

# Prognose ökologisch wichtiger Waldbodeneigenschaften mit Random Forest in der Nordwestschweiz : Vergleich der Vorhersagen mit wissensbasierter empirisch- statistischer Modellierung

Autor(en): **Herbst, Philipp / Mosimann, Thomas**

Objektyp: **Article**

Zeitschrift: **Geomatik Schweiz : Geoinformation und Landmanagement =  
Géomatique Suisse : géoinformation et gestion du territoire =  
Geomatica Svizzera : geoinformazione e gestione del territorio**

Band (Jahr): **108 (2010)**

Heft 4

PDF erstellt am: **20.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-236681>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

# Prognose ökologisch wichtiger Waldbodeneigenschaften mit Random Forest in der Nordwestschweiz

## Vergleich der Vorhersagen mit wissensbasierter empirisch-statistischer Modellierung

Flächendeckend hoch auflösende Karten von Waldbodeneigenschaften werden für die Waldentwicklung und für Umweltschutzbelange verstärkt benötigt. Die Bodeneigenschaften müssen prognostiziert werden. In dieser Studie erfolgt die Prognose der Bodeneigenschaften «Gründigkeit» und «Skelettgehalt» mit halbautomatisierten Random Forest Modellen. Die Ergebnisse werden anhand wissensbasierter Prognosen empirisch-statistischer Modelle flächendeckend validiert und beurteilt. Die Untersuchungen zeigen, dass Random Forest Modelle im Vergleich zu den wissensbasierten Vorhersagen gleichwertige Ergebnisse erzielen können. Sie zeigen aber auch deren Grenzen auf. Die Qualität der Stichprobe ist für einen rein automatisiert-statistischen Ansatz fast immer zu gering. Zuverlässige Prognosen sind bisher nur in Kombination mit einem wissensbasierten Modell möglich.

*Pour la gestion de la forêt et les besoins de la protection de l'environnement on a toujours plus besoin de cartes à haute résolution des aptitudes des sols boisés. La qualité des sols doit être déterminée par des pronostics. Dans cette étude la qualité des sols quant à leur «bonté» et à leur «teneur en squelette» est évaluée par des modèles semi-automatique Random Forest. Les résultats sont validés et jugés sur l'ensemble de la surface à l'aide de pronostics certifiés par des modèles empiriques et statistiques. Les recherches démontrent qu'en comparaison avec des pronostics certifiés des modèles Random Forest peuvent procurer des résultats équivalents. Mais elles démontrent également leurs limites. Pour une procédure purement automatisée et statistique la qualité de l'échantillon est presque toujours insuffisante. Jusqu'à présent des pronostics fiables sont seulement possibles avec un modèle certifié.*

Lo sviluppo dei boschi e per questioni di pressione dell'ambiente su tutto il territorio nazionale saranno necessarie delle carte in alta risoluzione relative alle proprietà del terreno del bosco. Le caratteristiche del terreno devono essere pronosticate. In questo studio da previsione delle caratteristiche del suolo «Gr» «Ske» sono base con dei modelli semi automatizzati di Random Forest. I risultati sono convalidati giudicati in base a previsioni scientifiche e si basano su demodè belli e empirici statistici. Ricerche mostrano che i modelli rendono foreste a confronto delle previsioni scientificamente basate portano a risultati affini. Però dimostrando anche di avere dei limiti. La qualità delle prove campione sono praticamente sempre troppo limitate per un approccio statistico puramente automatizzato. Le previsioni affidabili finora sono solo possibile in combinazione con un anello basato sulla scienza.

Ph. Herbst, Th. Mosimann

### Einleitung

Flächendeckend hoch auflösende Karten wichtiger Bodeneigenschaften fehlen in vielen Gebieten Europas, so auch in der Schweiz. Grossmassstäbige Waldbodenkarten werden jedoch für die Waldbewirtschaftung und den Umweltschutz verstärkt benötigt. Gleichzeitig sind detaillierte Feldkartierungen durch die zuständigen Institutionen nicht mehr finanzierbar. Die Bodeneigenschaften müssen also prognostiziert werden. Hierfür werden Zusammenhänge zwischen Standorteigenschaften und Bodeneigenschaften modelliert.

