

# Sampling Design in Bodenkartierungen

Grundlagen und Beispiele aus Projekten des KOBO

Thorsten Behrens und Karsten Schmidt



## Grundlagen

- Eine Stichprobe wird verwendet, wenn die vollständige Untersuchung einer Grundgesamtheit nicht möglich ist.
- Eine Stichprobe sollte repräsentativ für ihre Grundgesamtheit sein.
- Das Sampling Design\* bestimmt, wie die Stichprobe ausgewählt wird.

\* Stichprobenverfahren, Beprobungskonzept, Messnetzdesign, ...

# Grundlagen

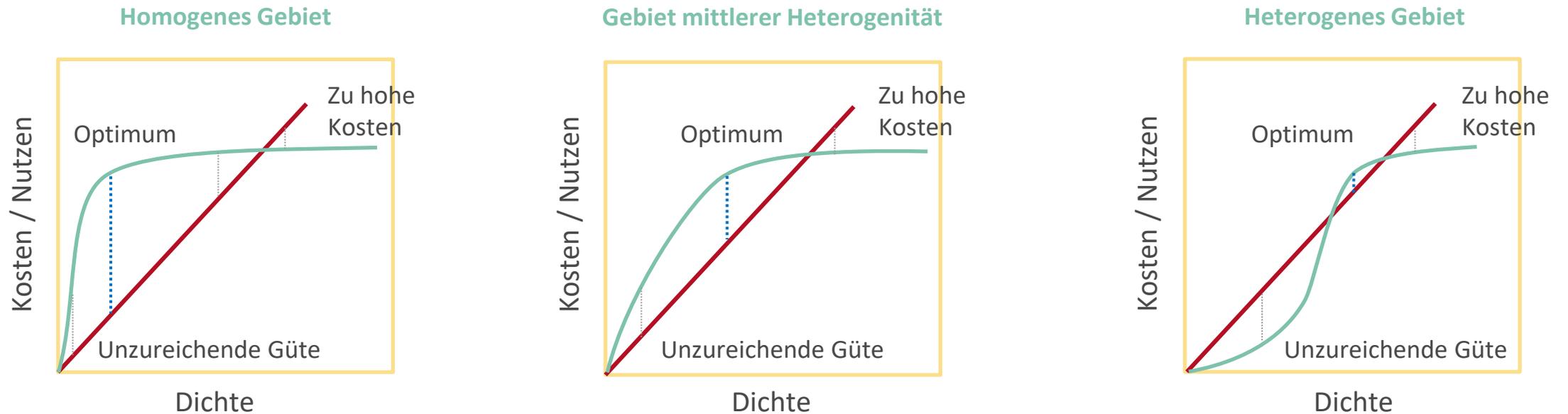
- **Inhalt**
  - Allgemeine Fragen bez. eines Sampling Designs
  - Kosten und Dichte
  - Kleinräumige Variabilität
  - Generalisierungsziel (Massstab) und Support
  - (Gebietsgrösse und Gebietsinventar)
  - Projektziele

# Grundlagen

- **Allgemeine Fragen?**
  - Welche Aussagen sollen mit den Daten möglich sein?
  - Welches Budget steht zur Verfügung?
  - Welches Sampling Design Verfahren soll genutzt werden?
- **Anschlussfragen:**
  - Welche Qualität kann bei dem vorhandenen Budget erwartet werden?
  - Können die gewünschten Aussagen mit dem gewählten Ansatz getroffen werden?
- **Allgemeine Antwort:**
  - Leider wissen wir es vorher oft nicht und hoffen auf eine gute Qualität...

# Grundlagen

- **Kosten und Beprobungsdichte**
  - Abhängig vom Untersuchungsgegenstand/-gebiet



Eine steigende Heterogenität führt zu höheren Kosten  
oder eine steigende Heterogenität führt zu höheren Unsicherheiten

Kosten  
Informationsgehalt

# Grundlagen

- **Kleinräumige Variabilität**

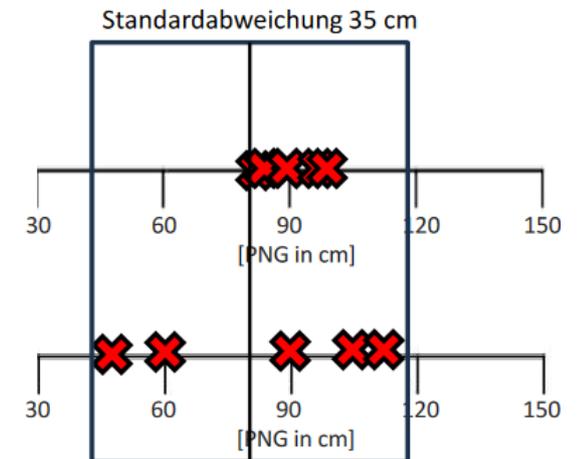
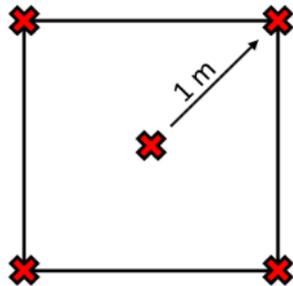
## Repräsentativität und kleinräumige Variabilität - Fokus Lagegenauigkeit

Stefan Oechlin  
BFH-HAFL Bodennutzung und Bodenschutz

07.05.2024  
Kolloquium 2 Boden 4.0

Beim Kartieren ist es sehr wichtig, die **Gesetzmässigkeiten zwischen Ausgangsmaterial, Relief und Bodenform zu erkennen** und **mit Analogieschlüssen zu arbeiten** (Abb. 7.3f). **So sollten die Kartierenden versuchen**, in ähnlichen Lagen auf gleichem Ausgangsmaterial die gleichen Böden zu kartieren und **sich nicht durch lokal geringe Bodenunterschiede** oder durch Bohrungen in Übergangsbereichen **verunsichern zu lassen**.

Schriftenreihe der FAL (24), 1997



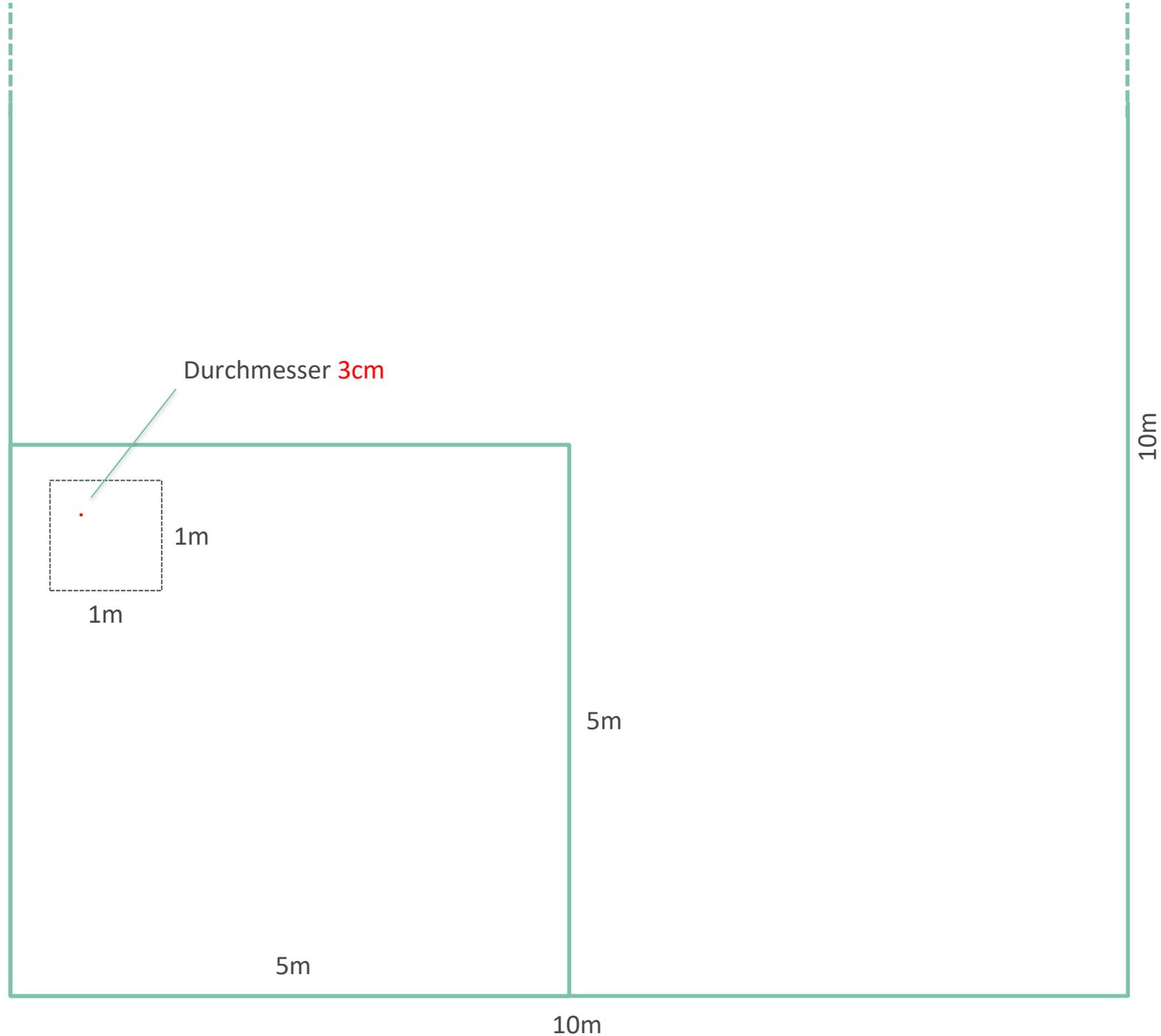
# Grundlagen

- **Generalisierung, Support und Probennahme**

- **Flächenverhältnis:**

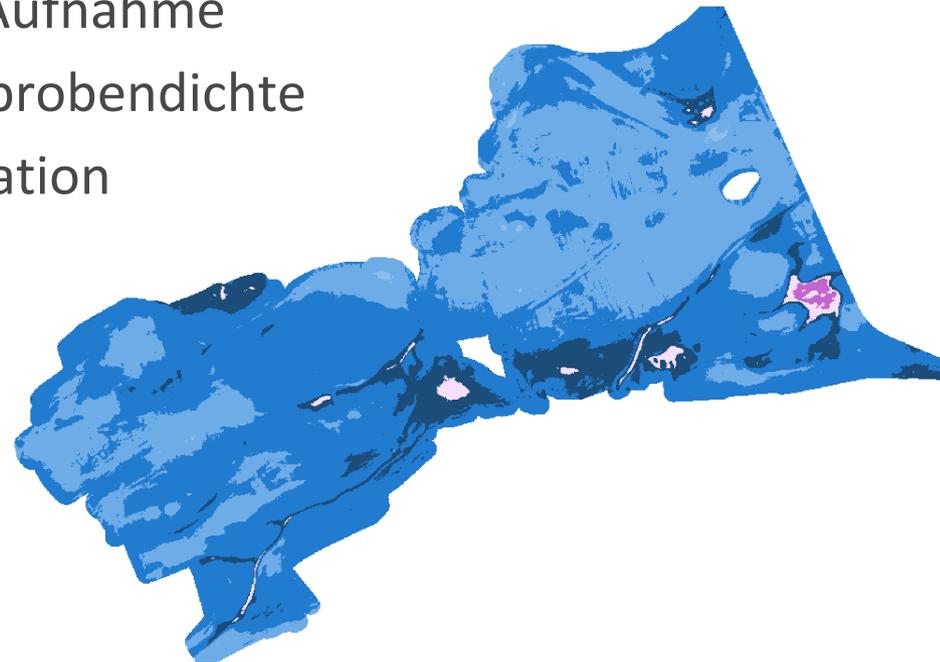
## Bohrstock vs. „Massstab“

- 1:5.000      1:10.000
- 5x5m      10x10m
- 1/35.000      1/140.000



# Grundlagen

- **Projektziele**
  - Was erwartet ein Kanton bei einer Vergabe eines Kartierloses?
    - Eine gute Bodenkarte
  - Was wollen wir in den KOBO-Projekten erreichen?
    - Feldtools für die Pedologische Aufnahme
    - Analyse der notwendigen Stichprobendichte
    - Einführung der rKLABS Klassifikation
    - *Validierung?*



