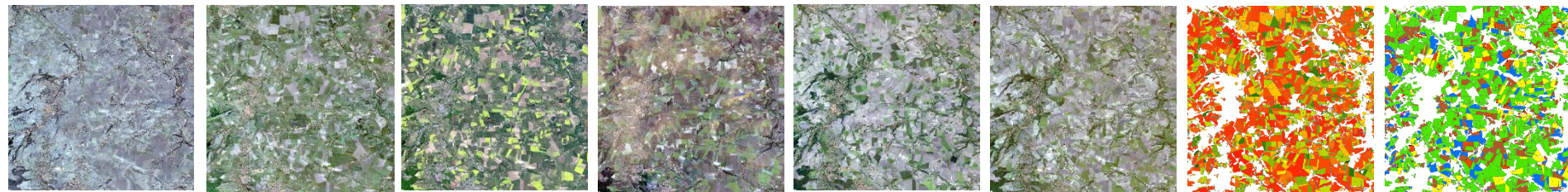


Einfluss stratifizierter Sampling Methoden auf die Genauigkeit multitemporaler objektbasierter Landwirtschaftsklassifikation

Tagung „Daten - Informationen - Entscheidungen“

4. Gemeinsamer Workshop der AKs

„Auswertung von Fernerkundungsdaten“ der DGPF e.V. und „Fernerkundung“ der DGfG e.V.



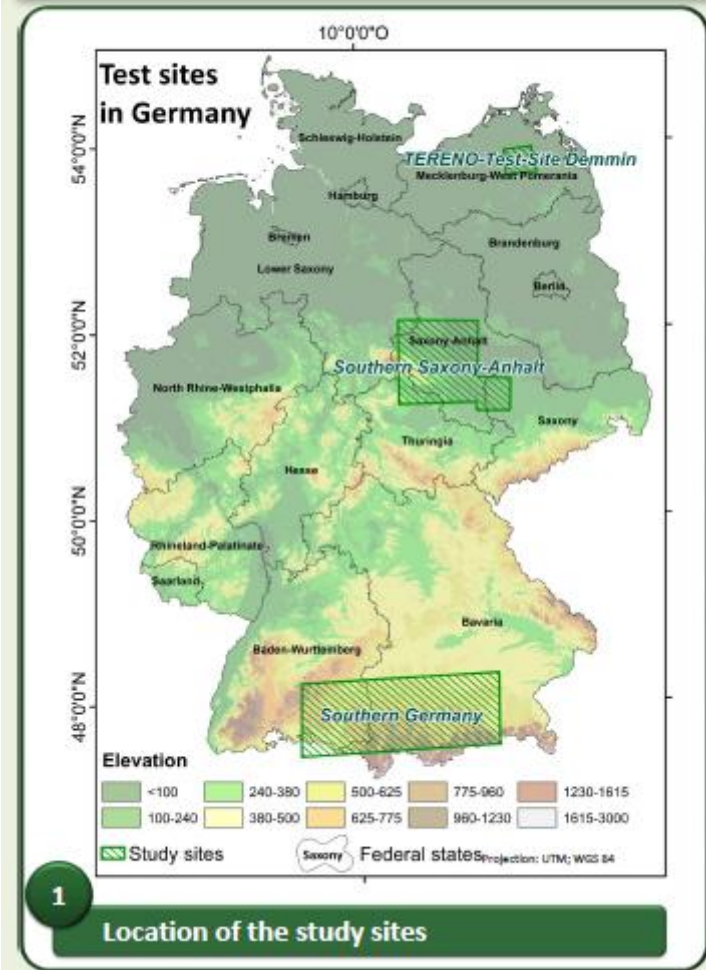
Jörn v. Uslar, Patrick Knöfel, Christopher Conrad

Universität Würzburg, Institut für Geographie und Geologie, Lehrstuhl für
Fernerkundung, Oswald-Külpe-Weg 86, 97074 Würzburg

Hauptziel des Projekts: **Klassifikation von landwirtschaftlichen Ackerfrüchten mit Sentinel-2-Daten**