Im Rahmen eines mehrjährigen Projekts erfolgt für die gesamte Waldfläche des Kantons Basel-Landschaft (ca. 200 km<sup>2</sup>) die Entwicklung von Vorhersagemodellen und die Erstellung von digitalen Karten der Waldbodeneigenschaften Gründigkeit, Skelettgehalt, Vernässung, Azidität und Wasserspeichervermögen. In den entscheidungsbasierten Modellen werden empirisch-statistisch ermittelte Zusammenhänge mit Expertenwissen kombiniert.

Seit einigen Jahren gelangen neben den empirisch-statistischen Modellansätzen halb- und vollautomatisierte Prognosemodelle verstärkt zum Einsatz (vgl. z.B. Lagacherie et al. 2007). Mit dem immer rascher wachsenden Umfang der Datenbestände können diese Verfahren der automatisierten Mustererkennung eine zeit- und kosteneffektive Alternative für die Bereitstellung hoch auflösender Bodenkarten darstellen (vgl. Behrens & Scholten 2006). Deshalb wurden in der Pilotphase des Projekts die Bodeneigenschaften Gründigkeit und Skelettgehalt auch durch die Implementierung halbautomatisierter Random Forest Modelle prognostiziert. Trotz des vermehrten Einsatzes automatisierter Modelle mangelt es jedoch an Studien, bei denen eine detaillierte Überprüfung der Vorhersagen direkt im Prognoseraum erfolgt. Diese Möglichkeit besteht hier durch die flächendeckende Validierung und Prüfung der

Random Forest Vorhersagen anhand der wissenschaftlichen Gründigkeit- und Skelettgehaltsprognosen im Untersuchungsraum. Dieser Beitrag stellt Ergebnisse aus vier Pilotgebieten des gesamten Modellierungsraumes dar.

## Datengrundlage und Methodik

Für die Generierung der Random Forest Modelle steht eine unter MS Access umgesetzte Waldbodendatenbank zur Verfügung. Sie enthält alle im Bodenaufschluss beschriebenen Waldbodenprofile des Kantons Basel-Landschaft und der unmittelbaren Nachbargebiete. Somit liegen in rund 450 Fällen profil- und horizontbezogene Punktinformationen zu den Waldböden vor, von denen etwa 90

Standorte im Projekt feldbodenkundlich neu aufgenommen wurden. Für die Entwicklung der Vorhersagemodelle und für die Prognose stehen rund 30 Vorhersagegrößen (Prädiktoren) in Form von Vektor- und Rasterdatensätzen unter ArcGIS und SAGA zur Verfügung. Hierzu zählen neben umfangreich modellierten Reliefparametern, Informationen zum Gestein, zur Waldgesellschaft und besonderen Bestandstypen (vgl. Burnand & Hasspacher 1999), zur Höhenlage und weitere Positionsparameter. Mittels empirisch-statistischer Untersuchungen von Zusammenhängen zwischen Standortfaktoren und Bodeneigenschaften und der Analyse von Regressionsbeziehungen, wurden die für die wissenschaftliche Entwicklung der Entscheidungsbäume relevanten Vorhersagegrößen ermittelt. Diese Prädiktoren

werden für die Random Forest Modelle vollständig übernommen und zum Teil ergänzt. Eine Vergleichbarkeit der Ergebnisse zwischen den Modellen ist damit gewährleistet.

### Random Forest

Random Forest (Breiman 2001) ist ein Klassifizierungs- und Regressionsmodell in Form eines Ensembles einzelner, binärer und voneinander unabhängiger Entscheidungsbäume (vgl. Breiman et al. 1984, Liaw & Wiener 2002). Mittels *bagging* wird jeder dieser Entscheidungsbäume auf Basis einer zufällig ausgewählten Unterstichprobe des Trainingsdatensatzes entwickelt. Die Unterstichproben sind voneinander unabhängig und umfassen immer zwei Drittel der Trainingsdaten. Das übrige Drittel (*out of bag*) dient zur unabhängigen Validierung und Fehlereinschätzung des Ensembles (vgl. Breiman & Cutler 2004). Als Prognose werden im Modellierungsraum die Merkmalsausprägungen ausgegeben, welche im Ensemble auf Basis der erlernten Klassifizierungsregeln am häufigsten ausgewiesen werden. Der Random Forest Algorithmus wird in ein halbautomatisiertes Modell für die Vorhersage der Bodenmerkmale Gründigkeit und Gesamtskelettgehalt implementiert (Abb. 1). Die Prognosen werden auf die Rasterzellen übertragen und unter ArcGIS visualisiert, kartographisch umgesetzt und validiert.