# Stichprobenverfahren

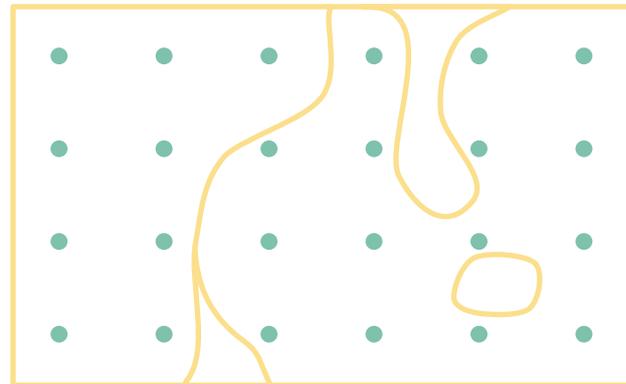
Wikipedia

# Stichprobenverfahren

- **Inhalt**
  - Räumliche und inhaltliche Verfahren
  - Beispiele
  - Algorithmen
  - Probleme
  - (Code)

# Stichprobenverfahren

- **Räumliche Verfahren**
  - Ohne Vorinformationen
    - bzw. wenn keine relevanten Vorinformationen vorliegen



Raster / systematisch



Zufall

# Stichprobenverfahren

- **Inhaltliche Verfahren**

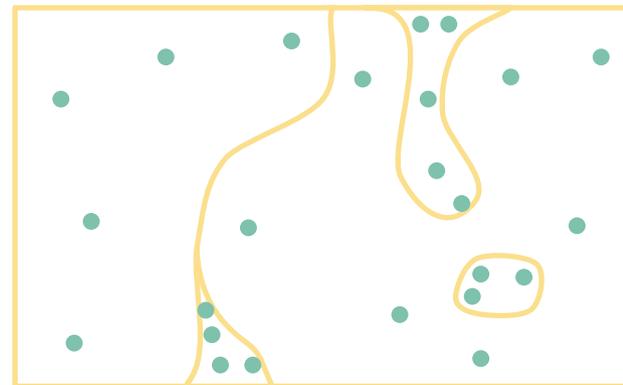
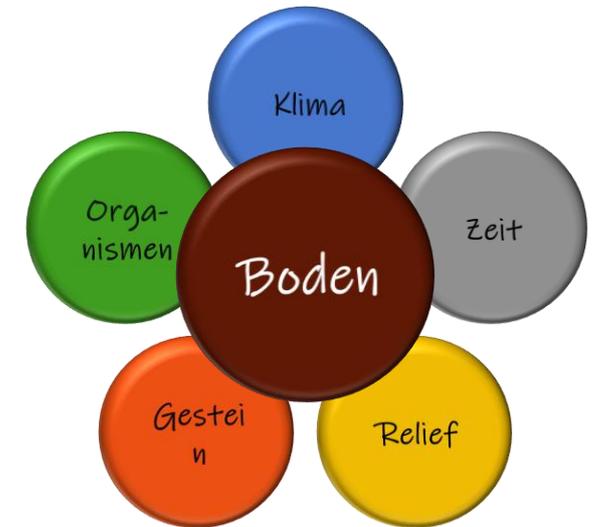
- Mit Vorinformationen / systematisch

- Bodenfunktionsgleichung nach Jenny (1941)

- Boden ist eine Funktion aus:

- Klima, Organismen, Relief, Ausgangsgestein und Zeit

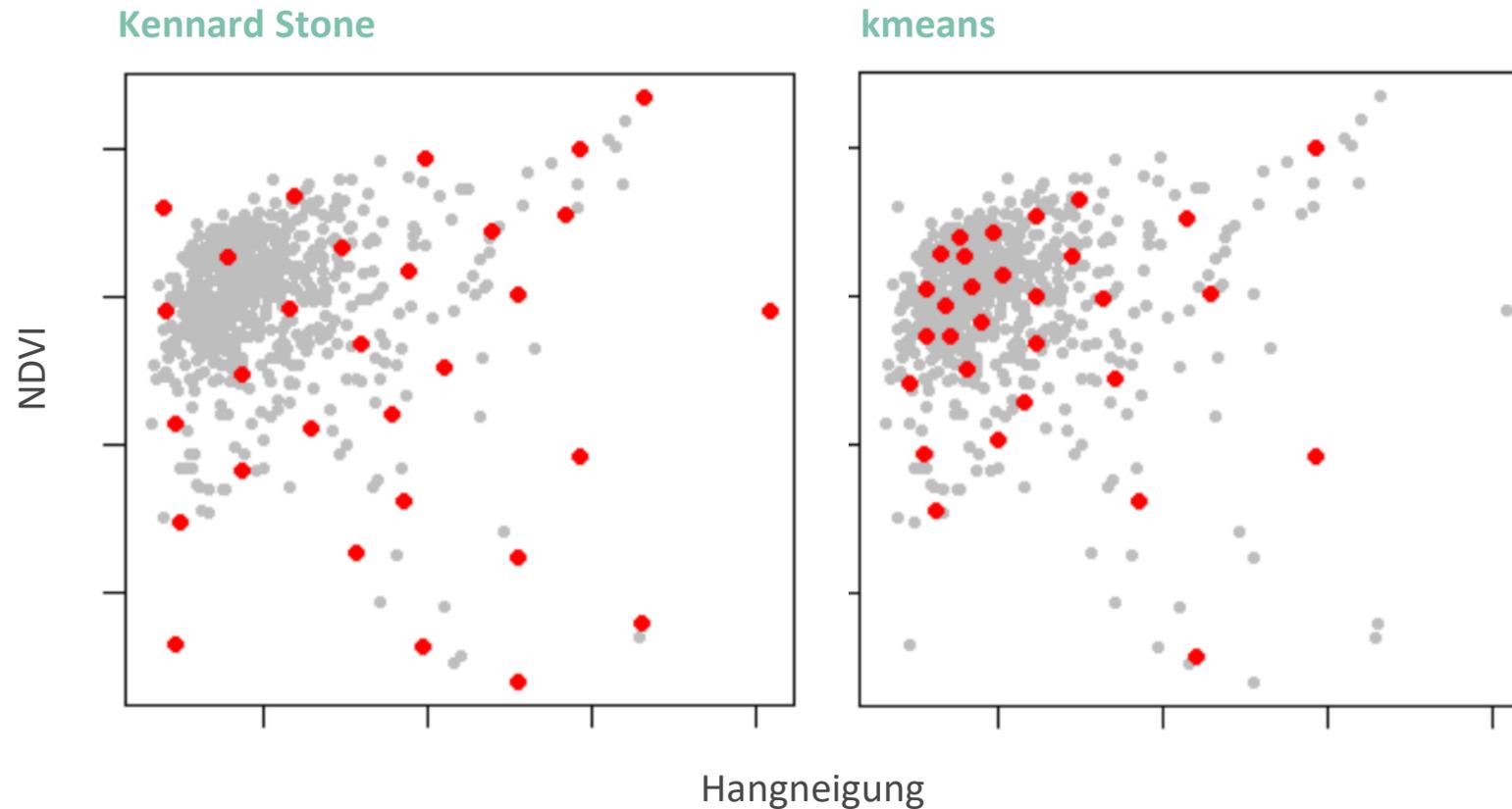
- $s = f(c_l, o, r, p, t)$



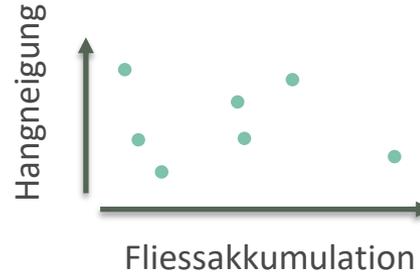
Geschichtet/stratifiziert

# Stichprobenverfahren

- **Inhaltliche Verfahren**
  - Kmeans, Kennard Stone, Latin hypercube, ...



# Stichprobenverfahren



- Algorithmen

## Kennard Stone

Schritt 1

Punkt, der im Zentrum liegt (oder der am weitesten von allen anderen entfernt liegt)



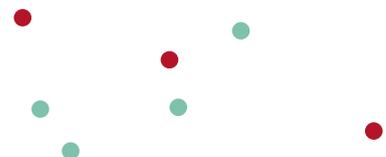
Schritt 2

Punkt, der am weitesten vom ersten Punkt entfernt liegt



Schritt 3

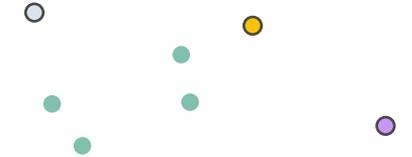
Punkt, der am weitesten von den ersten beiden Punkten entfernt liegt



## kmeans

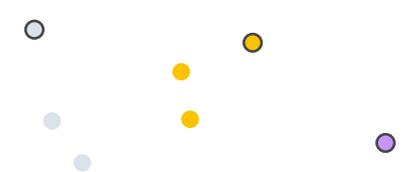
Schritt 1

Auswahl n zufälliger Punkte (hier n = 3)



Schritt 2

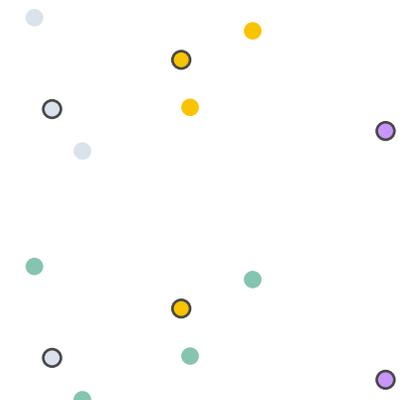
Alle anderen Punkte werden dem zufällig gezogenen Punkt zugewiesen, der am nächsten liegt