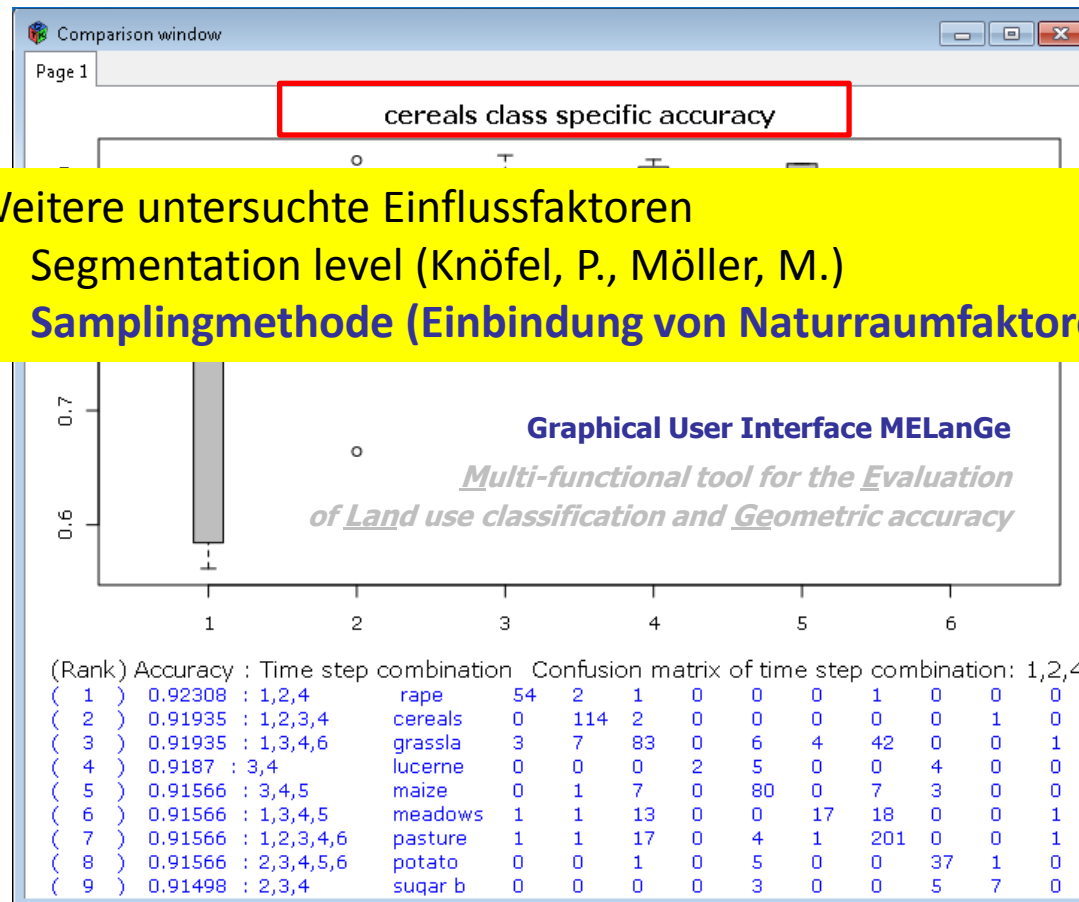
- Umfassende Modellierung der Phänologischen Entwicklungen von Agrarkulturen
- Simulation und Verwendung von Sentinel-2 Daten für neue Indizes
- Erfassung der relevanten Indizes und geeigneter phänologischer Zeitfenster für optimale Klassentrennung
- **Bewertung des Einflusses naturräumlicher Ausstattungen auf Klassifikation und deren Ergebnisse**

Study Sites



Projektpartner:

Einflüsse des Klassifikationsschlüssels auf die Genauigkeit und die Auswahl relevanter Zeitschnitte für die Klassifikation



Weitere untersuchte Einflussfaktoren

- Segmentation level (Knöfel, P., Möller, M.)
- **Samplingmethode (Einbindung von Naturraumfaktoren)**

Class-wise accuracies for winter barley and wheat and winter cereals, study area Demmin. The x-axis depicts the number of time steps, the y-axis the accuracy. The time step combinations in the lower left table refer to specific acquisition dates of remote sensing data (in this case: RapidEye)

Welchen Einfluss hat die naturräumliche Ausstattung (z.B. Topographie, Böden, etc.) auf die Klassifikationsergebnisse?

Zwei Perspektiven sind möglich:

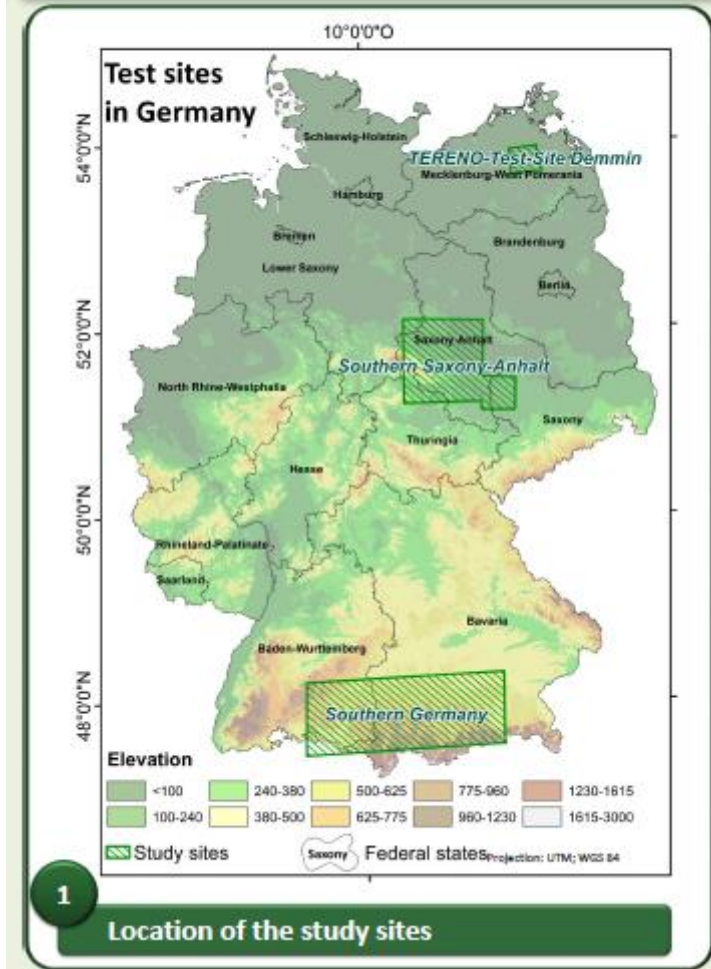
- Wie stark und in welcher Richtung (positiv/negativ) wirkt sich die Berücksichtigung einzelner Umweltfaktoren beim Sampling auf die Klassifikationsgüte aus?
 - ⇒ Stratified sampling
 - ⇒ Einbindung von Zusatzvariablen in die Klassifikation
- Wie sicher ist eine Klassifikationsentscheidung (für ein Pixel/Objekt)? Lassen sich Zusammenhänge der naturräumlichen Ausstattung/Lage mit der Treffsicherheit von Klassifikationsentscheidungen erkennen?