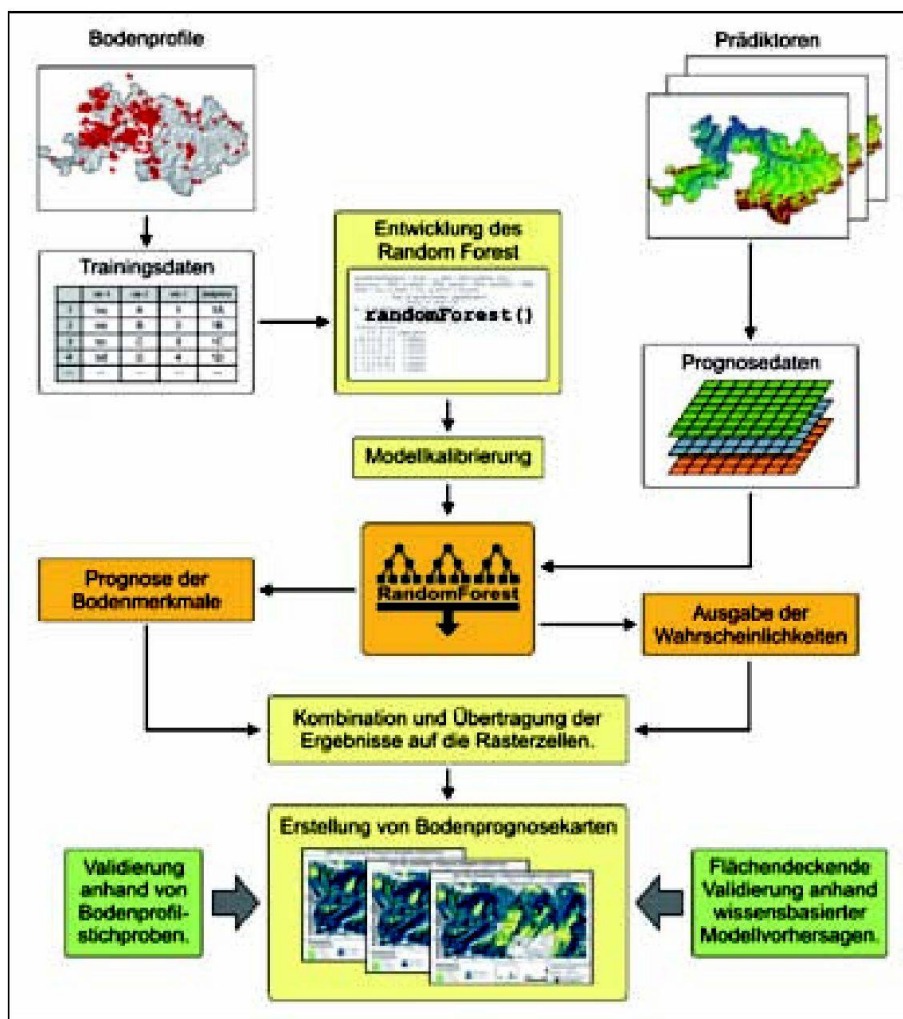


Abb. 1: Schema der Entwicklung und Validierung halbautomatisierter Random Forest Modelle.

### Vergleich von halbautomatisierten und wissenschaftlichen Modellergebnissen

Für jedes Untersuchungsgebiet liegen von Mosimann et al. (2007), auf langjähriger Kartiererfahrung und umfangreichen bodenkundlichen Neuaufnahmen von Waldstandorten gründende, wissenschaftliche Modelle (classification trees) für die Vorhersage der Bodeneigenschaften Gründigkeit und Skelettgehalt vor. Auf Grundlage dieser Vorhersagen werden die Prognosen der Random Forest Modelle flächendeckend einer lagebezogenen Plausibilitätsprüfung unterzogen. Zunächst erfolgt eine einheitliche Kategorisierung der Prognosen in vier bis fünf