Schritt 3ff

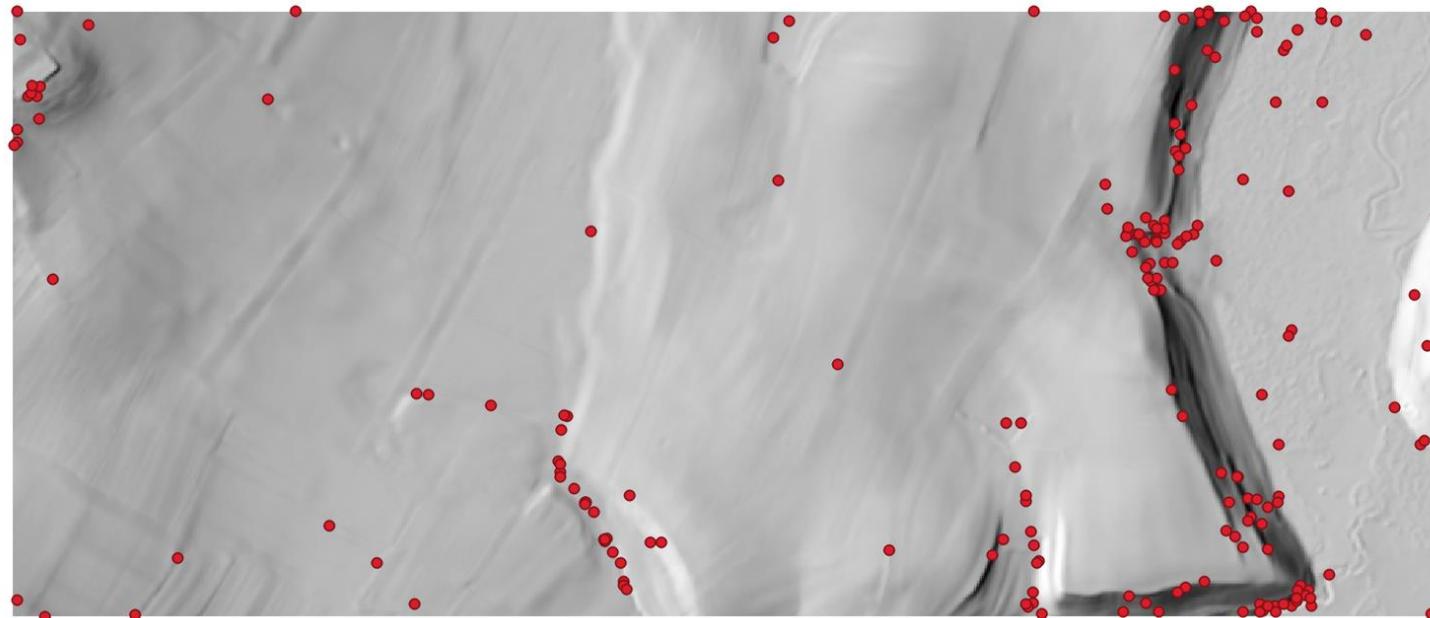
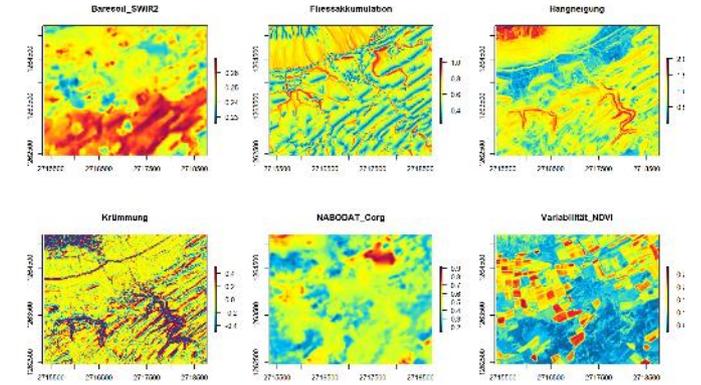
Optimierung der Cluster:

- Neues Zentrum mit kürzestem Abstand zu allen Punkten aus der selben Klasse
- Neue Zuweisung aller Punkte zum neuen Zentrum



# Stichprobenverfahren

- Probleme inhaltlicher Ansätze
  - Abdeckung
    - Nur inhaltlich?



# Sampling Design am KOBO - Analyse und Kombination von Methoden

# Sampling Design am KOBO

- **Inhalt**
  - Hintergrund
  - Eingesetzte Methoden
    - Regionalisierung
    - Kalibration
    - Bodeneigenschaftskarten
    - (Pedologische Bohrungen)
  - Ausweichflächen

# Sampling Design in KOBO-OPs

- **Hintergrund**

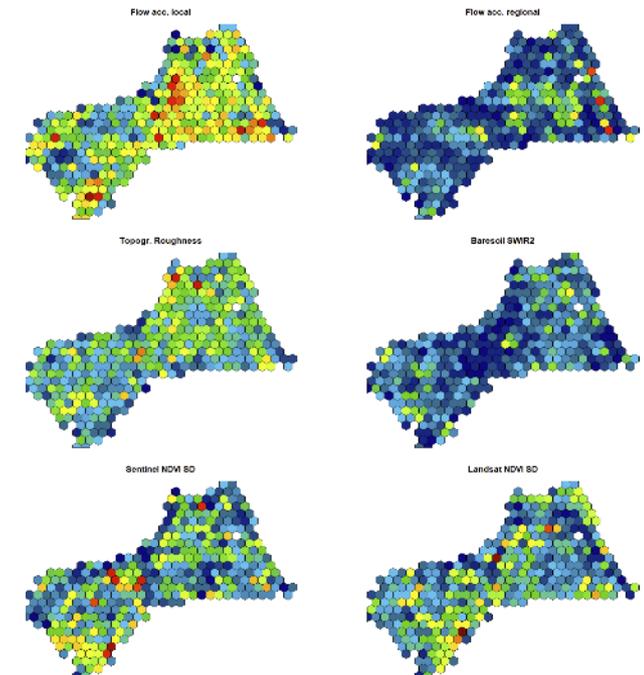
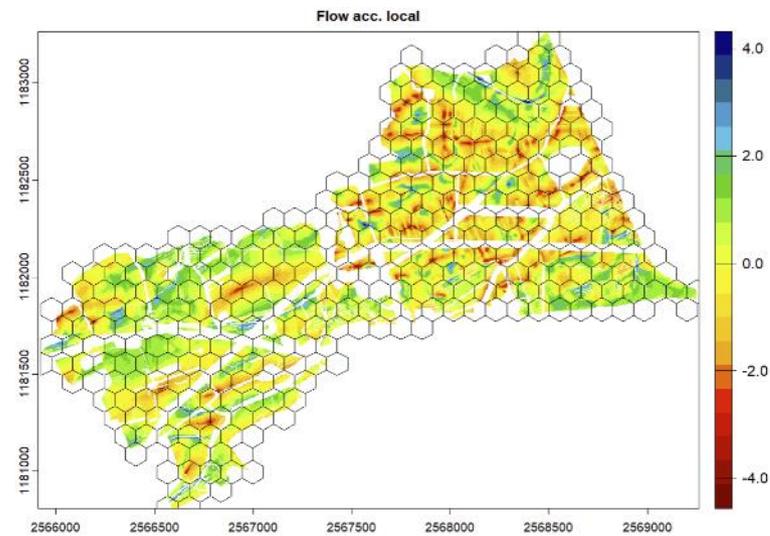
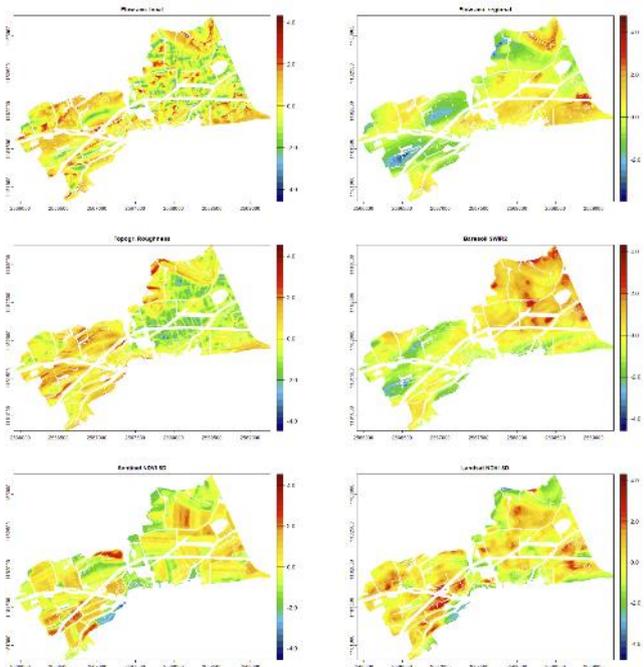
- H3: Bodeneigenschaftskarten
  - Vor der pedologischen Aufnahme
    - (H1: Profile, H2: Bohrungen)
  - 3-5 Tiefenstufen
  - Ton, Schluff, Sand, pH, C<sub>org</sub>, Carbonat
  - Ermöglicht viel besseres Boden-Landschaftsverständnis
  - Grundlage für Polygonabgrenzung
  - Grundlage für Sampling Design für H2
  - Grundlage für die meisten Themenkarten
- H2: Bohrungen
  - Vorgegebene Bohrungen
  - Freie Bohrungen
- Validierung der Modelle

- **Umsetzung**

- Hybrides Sampling Design
  - Erfassung der **lokalen inhaltlichen** Variabilität
  - Sicherstellung einer **räumlich gleichmässigen** Verteilung
- Option die relevante Dichte systematisch analysieren zu können

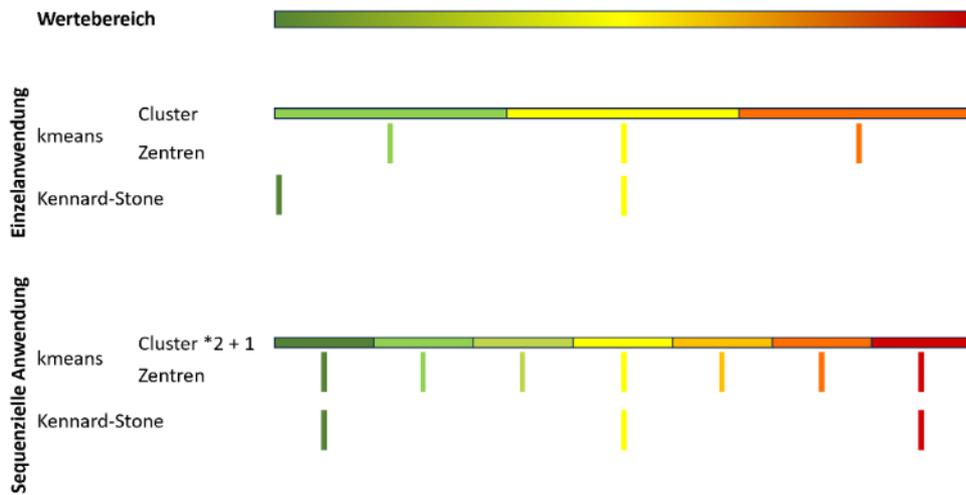
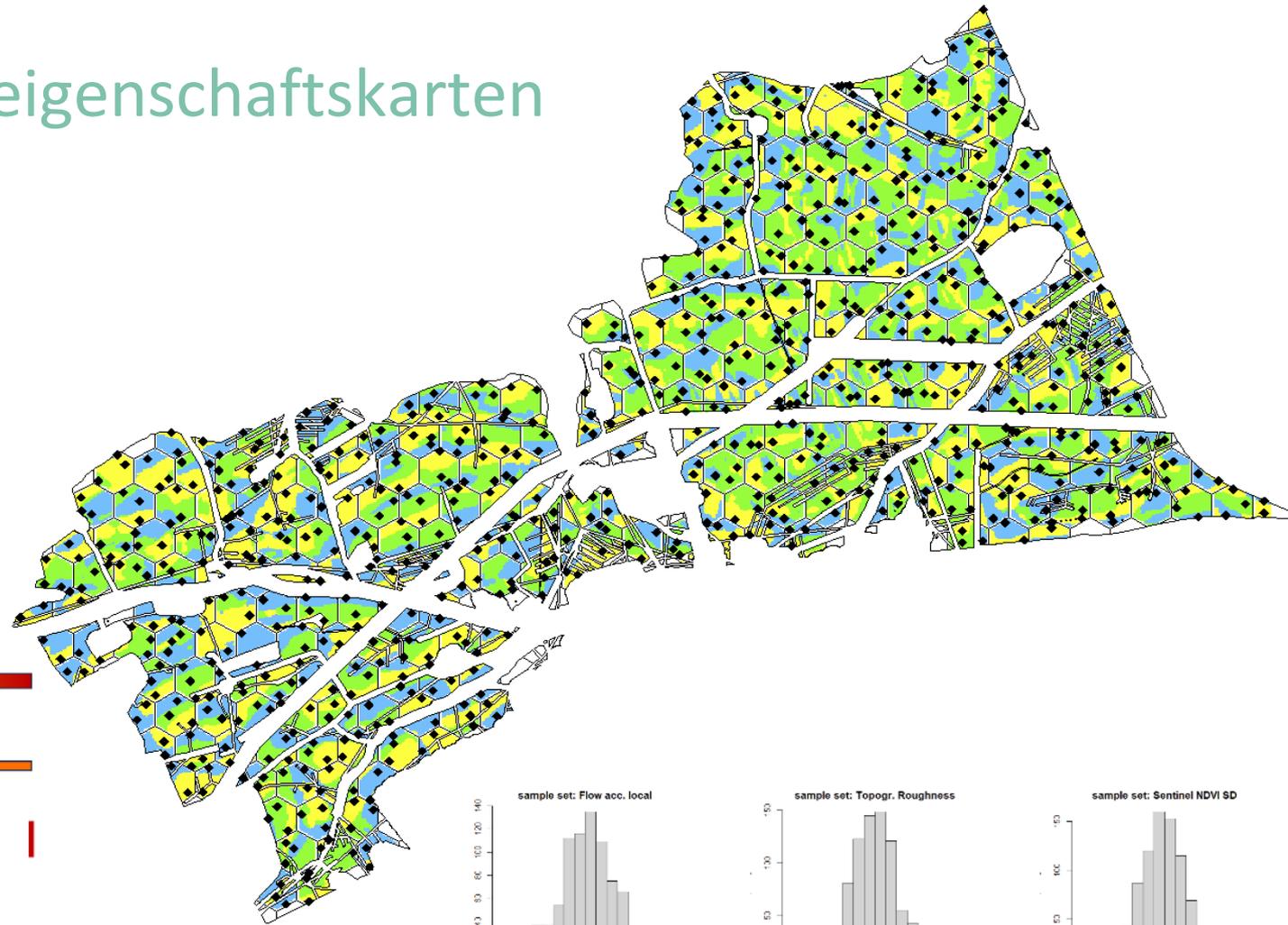
# Sampling Design | H3 Bodeneigenschaftskarten

