten und natürlichen Altersklassen*. Dissertation, Albert-Ludwigs-Universität Freiburg. Forschungsbericht, herausgegeben vom DLR.
- SCHARDT, M. (1998): Erfassung forstlicher Parameter mittels Landsat/TM-Daten. *AFZ/Der Wald* **24**: 1461–1463.
- SÄCHSISCHES STAATSMINISTERIUM FÜR LANDWIRTSCHAFT, ERNÄHRUNG UND FORSTEN (1992): Waldgesetz für den Freistaat Sachsen.
- SCHÜTT, P., SCHUCK, H. J. und STIMM, B. (1992): *Lexikon der Forstbotanik (Morphologie, Pathologie, Ökologie und Systematik wichtiger Baum- und Straucharten)*. Ecomed Verlagsgesellschaft mbH (Landsberg/Lech), Justus-von-Liebig-Str. 1, 8910 Landsberg/Lech, 581 Seiten.
- SCHWARZ, M., STEINMEIER, C., HOLECZ, F., STEBLER, O. und WAGNER, H. (2003): Detection of Windthrow in Mountainous Regions with Different Remote Sensing Data and Classification Methods. *Scandinavian Journal of Forest Research* **18**: 525–536.
- SMIATEK, G. (1993): *Erfassung der Flächennutzung mit Hilfe von Landsat/TM-Stichproben*. Dissertation, Universität Stuttgart, Institut für Navigation.
- SMITS, P. C., DELLEPIANE, S. G. und SCHOWENGERDT, R. A. (1999): Quality Assessment of image classification algorithms for land-cover mapping: a review and a proposal for a cost-based approach. *International Journal of Remote Sensing* **20**(8): 1461–1486.
- SPOT (2005): Spot Products & Services—Price list 2005. Herunterzuladen in der SPOT-Internetpräsenz, Stand: 9. April 2005.
URL: http://www.spotimage.fr/automne_modules_files/standard/public/p336_fileLINKEDFILE_Liste_de_Prix_A4US-mp.pdf.
- STÜMER, W. (2004): *Kombination von terrestrischen Aufnahmen und Fernerkundungsdaten mit Hilfe der kNN-Methode zur Kartierung von Waldökosystemen*. Dissertation, Technische Universität Dresden, Fakultät für Forst-, Geo- und Hydrowissenschaften (Tharandt).
- UNITED NATIONS (1997): Kyoto Protocol to the United Nations Framework Convention on Climate Change.
- UNITED NATIONS (2002): A Guide to the Climate Change Convention Process - Preliminary Second Edition.

- UNITED STATES GEOLOGICAL SURVEY (2003): Landsat Monthly Update, June 2003. Herunterzuladen in der Internetpräsenz des Landsat-Projektes, Stand: 9. April 2005.
URL: <http://landsat.usgs.gov/monthnews.html>.
- UNITED STATES GEOLOGICAL SURVEY (2004): Landsat Monthly Update, April 2004. Herunterzuladen in der Internetpräsenz des Landsat-Projektes, Stand: 9. April 2005.
URL: <http://landsat.usgs.gov/monthnews.html>.
- WANG, L., SOUSA, W. P., GONG, P. und BIGING, G. S. (2004): Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of Panama. *Remote Sensing of Environment* **91**: 432–440.
- WERNER, C. (2002): Einsatzmöglichkeiten räumlich hochauflösender Satellitenbilder für Landschaftsplanung und Naturschutz—Beispiele der Auswertungen von IRS-1C/D-Daten—. Habilitationsschrift an der Fakultät VII (Architektur Umwelt Gesellschaft) der Technischen Universität Berlin.
- WINTER, R. (1987): Überblick zu den Walduntersuchungen mit Satellitendaten. 2. DFVLR-Statusseminar: „Untersuchung und Kartierung von Waldschäden mit Methoden der Fernerkundung“. DFVLR-Tagungsband, 415–432.
- WUENSCH, K. L. (2005): Binary Logistic Regression with SPSS®. Internetpräsenz von Dr. Karl L. Wuensch, Übungsaufgaben und Anleitungen für statistische Fragestellungen, Stand: 27. Februar 2006.
URL: <http://core.ecu.edu/psyc/wuenschk/SPSS/SPSS-MV.htm>.
- WYNNE, R. H., ODERWALD, R. G., REAMS, G. A. und SCRIVANI, J. A. (2000): Optical Remote Sensing for Forest Area Estimation. *Journal of Forestry* **98**(5): 31–36.
- ZANONI, V., STANLEY, T., RYAN, R., PAGNUTTI, M., BALDRIDGE, B., ROYLANCE, S., SNYDER, G. und LEE, G. (2003): The Joint Agency Commercial Imagery Evaluation team: overview and IKONOS joint characterization approach. *Remote Sensing of Environment* **88**: 17–22.