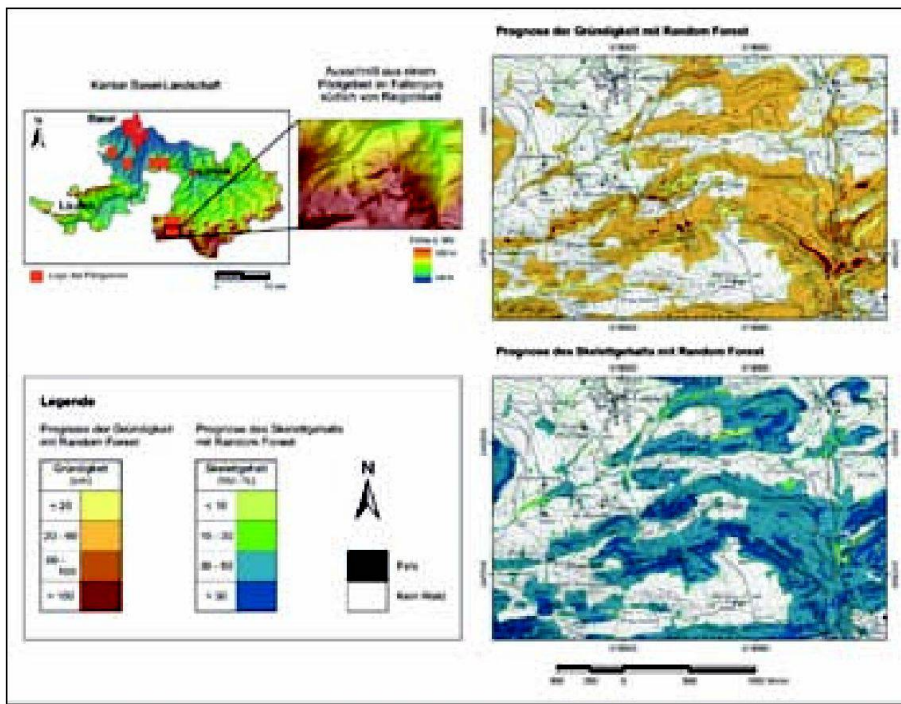


Abb. 2: Kartenausschnitte der Prognose von Gründigkeit und Skelettgehalt mit Random Forest.

Merkmalsklassen. Der explorative Vergleich der Prognosekarten und die Prüfung der räumlichen Verteilungsmuster der Bodeneigenschaften und ihrer Eintrittswahrscheinlichkeiten schliesst sich an. Folgend wird ein Kontrast-Index berechnet. Die Klassendifferenz veranschaulicht für jede Rasterzelle, wie stark die Random Forest Vorhersage von der wissensbasierten Prognose abweicht. Zudem werden die Random Forest Vorhersagen auf wissensbasiert nicht prognostizierbaren Flächen gesondert analysiert. Auf dieser Grundlage können Ergebniskarten und geeignete Histogramme der Flächenbilanzierungen erstellt werden. Sie dienen zur Identifizierung möglicher Regelmässigkeiten in der räumlichen Verbreitungsstruktur der Prognosedifferenzen.

## Ergebnisse und Diskussion

### Prognose von Gründigkeit und Skelettgehalt mit Random Forest

Im Lösshügelland werden fast flächendeckend hohe bis sehr hohe Gründigkeiten (> 80 cm) und sehr geringe Skelettgehalte ausgewiesen. Im Jura dominieren hingegen die Gründigkeiten < 40 cm. Die Verteilungsmuster sind im Lösshügelland und im Jura weitgehend nachvollziehbar. Beispielsweise weisen stark konvex gekrümmte Hangschultern sehr geringe Gründigkeiten auf, wohingegen in Tiefenlinien höhere Werte der Gründigkeit

prognostiziert werden. Einige Hangbereiche weisen jedoch auch nicht nachvollziehbare feingliedrige Verteilungsmuster auf. Die vielfältigen Prädiktorenkombinationen sind hierfür als massgebliche Ursache anzusehen. Ferner sind Fehler in der Prognose auszumachen. Im Jura wird beispielsweise in einem durch Gehänge schutt überdeckten steilen Hangbereich unterhalb der *Bürtenflue*, die höchste Gründigkeit im dortigen Pilotgebiet prognostiziert (Abb. 2). Dies gibt einen deut-