- Sampling am KOBO: OP3-OP8
  - 3 Standorte / ha
  - Auswahl von *cloprt* (Jenny) Faktoren auf Basis der lokalen Variabilität
    - Bezugseinheit: 1ha

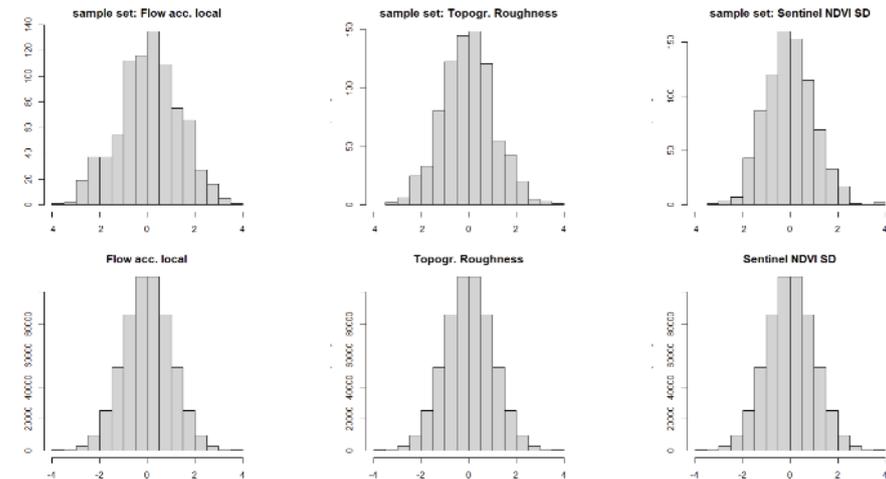


# Sampling Design | H3 Bodeneigenschaftskarten

- Sampling am KOBO: OP3-OP8
  - Inhaltliches Sampling
    - Innerhalb der 1ha Hexagone
    - Kmeans + KenStone
      - \_ sequenziell

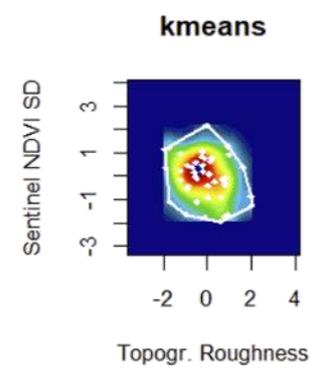
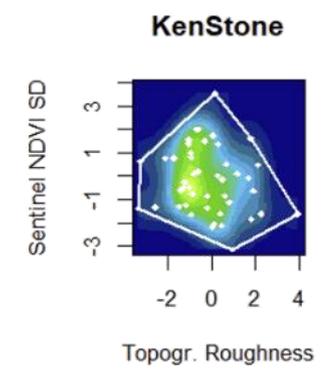
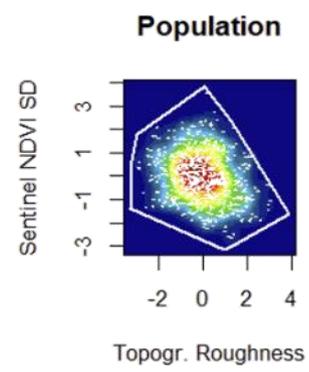
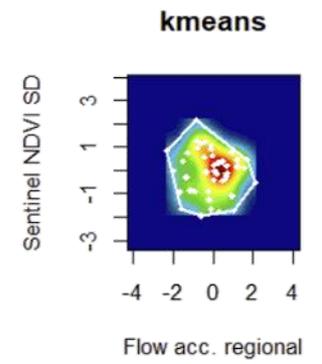
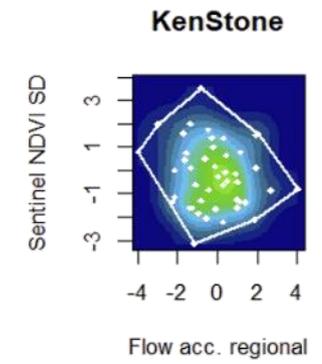
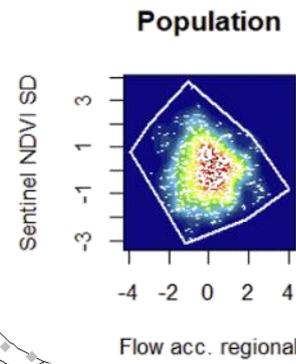
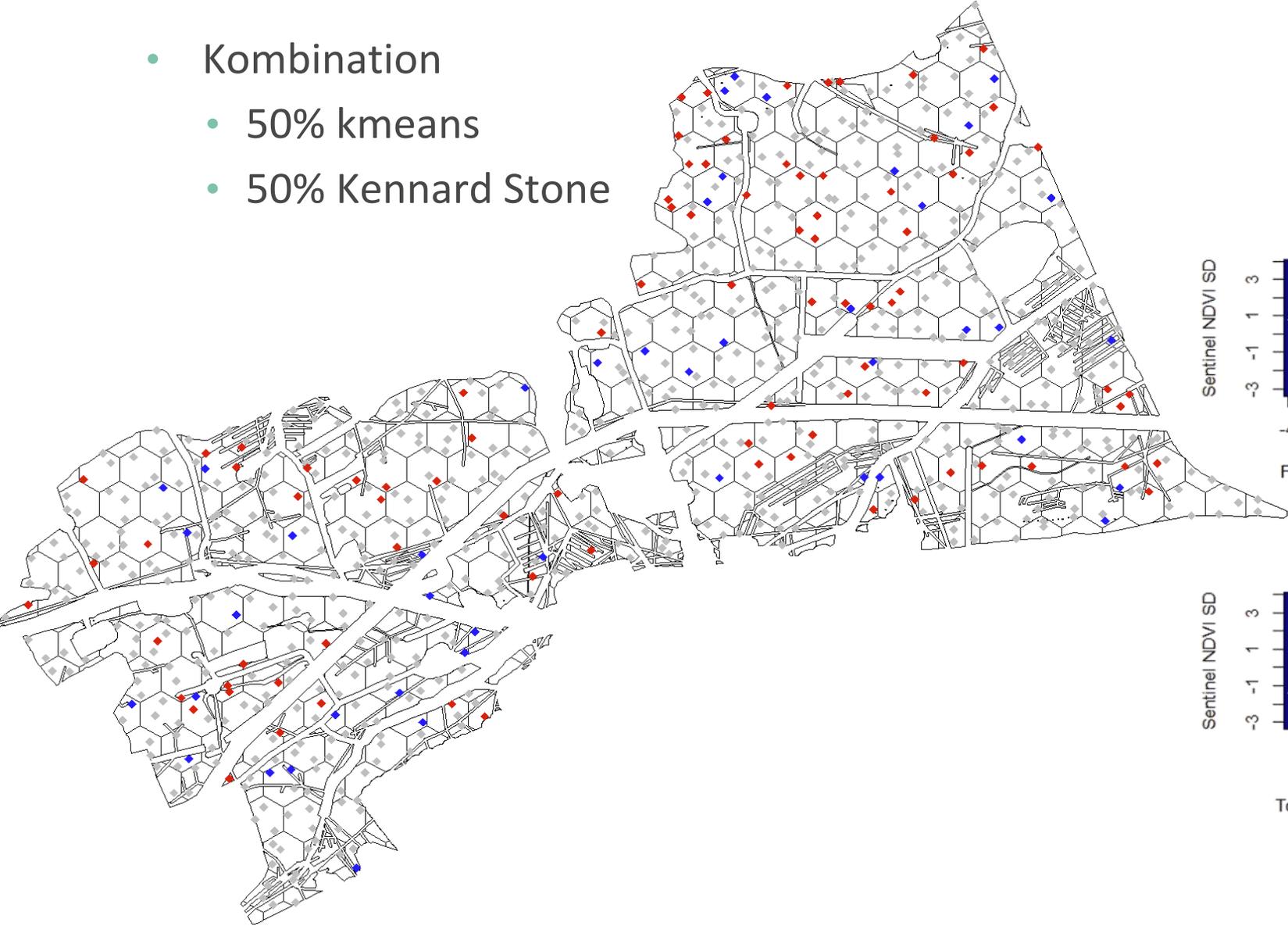


Puffern kritischer Extreme  
Stabile Extreme  
und der Durchschnitt



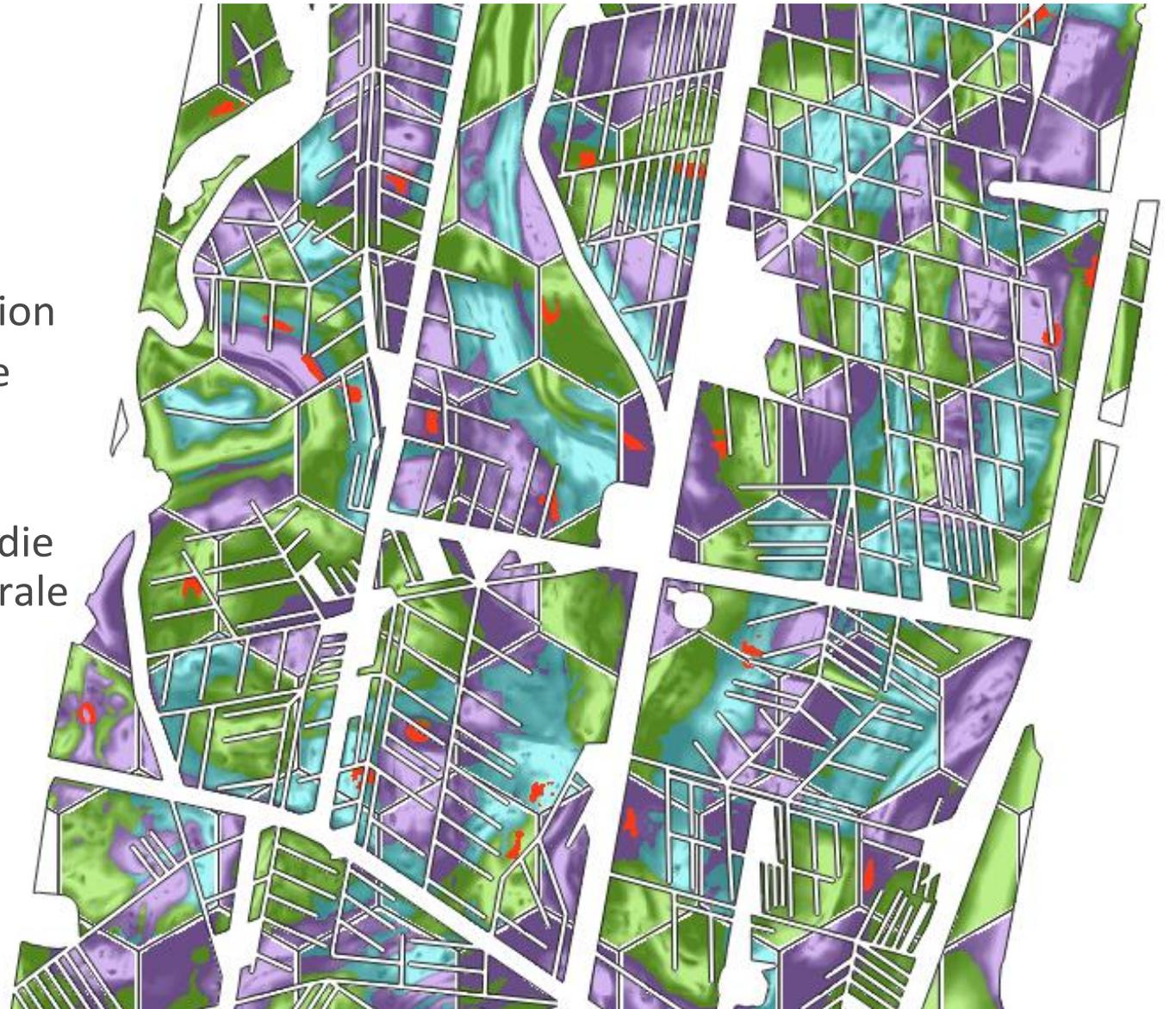
# Sampling Design | H3 Kalibrationsstandorte