Mitteldeutsches Tiefland /Harzvorland



 Untersuchungsgebiet

Study Sites



Winterweizen- und Rapsfeld im Untersuchungsgebiet

Feldgrenzen:

- Anbauinformationen des Landes Sachsen-Anhalt (Kooperation mit Universität Halle)
- Jede Feldparzelle als Objekt mit zugehöriger Feldfrucht

Schlaggrenzen:

- Erzeugte Objekte der Segmentation
- Fasst beieinander liegende Parzellen mit gleichem Anbau als Objekt zusammen

DGM:

- Auflösung 20mx20m. Basierend auf dem offiziellen DGM 10 des Landes Sachsen-Anhalt

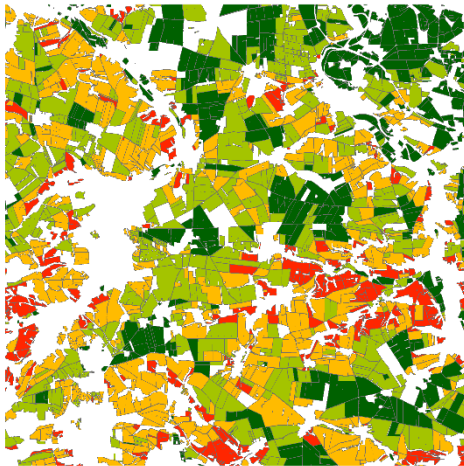
Bodenkarte:

- VBK des Landes Sachsen-Anhalt (2008)
- 97 Bodenarten und Untertypen im Untersuchungsgebiet

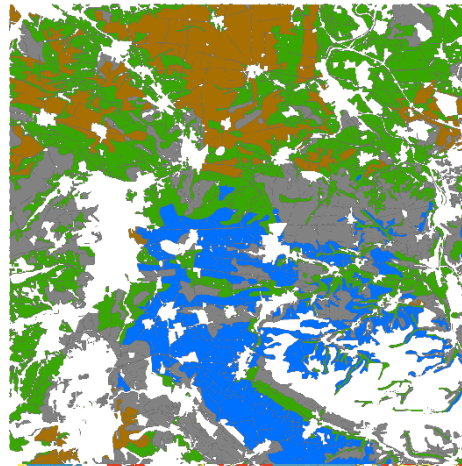
Daten auf die zu klassifizierenden Objekte anpassen: Schlaggrenzen

DGM

Slope, Aspect, Range

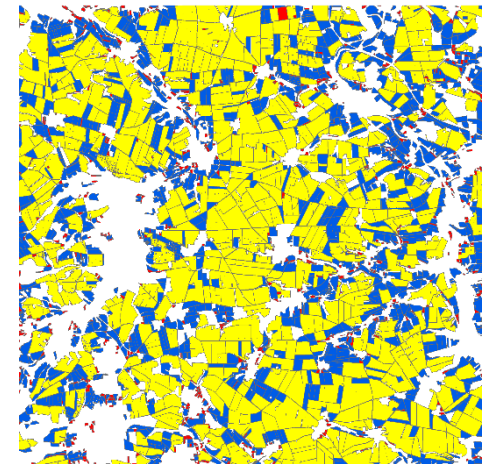


Bodenkarte



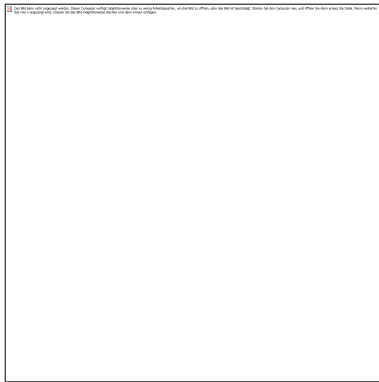
Feldparzellen

Größe

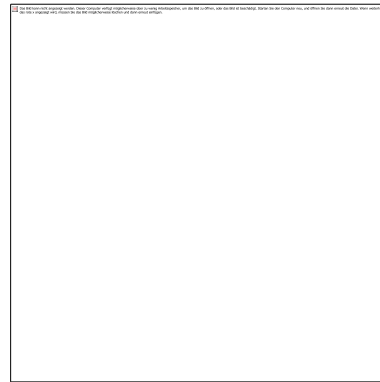


Rapid Eye:

- 5mx5m, Multispektral (Blue, Green, Red, Red Edge, Near Infrared)
- 2011, Multitemporal
- Atmosphärenkorrektur