lichen Hinweis auf die zum Teil unzureichende Qualität der Trainingsdatensätze. Bei den Skelettgehaltsprognosen sind die grundlegenden Verteilungsformen und der Einfluss relevanter Prädiktoren im Jura nachvollziehbar. Im Bereich von Hangschultern, in Kammlagen und in Steilhängen werden die höchsten Skelettgehalte prognostiziert. In schwach geneigten Hangbereichen und Hochflächen der Tafeln werden geringere Skelettgehalte ausgewiesen. In den Hangbereichen sind jedoch teilweise nicht nachvollziehbare, häufig in Neigungsrichtung gestaffelte Verteilungsmuster festzustellen. Als mögliche Ursache für diese Strukturen sei eine mangelnde Qualität des digitalen Geländemodells genannt. Hier zeichnen sich Fehlerresiduen (Terrassierungseffekte) über die Reliefparameter in den Prognosen ab. In einem der Untersuchungsgebiete im Tafeljura wird auf fast 20% der Waldfläche ein Skelettgehalt < 10 Vol.-% prognostiziert. Diese auffälligen, weiträumigen Bereiche liegen auf den Hochflächen der Tafeln und zum Teil in schwach geneigten Hangbereichen. Aufgrund der vorherrschenden Bodenausgangssubstrate mit hoher räumlicher und Tiefen-

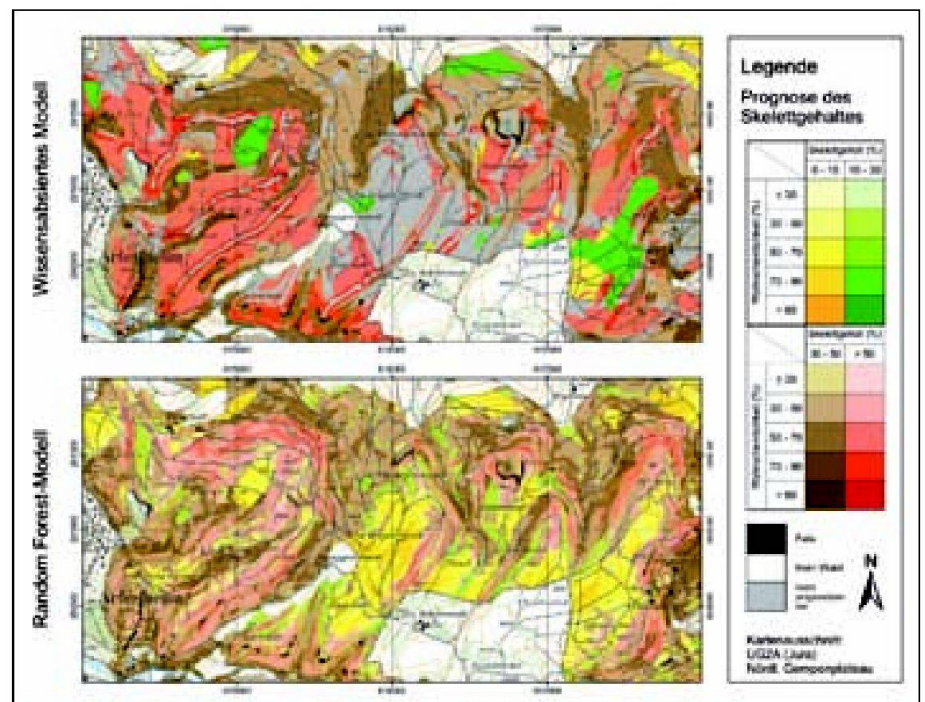


Abb. 3: Gegenüberstellung der Skelettgehaltsprognosen für ein Pilotgebiet im Tafeljura.

variation der Skelettgehalte, muss ein Teil dieser sehr geringen Skelettgehaltprognosen als falsch angenommen werden. Dieser Umstand ist ein deutlicher Hinweis auf die teilweise auftretenden Inkonsistenzen und Qualitätsmängel innerhalb der Trainings- und Prognosedatensätze. Kommt es zudem zu einer Überlagerung von Prädiktor- und Zielgrösseninformationen, die von der Stichprobe nur mangelhaft abgebildet werden, dann können falsche Prognosen, unabhängig von der Qualität des Klassifizierer-Algorithmus, nicht mehr ausgeschlossen werden. Es wurde erfolglos versucht, mögliche Einflussfaktoren oder Ursachen für diese Fehlprognosen in der Struktur der Datengrundlage auszuweisen. Die Stichprobe mit 400 untersuchten Standorten hat sich somit für einen rein automatisiert-statistischen Ansatz als unzureichend erweisen.