- Kombination
  - 50% kmeans
  - 50% Kennard Stone



# Sampling Design | H3 / H2

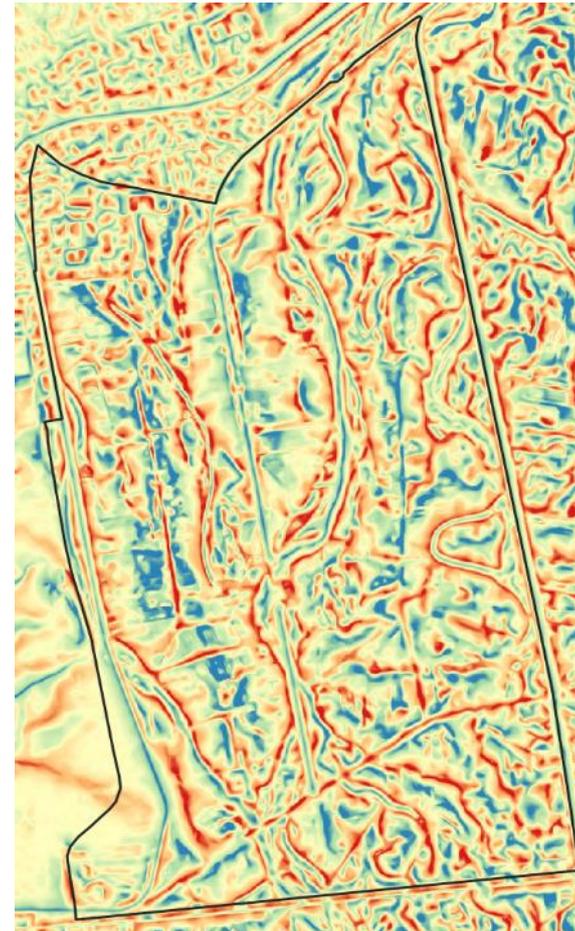
- Sampling am KOBO: OP3-OP8
  - Ausweichflächen
    - Mit Repräsentanzinformation
      - Ermöglicht eine pedologische Interpretation der Geländesituation
      - Ermöglicht eine optimierte Suche nach Ausweichstandorten
    - Der ausgewiesene Punkt dient für die H3 Techniker (Beprobung) als zentrale Vorgabe
    - Für die Pedolog:innen dient der ausgewiesene Punkt nur der Orientierung innerhalb des Repräsentanzraums



# Sampling Design am KOBO ...Aufbereitung der Basisdaten

# Sampling Design | Basisdaten

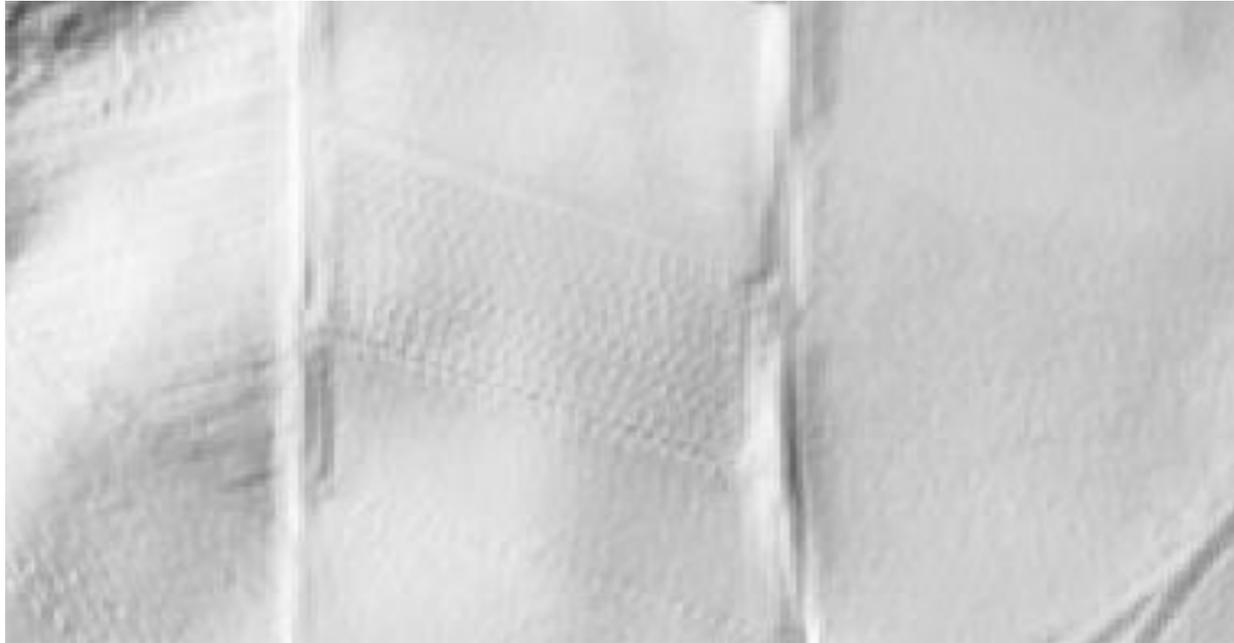
- Auswahl relevanter Basisdaten ja nach Gebiet und Pedogenese