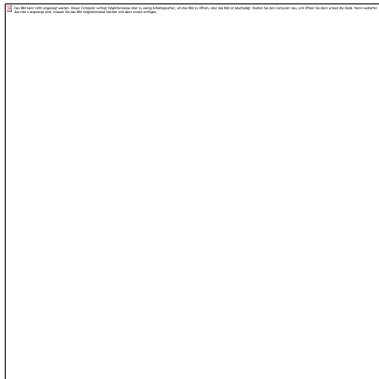
März



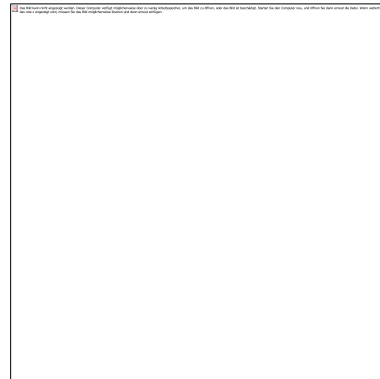
April



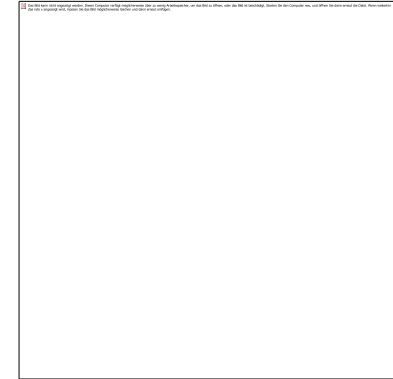
Mai



Juni



August



Oktober

Anpassen auf Feldobjekte

Berechnung 114 verschiedener spektraler Produkte:

- Mittelwerte sämtlicher Bänder
- Ratio verschiedener Bänder
- Indizes: NDVI, RNDVI, EVI, SAVI etc.

| Field_ID | NDVI_mean_ts1 | EVI_mean_ts1 | GI_mean_ts1 | BGI_mean_ts1 | RGI_mean_ts1 | DVI_me |
|----------|---------------|--------------|-------------|--------------|--------------|---------|
| 1 | -0.055409843 | 0.183800139 | 0.996004661 | 1.137661399 | 1.005651324 | -495.4 |
| 2 | -0.051924441 | 0.064735342 | 1.094539926 | 1.319944491 | 0.915174498 | -319.99 |
| 4 | -0.113093868 | 0.074958604 | 1.253047102 | 1.44464644 | 0.801143094 | -486.32 |
| 6 | -0.009022092 | 0.011001087 | 1.073373609 | 1.275555607 | 0.932896309 | -50.101 |
| 7 | -0.052124649 | 0.116391832 | 1.029850467 | 1.200328995 | 0.972974942 | -402.40 |
| 9 | -0.066215086 | 0.109453808 | 1.041425985 | 1.254398107 | 0.961885503 | -455.61 |
| 10 | -0.047105371 | 0.067877683 | 1.063644304 | 1.289824023 | 0.94138492 | -307.63 |
| 11 | -0.05518083 | 0.072765072 | 1.111698952 | 1.287856989 | 0.903463715 | -343.01 |
| 12 | -0.072852711 | 0.063950906 | 1.186353261 | 1.378519575 | 0.844587988 | -377.53 |
| 13 | -0.085336037 | 0.095521536 | 1.1248221 | 1.317528441 | 0.893755506 | -483.26 |
| 14 | -0.054586948 | 0.09358974 | 1.049747493 | 1.248807513 | 0.95391195 | -388.83 |
| 15 | -0.053146893 | 0.065185415 | 1.108936315 | 1.312287147 | 0.903863873 | -322.6 |
| 16 | -0.049365837 | 0.054008938 | 1.132949417 | 1.340350812 | 0.885300139 | -289.78 |

...

Überprüfung der Klassen auf spektrale Ähnlichkeit

Viele Klassen im Untersuchungsgebiet vorhanden:

Große Ähnlichkeit einiger Klassen

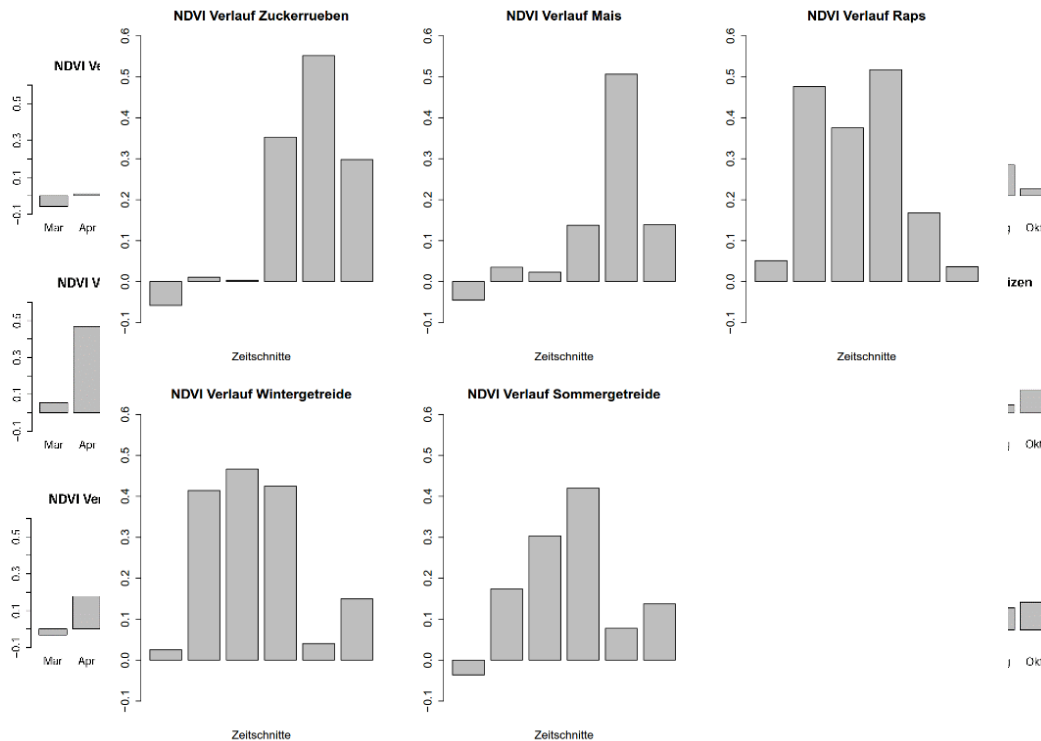
Teilweise sehr kleine Klassen

Reduktion der Klassenzahl

-> Schlechte Trennbarkeit

-> Zu geringe Objektanzahl für Klassifikation

-> Test für gut unterscheidbare Klassen



| Klasse | Anzahl |
|------------------------|---------|
| Weizen (Sommer/Winter) | 1250/51 |
| Raps | 416 |
| Winterroggen | 17 |
| Mais | 221 |
| Zuckerrüben | 142 |
| Kartoffeln | 24 |
| Hafer | 126 |
| Hartweizen | 95 |

Überwachte, objektbasierte Random Forest Klassifikation von multitemporalen optischen Satellitendaten (RapidEye)

Überwacht:

Zu ermittelnde Klassen sind bekannt

Objektbasiert (Turker and Arıkan 2005):

Vermeiden von Missklassifikation innerhalb der Felder

Multitemporale Klassifikation:

Unterscheidung vieler ähnlicher spektraler Signaturen notwendig

Multitemporaler Ansatz, um unterschiedliche phänologische Entwicklung zur Differenzierung zu nutzen

Random Forest (Breiman 2001) :

Hierarchische Klassifikationsbäume

Auswahl der Prädiktoren zufällig

N-fache Durchführung und Ermitteln der besten Ergebnisse durch majority vote (finale Klassenentscheidung pro Objekt)

Random Sampling:

Zufälliges Ziehen von 30% der gesamten Feldobjekte (2879) als Trainingsdaten

Stratifiziertes Sampling:

Jeweils 30% aus Teilmengen anstatt aus der Gesamtmenge ziehen:

- Zufälliges Ziehen z.B. aus der Menge der kleinen (1224), mittleren (1183) und großen Felder (472)
- Gleiche Vorgehensweise mit den anderen Straten und deren Kombinationen.
Z.B.: kleines Feld auf Braunerde mit Südexposition
- Gewährleistung sämtliche Aspekte der jeweiligen Klasse in den Trainingsdaten zu enthalten

| Object-ID | Landnutzung | Groesse | DGM-Klasse | Aspect | Slope-Klasse | Bodenklasse |
|-----------|--------------|---------|------------|--------|--------------|-------------|
| 1 | Zuckerrueben | mittel | 1 | 1 | 1 | 1 |
| 4 | Zuckerrueben | groß | 1 | 2 | 1 | 1 |
| 6 | Zuckerrueben | mittel | 1 | 2 | 1 | 4 |
| 7 | Zuckerrueben | groß | 2 | 3 | 1 | 2 |
| 9 | Zuckerrueben | mittel | 1 | 4 | 1 | 2 |
| 10 | Zuckerrueben | mittel | 1 | 2 | 1 | 2 |
| 12 | Zuckerrueben | mittel | 1 | 3 | 1 | 1 |
| 13 | Zuckerrueben | groß | 1 | 2 | 1 | 1 |
| 14 | Zuckerrueben | groß | 3 | 2 | 2 | 2 |

Durchführen und Vergleichen von Klassifikationen basierend auf den verschiedenen Samplingplänen

Zehnmaliges Durchführen des Samplings und der Klassifikation, um Aussagen über die Reproduzierbarkeit der Ergebnisse treffen zu können.

Genauigkeit:

Genauigkeit dargestellt durch die Overall, bzw. Classwise Accuracy:

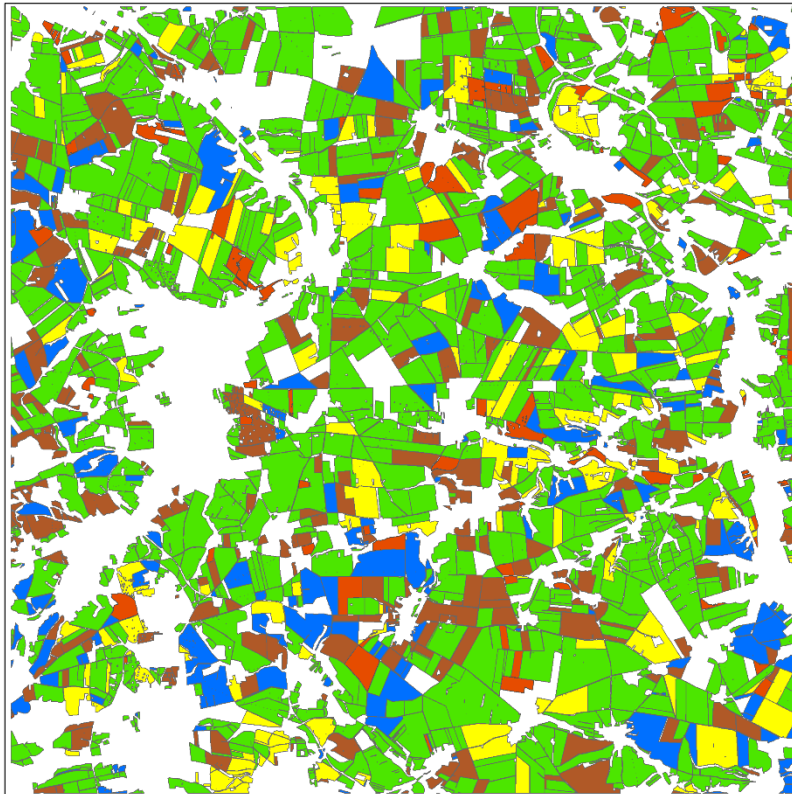
$$ACC = \frac{n_c}{n}$$

Robustheit:

- Stabilität der Klassifikationsergebnisse
- Range der Genauigkeit bei mehreren Läufen des Klassifikationsalgorithmus