### Vergleich von Random Forest und wissensbasierten Modellvorhersagen

Bei den räumlichen Verteilungen der Random Forest Prognosen sind die grundlegenden Strukturen der wissensbasierten Prognosen erkennbar und nachvollziehbar. Grossflächige Verteilungsmuster, beispielsweise bedingt durch Tiefenbereiche oder einen einheitlichen Substrattyp, werden auch durch die Random Forest Modelle in sehr ähnlicher Struktur ausgewiesen. Die deutlichsten Übereinstimmungen räumlicher Verteilungsmuster sind im Löss festzustellen. Die geringste Übereinstimmung ist für die Skelettgehaltsprognosen im Jura festzuhalten (Abb. 3). Bei den wissensbasierten Modellen fallen die Prognosen der Bodeneigenschaften insgesamt deutlich strukturierter aus, als es bei den Random Forest Vorhersagen der Fall ist. Die Random Forest Modelle weisen innerhalb der Grossstrukturen kleinflächig stark variierende Prognosen auf. Vor allem bei den Skelettgehaltsprognosen im Jura sind nicht nachvollziehbare Hangmuster der prognostizierten Bodeneigenschaften feststellbar. Hierzu zählen u.a. bänderartige Wechsel der Skelettgehalte in Gefälle- richtung eines Hanges. Bei den wissensbasierten Vorhersagen treten sie nicht

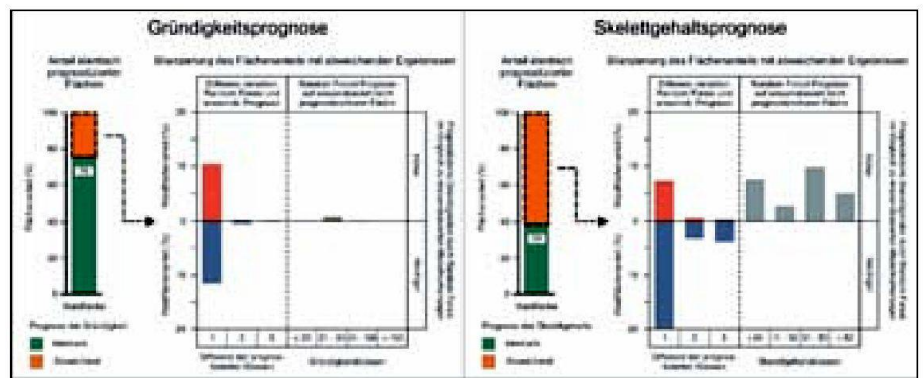


Abb. 4: Typische Beispiele der Flächenbilanzierung abweichender Modellprognosen aus dem Jura.

auf. Sie sind aufgrund der Reliefsituation und nach der Prüfung der Standortbedingungen auch als unplausibel einzu- stufen.

Die lagebezogene Prüfung der Prognose- differenzen erfolgt auf Basis der Klassen- differenzen zwischen den Random Forest Vorhersagen und den wissensbasierten Modellprognosen. Insgesamt ist festzu- stellen, dass der Anteil der Waldflächen mit abweichenden Vorhersagen von den Skelettgehaltsprognosen im Lösshügel- land (< 2 %), über die Gründigkeitsprog- nosen in beiden Bodenregionen (ca. 25%), bis hin zu den Skelettgehaltsprog- nosen im Jura (ca. 60%) deutlich zunimmt (Abb. 4). Bei den Gründigkeitsprognosen beträgt die Abweichung der Random For- est Vorhersagen in allen vier Unters- suchungsgebieten grösstenteils eine Klasse. Die Flächen mit abweichenden Progno- sen der Skelettgehalte fallen im Jura in- dessen wesentlich grossräumiger aus. Die Anteile der wissensbasiert nicht prognos- tizierbaren Bereiche betragen in beiden Untersuchungsgebieten in der Pilotphase ca. 25% der gesamten Waldfläche. Die empirisch-statistischen Modelle weisen dann nicht prognostizierbare Flächen aus, wenn für die Prädiktorenkombinationen unscharfe, ungleich verteilte oder zu klei-

ne Stichproben vorliegen, deren Häufig- keitsverteilungen wissensbasiert nicht be- stätigt werden können, oder wenn die Felduntersuchungen belegen, dass bei bestimmten Standortbedingungen die Bodeneigenschaften sehr kleinräumig wechseln und damit nicht vorhersagbar sind. Auf den prognostizierbaren Wald- flächen sind auf weiteren ca. 30 Flächen- % Prognosedifferenzen zu verzeichnen. Diese treten in den stärker geneigten Hangbereichen grossflächig auf (Abb. 5). Sie sind auf die insgesamt deutlich hetero- generen Random Forest Vorhersagen in den Hanglagen zurückzuführen. Die flä- chenbezogene Prognosequalität der Ran- dom Forest Modelle ist für den Skelett- gehalt im Jura damit stark eingeschränkt.