# Sampling Design | Basisdaten

- Aufbereitung der Grundlagendaten

SwissTopo 2m



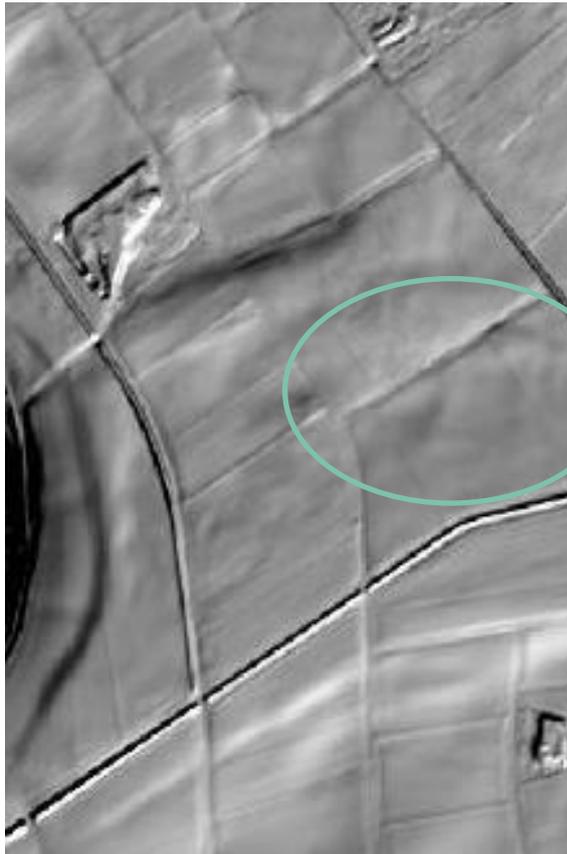
SwissTopo 0.5m -> KOBO 2m



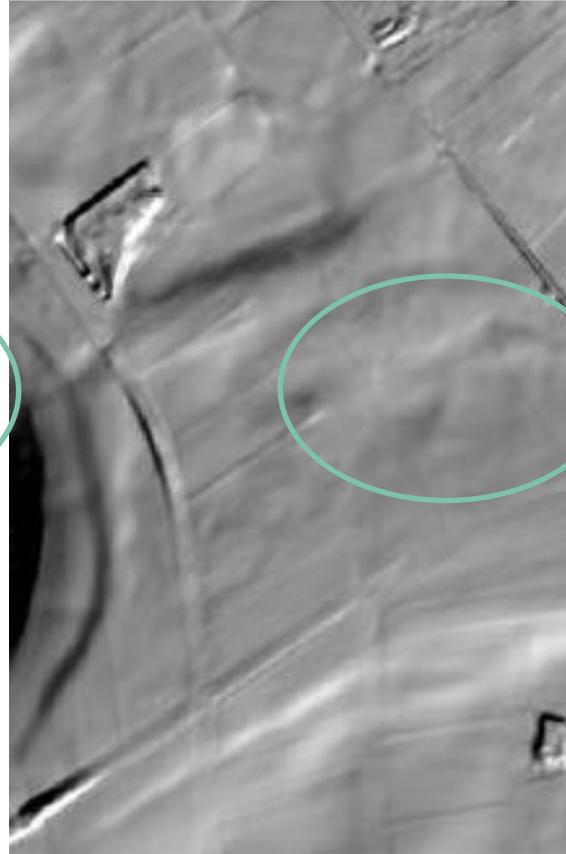
# Sampling Design | Basisdaten

- Aufbereitung der Grundlagendaten

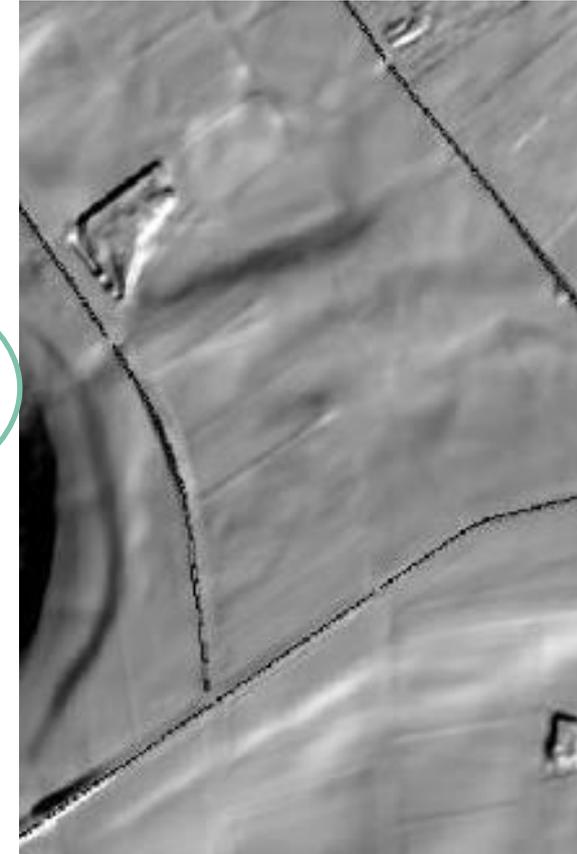
Rohdaten



Wege entfernt

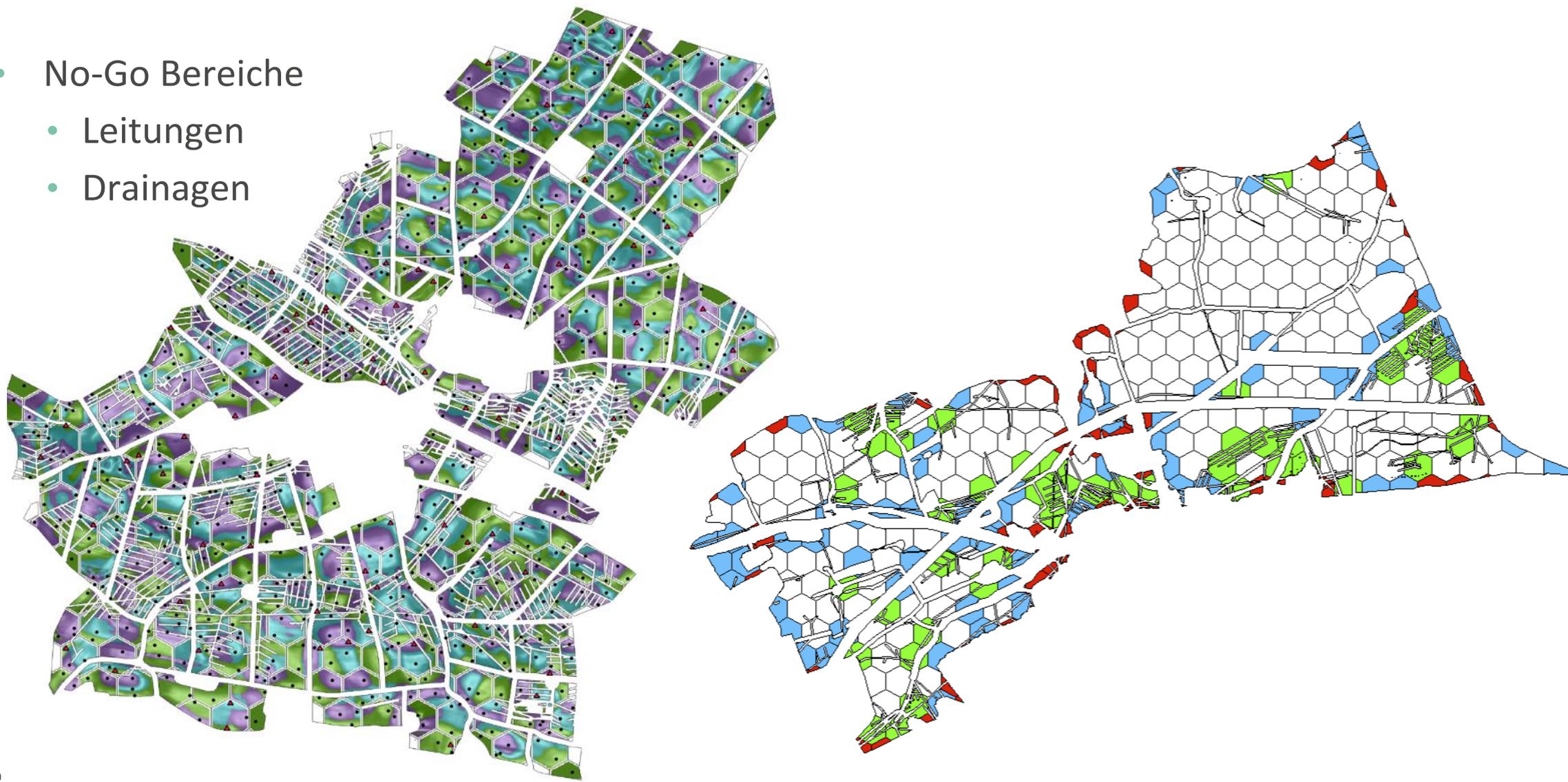


Gewässer eingebraunt



# Sampling Design | Ausschlussflächen

- No-Go Bereiche
  - Leitungen
  - Drainagen



# Modellierung

## Analyse der notwendigen Dichte

# Sampling Design am KOBO II

- **Inhalt**
  - Prozesskette
    - Daten- und Probenmanagement
    - Übersicht
      - \_ Spektroskopie
      - \_ Räumliche Modellierung
  - Analyse der notwendigen Beprobungsdichte
  - Grenzen der Modellierung

# Maschinenraum

- Daten- und Probenmanagement
- Digitalisierung und Skalierung



Extern

## Ihr Spezialist in der Datenverarbeitung

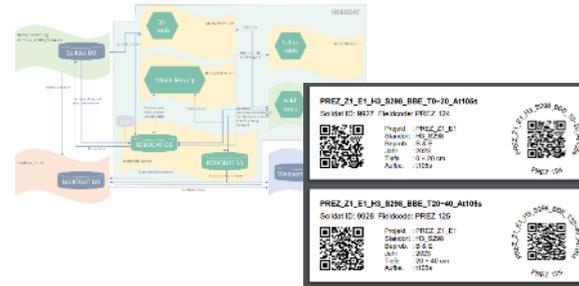
Ihre Daten sind von zentraler Bedeutung für Ihr Unternehmen? Wir können Ihnen helfen! Wir sind für Sie da, um Ihre Geodaten aufzuwerten, um maßgeschneiderte Anwendungen zu entwickeln, um eine Cloud-Infrastruktur zu konzipieren oder Sie bei der Migration Ihres bestehenden Systems zu beraten.

Extern

## Object Storage

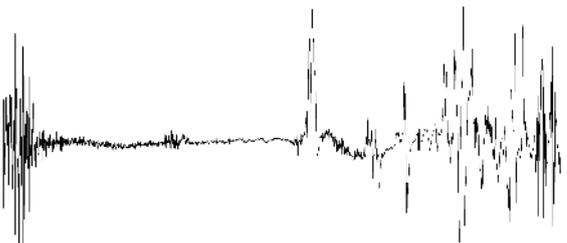
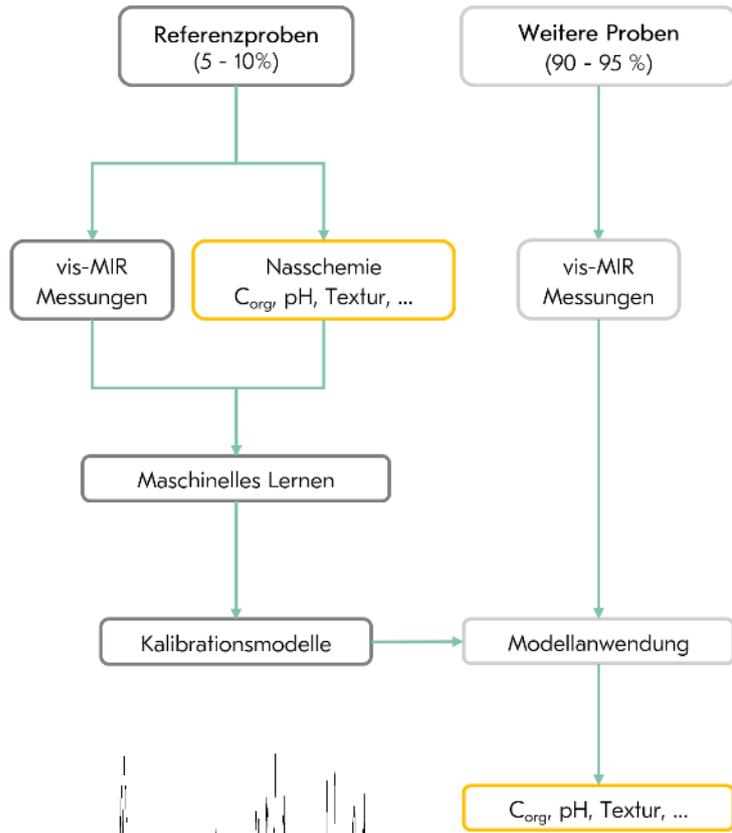
An S3 Compatible Object Storage to store your assets, files, and metadata at a competitive price

A scalable, reliable, and cost effective Object Storage solution to support your application. Backup or store your data from any Exscale zone with no hidden fees, using your existing S3 compatible tooling and a familiar API.