Klassifikationsergebnis

Verwendete Straten: Aspect, Dgmclass, Groesse, Untergrund, Slope



Legende

Feldfrüchte

-  Mais
-  Raps
-  Sommergetreide
-  Wintergetreide
-  Zuckerrueben

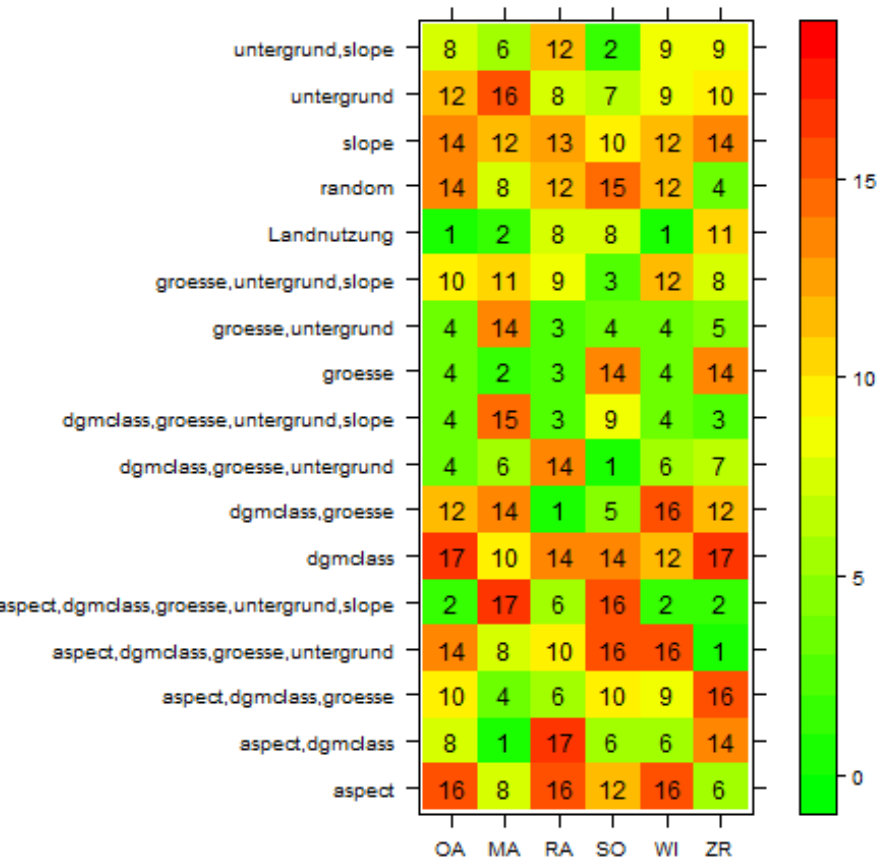
0 1.25 2.5 5 Kilometers



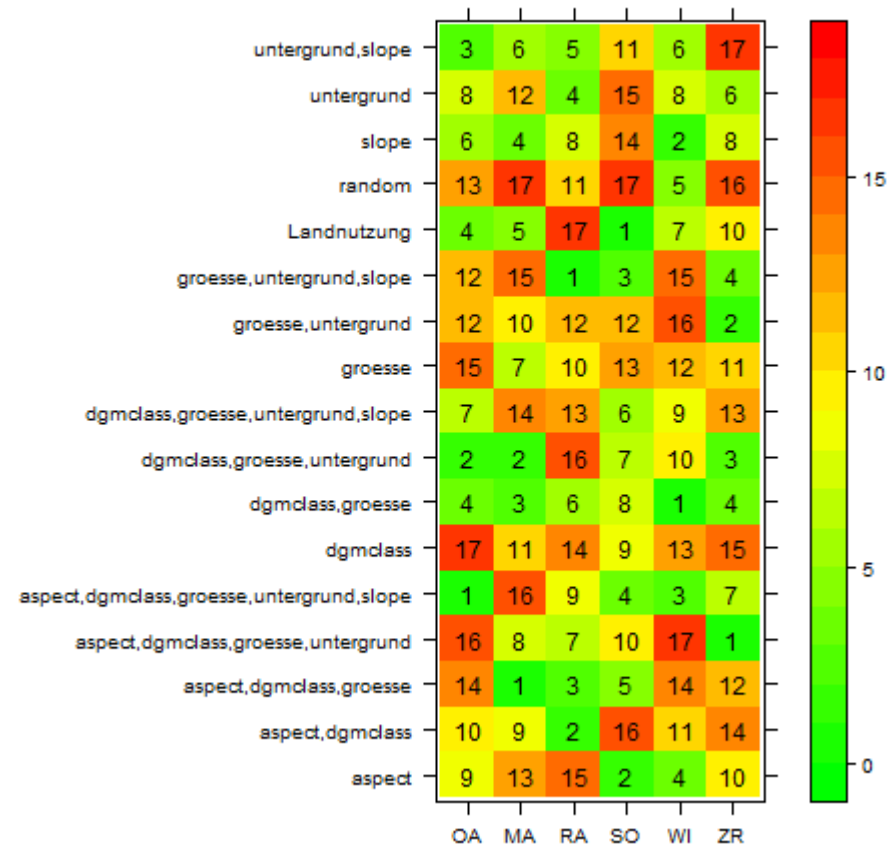
Accuracy

| | |
|----------------|-------|
| Overall | 93.7% |
| Mais | 92.7% |
| Raps | 92.7% |
| Sommergetreide | 73.9% |
| Wintergetreide | 95.7% |
| Zuckerrüben | 97.0% |

Accuracy der einzelnen Klassen, sortiert



Range der einzelnen Klassen, sortiert

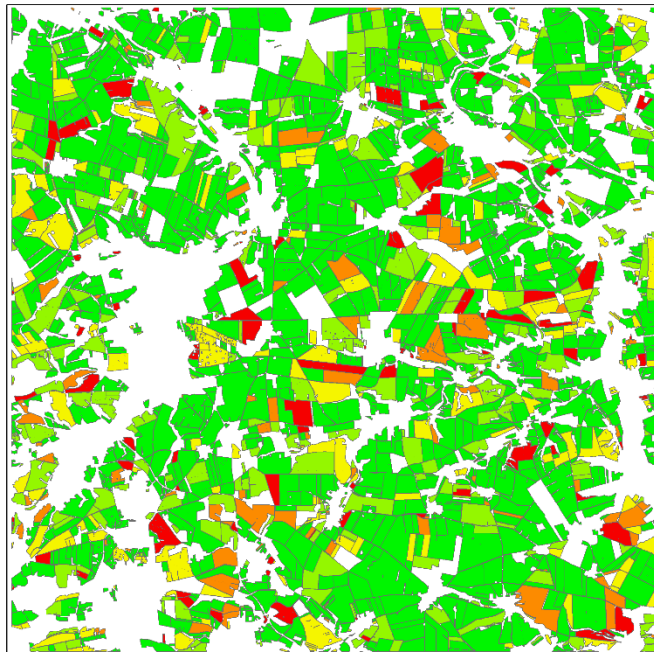


OA: Overall, MA: Mais, RA: Raps, SO: Sommergetreide, WI: Wintergetreide, ZR: Zuckerrüben

Wie unterscheiden sich die verschiedenen Samplingmethoden in der Sicherheit der Klassifikation?