### Schlussfolgerungen

Die flächendeckende Prognose der Bo- denmerkmale Gründigkeit und Skelett- gehalt wird mit der Hilfe von Random For- est durchgeführt. Damit erfolgt die An- wendung einer *Machine Learning* Methode, über deren GIS-Kopplung und Verwendbarkeit für physisch-geographi- sche und landschaftsökologische Frage- stellungen bisher nur wenige Studien vor- liegen (vgl. z.B. Benito Garzón et al. 2006,

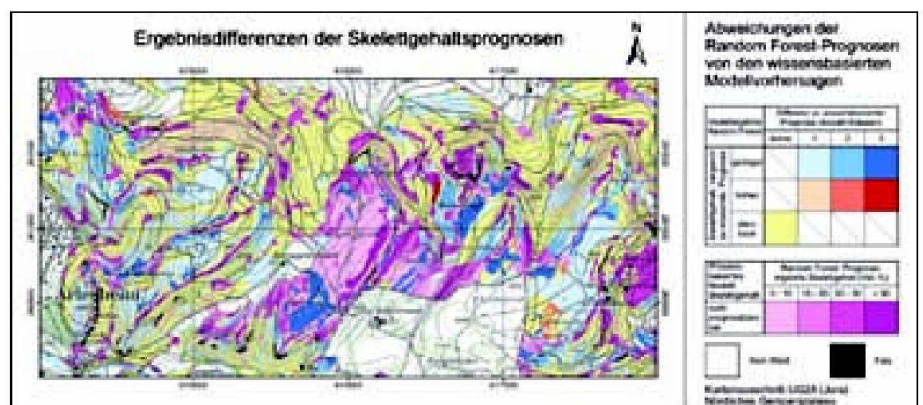


Abb. 5: Lagebezogene Ergebnisdifferenzen der Skelettgehaltsprognose für ein Pilotgebiet im Tafeljura.

Grimm et al. 2008). Die vorliegende Untersuchung belegt die grundsätzliche Anwendbarkeit der Random Forest Modelle, zeigt aber auch deren Grenzen auf. Die Ergebnisvalidierung verdeutlicht, dass Random Forest im Vergleich zu den wissensbasierten Vorhersagen empirisch-statistischer Modelle gleichwertige und plausible Ergebnisse erzielen kann, wenn die Zusammenhänge der Prädiktor- und Zielgrösseninformationen von der Stichprobe ausreichend abgebildet werden. Bei den Random Forest Vorhersagen im Jura sind jedoch gehäuft kleinräumige Wechsel und Einschränkungen in der Prognosequalität bis hin zu Fehlprognosen festzustellen. Hier zeigt sich, dass die Qualität der Stichprobe für einen rein automatisierten Algorithmus zu gering ist. Random Forest geht von einer Qualität der Stichprobe aus, die häufig nicht erfüllt wird.

Damit ergeben sich im Vergleich zu einem wissensbasiert empirisch-statistischen Modellansatz klare Nachteile bei der Anwendung von Random Forest (RF):

- RF erkennt nicht Inkonsistenzen in der Datengrundlage oder zu kleine Stichproben.
- RF ist nicht in der Lage Ausreisser und Sonderfälle zu erkennen.
- RF kann Zusammenhänge von Standortbedingungen und Bodeneigenschaften die für kleine Flächen gelten, nicht von denen differenzieren, die bei grossen Flächen bestehen.
- Für RF ist nicht erkennbar, ob die Stichprobe hinsichtlich der Flächenanteile der verschiedenen Prädiktorenkombinationen repräsentativ verteilt ist.

Damit liefert Random Forest auch Vorhersagen bei ungenügend gesicherten Zusammenhängen bis hin zu Fällen von Standortbedingungen, die nicht prognostizierbar sind.