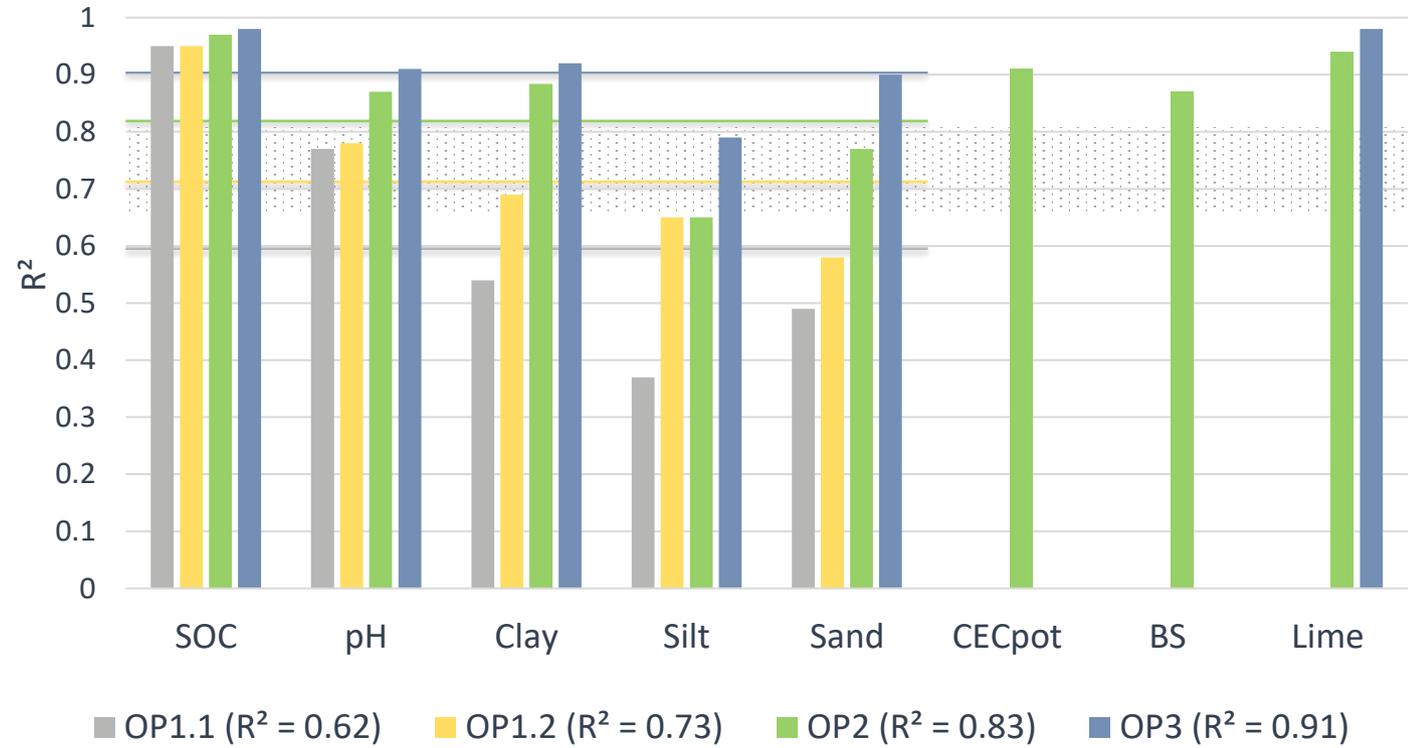



# Modellierung

- Spektroskopie

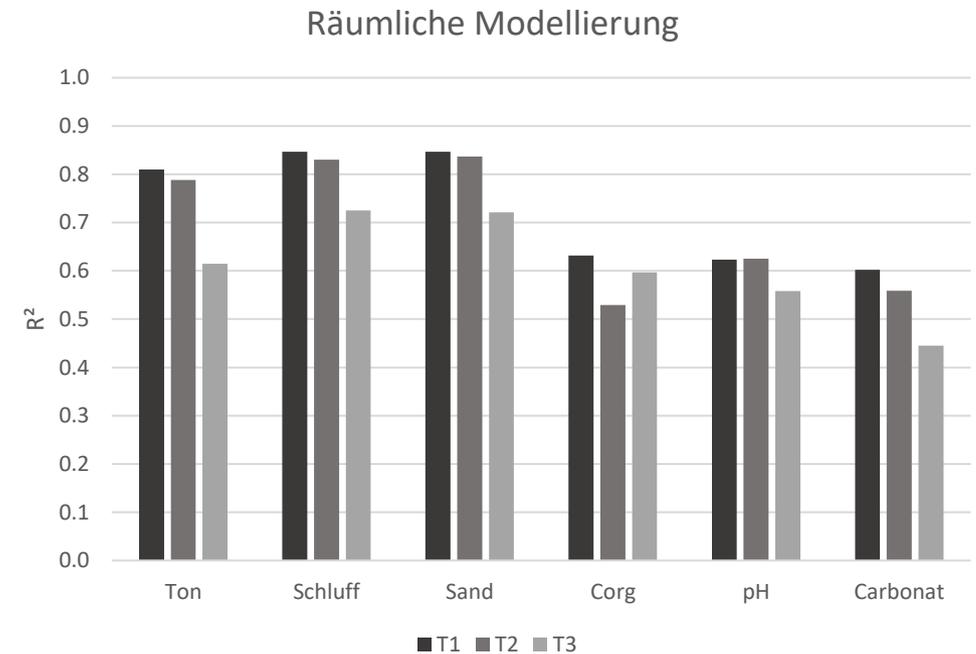
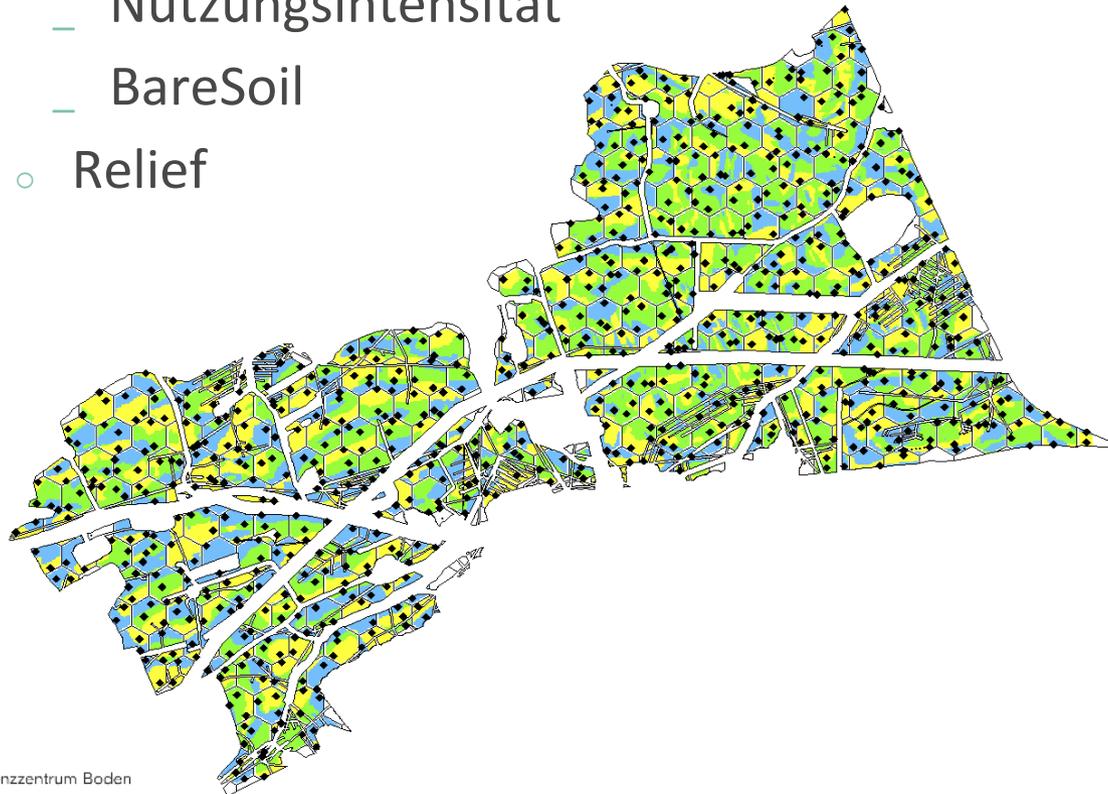


Modellgüte (5x10CV), Spektroskopie



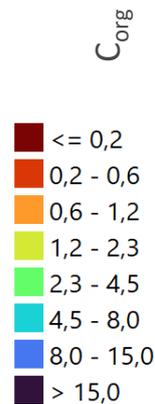
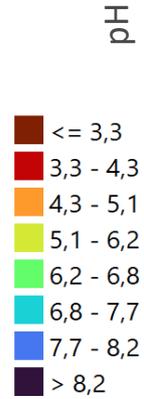
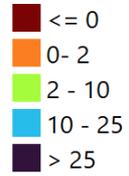
# Modellierung

- Räumliche Modellierung
  - Multi-skalige, multi-spektrale und multi-temporale Daten
    - Fernerkundung
      - \_ Nutzungsintensität
      - \_ BareSoil
    - Relief



# Modellierung

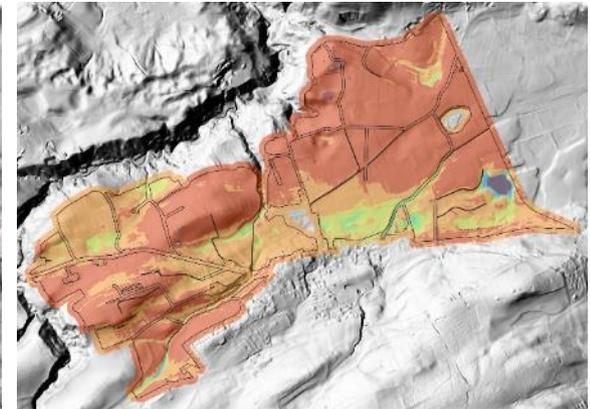
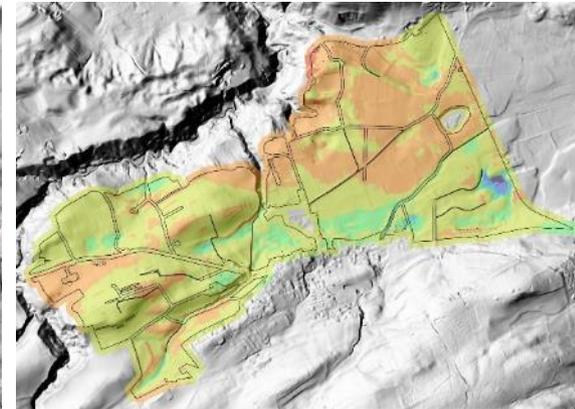
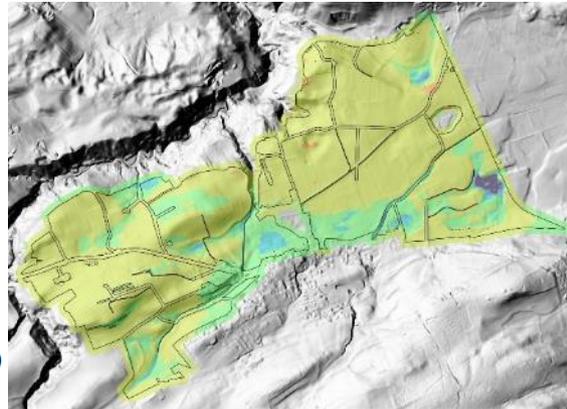
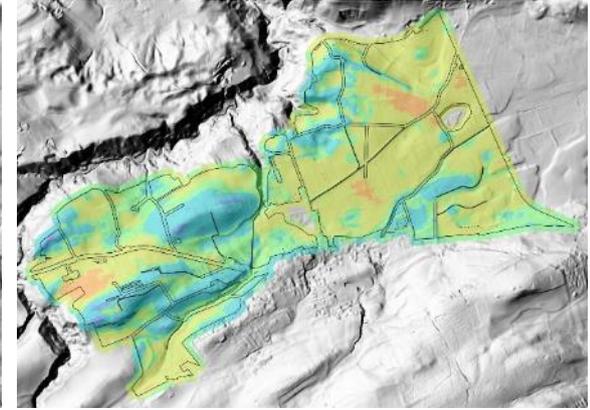
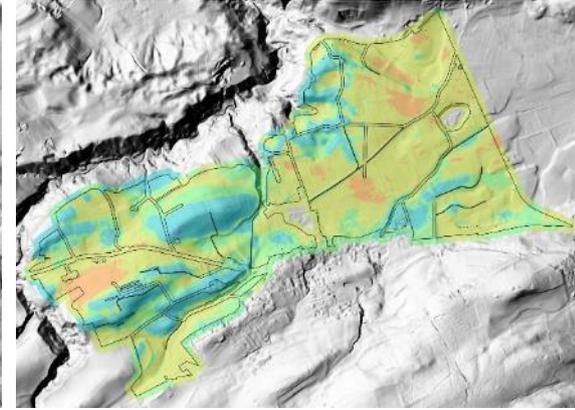
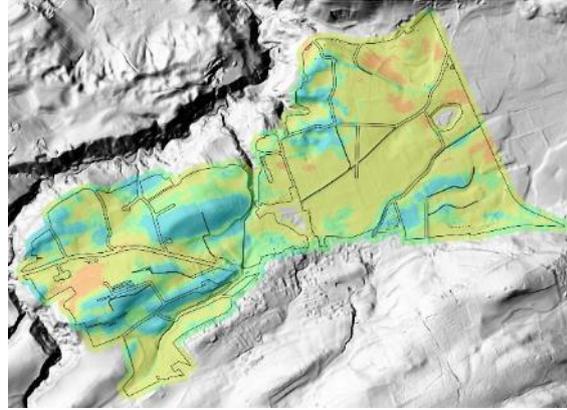
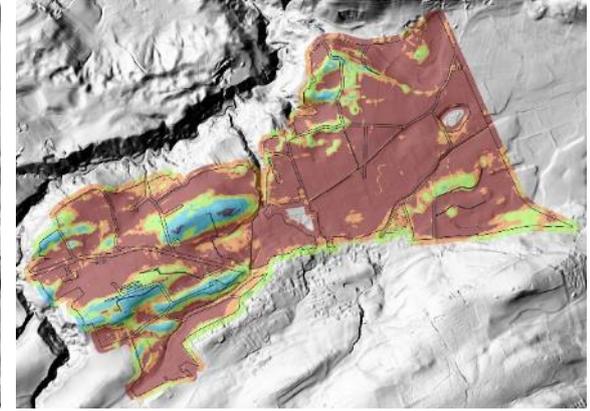
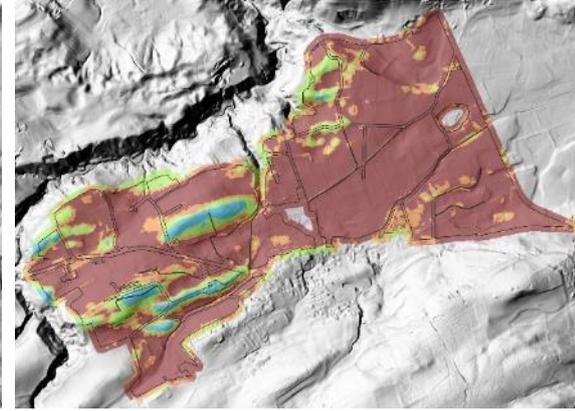
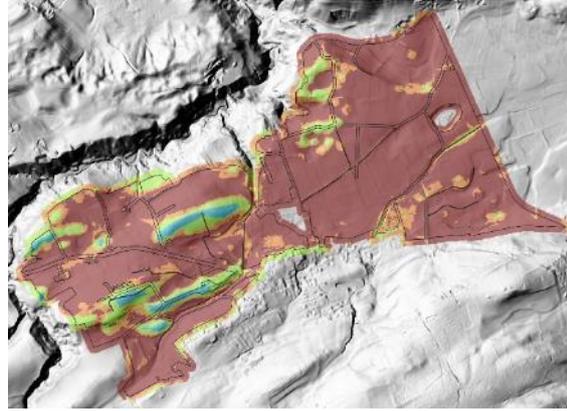
- Bodeneigen-  
schaftskarten