Klassifikationssicherheit

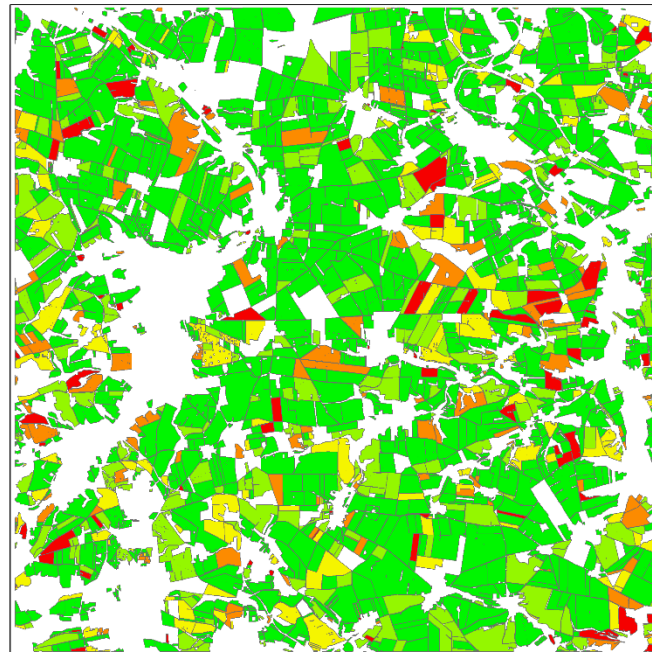
Random Klassifiziert



Mean= 0,899

Klassifikationssicherheit

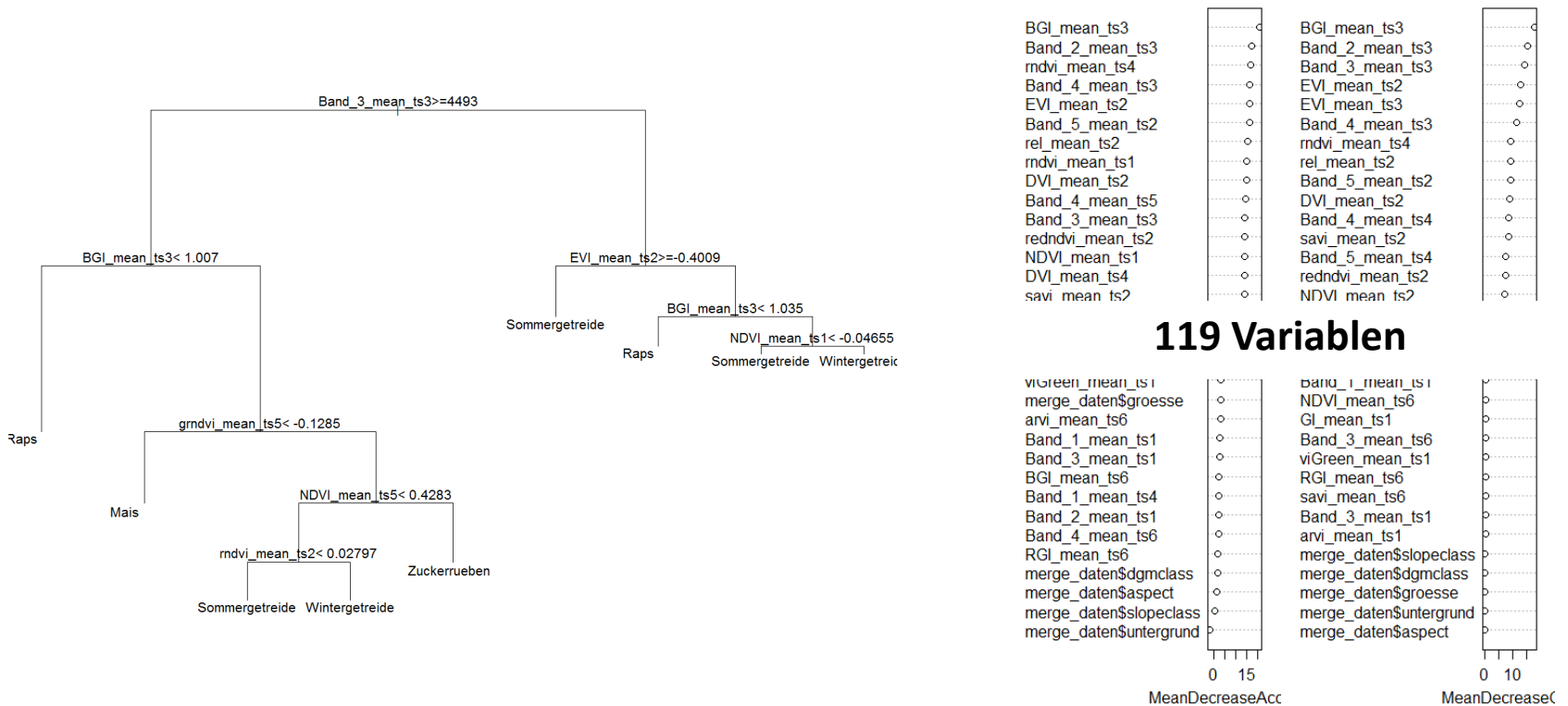
Stratifiziert: Aspect, Dgmclass, Groesse, Untergrund, Slope



Mean= 0,906

Es gibt Unterschiede, Problemfälle bleiben aber meist erhalten

Sekundärdaten werden wie spektrale Parameter für Random Forest verwendet
Entscheidungsbaum und verwendete Variablen der Klassifikation



Im Klassifikator spielen die Umweltvariablen kaum eine Rolle

Eignung des Untersuchungsgebiets:

- Teils sehr große Felder
- Kleinräumige Reliefierung
- Viele verschiedene naturräumliche Merkmale in einem Objekt
- Aussagekraft der Mittelwerte in großen Objekten verringert

Mögliche Daten in anderen Gebieten:

- Klimadaten
- Nähe zu Bewässerung
- Echte Höhenstufen

- Bisher keine deutliche Verbesserung der Ergebnisse durch stratifiziertes Sampling, Stabilität der Random Forest Klassifikation wird allerdings erhöht
- Evtl. Erfolge durch geeignetere Daten oder Übertragen auf ein anderes Untersuchungsgebiet
- Eventuell Modellierung der Klassifikationssicherheit des Random Samplings auf Basis der Umweltvariablen: Gibt es einen Zusammenhang der Umweltvariablen mit der Klassifikationssicherheit?
- Anwendung für andere Fragestellung als Klassifikation: Ertragsmodellierung

- Breiman, L. (2001): “Random Forests”. In: *Machine Learning* 45 (1), S. 5-32.
- Conrad, C., Dech, S., Dubovyk, O., Fritsch, S., Klein, D., Löw, F., Schorcht, G., Zeidler, J. (2014): “Derivation of temporal windows for accurate crop discrimination”. In: *Computers and Electronics in Agriculture*, 103, S. 63–74.
- Löw, F., Duveiller, G., Conrad, C., Michel, U. (2015): “Impact of Categorical and Spatial Scale on Supervised Crop Classification using Remote Sensing”. In: *“Photogrammetrie, Fernerkundung, Geoinformation”*, 1, S. 7-20.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012): „An assessment of the effectiveness of a random forest classifier for land-cover classification“. IN: *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, S. 93-104.
- Turker, M., & Arikan, M. (2005): „Sequential masking classification of multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey“. In: *International Journal of Remote Sensing*, 26(17), S. 3813-3830

Vielen Dank für die Aufmerksamkeit

23.09.2015

Jörn von Uslar, Patrick Knöfel, Christopher Conrad

Universität Würzburg, Institut für Geographie und Geologie, Lehrstuhl für
Fernerkundung, Oswald-Külpe-Weg 86, 97074 Würzburg