Entsprechend wichtig sind die lagebezogenen Plausibilitätsprüfungen der Modellprognosen. Sie belegen die grosse Bedeutung der flächendeckenden Validierung von Random Forest Vorhersagen anhand wissensbasierter Modellprognosen. Nur so wird eine fundierte und differenzierte Aussage über die raumbezo-

gene Prognosequalität ermöglicht und können Fehlprognosen identifiziert werden. Die Annahme der Gültigkeit stichprobenartiger oder anhand von Testdatensätzen und -arealen ermittelter Validierungsergebnisse für den gesamten Prognoseraum, hat sich als nur eingeschränkt zulässig erwiesen. Die raumbezogene Prognosequalität von Random Forest Vorhersagen und von automatisierten Modellprognosen im Allgemeinen, sollte deshalb für jedes Untersuchungsgebiet individuell und möglichst flächendeckend ermittelt werden. Diesbezüglich mangelt es noch an Studien. Vor dem Hintergrund der wachsenden Nachfrage nach hochauflösenden und automatisierten Bodenprognosemodellen (vgl. Behrens et al. 2005, Behrens & Scholten 2006), besteht deshalb der dringende Bedarf, hierfür geeignete Validierungsverfahren weiter zu entwickeln. Zudem sollten Verfahren implementiert werden, um nicht prognostizierbare Fälle automatisiert ausweisen zu können (z.B. Schwellenwerte für Eintrittswahrscheinlichkeiten). Dadurch würde sich die Nachvollziehbarkeit der automatisierten Modellvorhersagen, insbesondere für Anwender ohne genauere Kenntnis des Algorithmus, insgesamt erhöhen. Mögliche Fehlinterpretationen der Ergebnisse können so minimiert werden.

Es bleibt festzuhalten, dass die Random Forest Prognosen der Waldbodeneigenschaften bisher nur in Kombination mit einem wissensbasierten Modell genügend zuverlässig sind. Zusätzliche umfangreiche Felduntersuchungen zur gezielten Ergänzung der Datenbasis sind deshalb unerlässlich, um die Modellqualität verbessern zu können.

#### Literatur:

- Behrens, T. & T. Scholten (2006): Digital soil mapping in Germany – a review. In: Journal of Plant Nutrition and Soil Science, Vol. 169, S. 434–443.
- Behrens, T., H. Förster, T. Scholten, U. Steinbrücken, E.-D. Spies & M. Goldschmitt (2005): Digital soil mapping using artificial neural networks. In: Journal of Plant Nutrition and Soil Science 168, S. 21–33.

Benito Garzón, M., R. Blazek, M. Neteler, R. Sanchez de Dios & H. Sainz (2006): Predicting habitat suitability with Machine Learning Models: the potential area of *Pinus sylvestris* in the Iberian Peninsula. In: Ecological Modelling 197, S. 383–393.

Breiman, L. & A. Cutler (2004): Random Forest Description.

URL [http://www.stat.berkeley.edu/users/breiman/RandomForests/cc\\_home.htm](http://www.stat.berkeley.edu/users/breiman/RandomForests/cc_home.htm)

Breiman, L. (2001): Random Forests. In: Machine Learning 45, Boston, S. 5–32.

Breiman, L., J.H. Friedman, R.A. Olshen & C.J. Stone (1984): Classification and Regression Trees. In: P.J. Bickel et al. [Hrsg.]: The Wadsworth Statistics/Probability Series. Belmont, 358 S.

Burnand, J. & B. Hasspacher (1999): Waldstandorte beider Basel. Kommentar zur vegetationskundlichen Standortkartierung der Wälder. In: Quellen und Forschungen zur Geschichte und Landeskunde des Kanton-Basel-Landschaft, Bd. 72, Liestal, 266 S.

Grimm, R., T. Behrens, M. Märker & H. Elsenbeer (2008): Soil organic carbon concentrations and stocks on Barro Colorado Island – Digital soil mapping using Random Forests analysis. In: Geoderma, Vol. 146, S. 102–113.

Lagacherie, P., McBratney, A. & M. Voltz [Hrsg.] (2007): Digital Soil Mapping – An Introductory Perspective. In: Developments in Soil Science, Vol. 33, Amsterdam, 600 S.

Liaw, A. & M. Wiener (2002): Classification and Regression by random Forest. In: R News, Vol. 2/3, S. 18–22.

Mosimann, Th., Meer, U. & J. Gross (2007): Dokumentation der Struktur der entscheidungsbasierten Modelle für die Vorhersage von Bodeneigenschaften. Unveröffentlichte Dokumentation, Institut f. Physische Geographie u. Landschaftsökologie, Universität Hannover.

Dipl.-Geogr. Philipp Herbst  
Prof. Dr. Thomas Mosimann  
Gottfried Wilhelm Leibniz Universität Hannover  
Institut für Physische Geographie und Landschaftsökologie  
Schneiderberg 50  
DE-30167 Hannover  
herbst@phygeo.uni-hannover.de  
mosimann@phygeo.uni-hannover.de