0-20

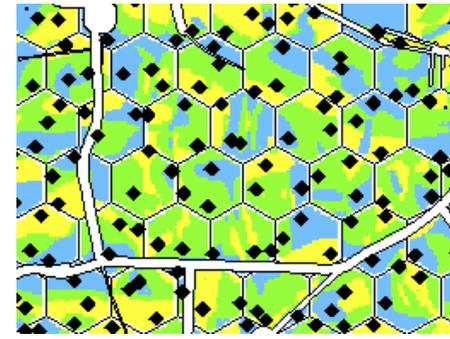
20-40

40-70

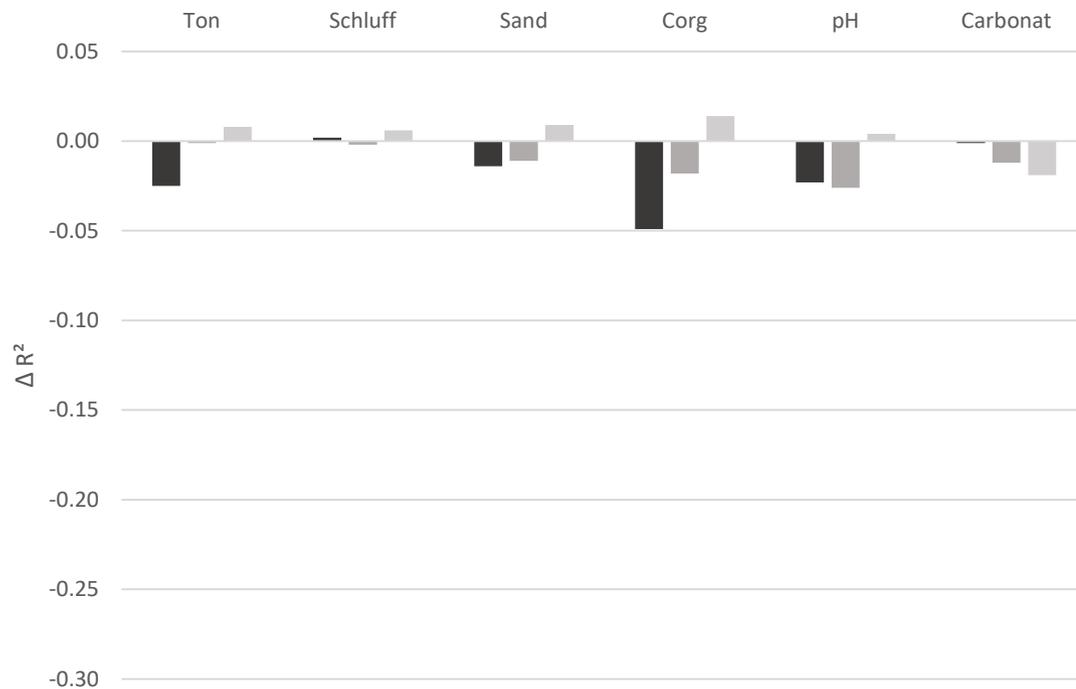


# Analyse

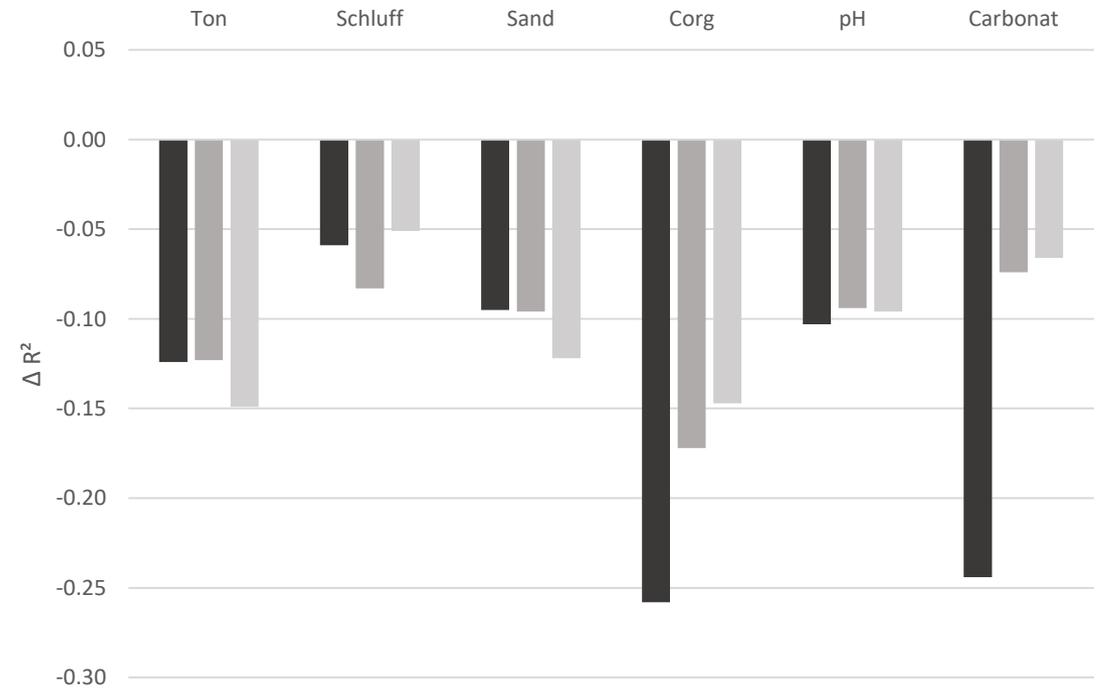
- Reduktion von 3 zu 2 und von 3 zu 1 Probe / ha



Modellgüte vs. Probendichte  
3 vs. 2

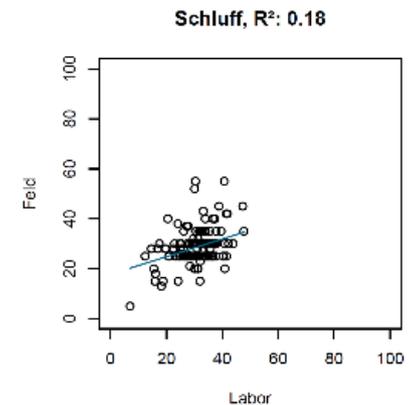
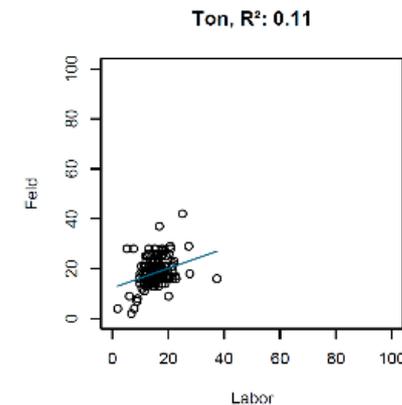
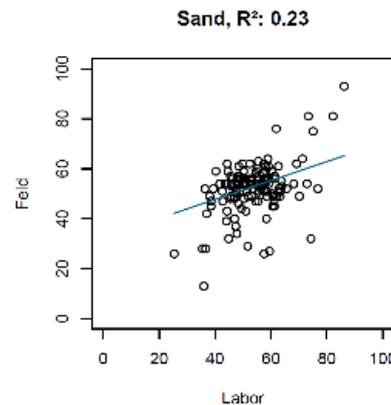
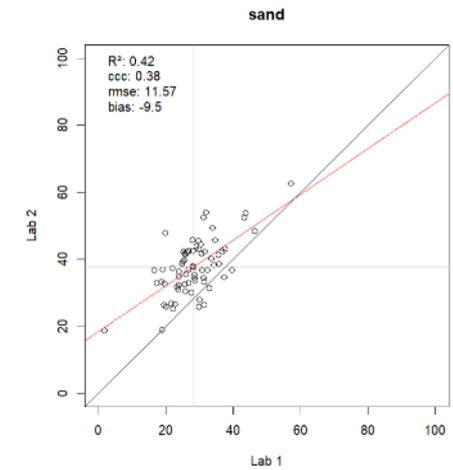
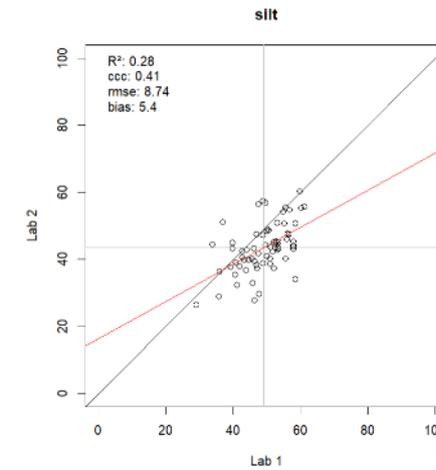
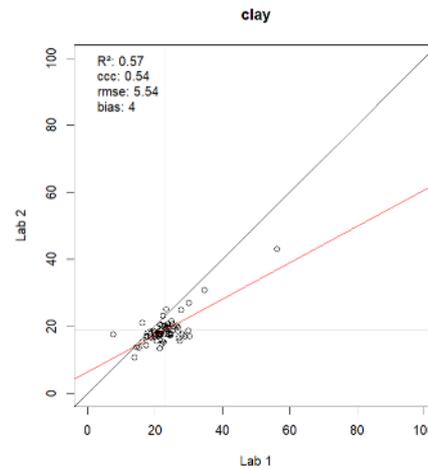
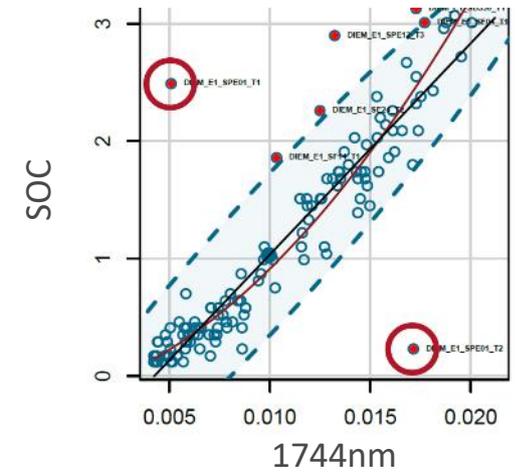
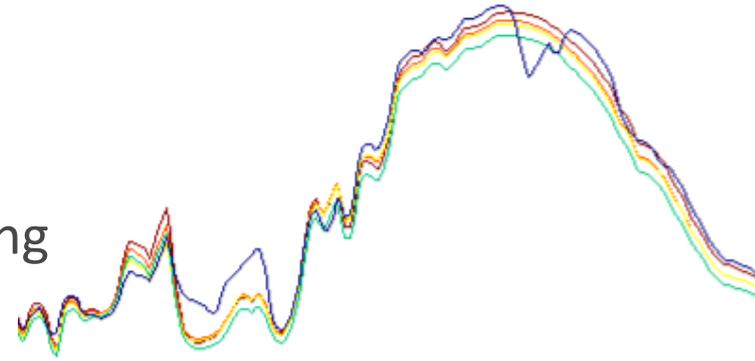


Modellgüte vs. Probendichte  
3 vs. 1



# Garbage in, garbage out...

- Grenzen der räumlichen Modellierung



# Validierung

Wie unabhängig muss sie sein?

# Validierung

- **Inhalt**
  - Methoden der internen Validierung
  - Unabhängige Validierung
    - Beispiele
    - Probleme

# Kreuzvalidierung

- Validierung an den Trainingsdaten



$R^2: 0.98$

---

- Validierung an einer Teilmenge, die nicht in die Modellbildung eingeht



$R^2: 0.66$

---

- Kreuzvalidierung



$R^2: 0.57$



$R^2: 0.83$



$R^2: 0.72$

$R^2: (0.57 + 0.83 + 0.72) / 3$

$R^2: 0.71$

# Unabhängige Validierung

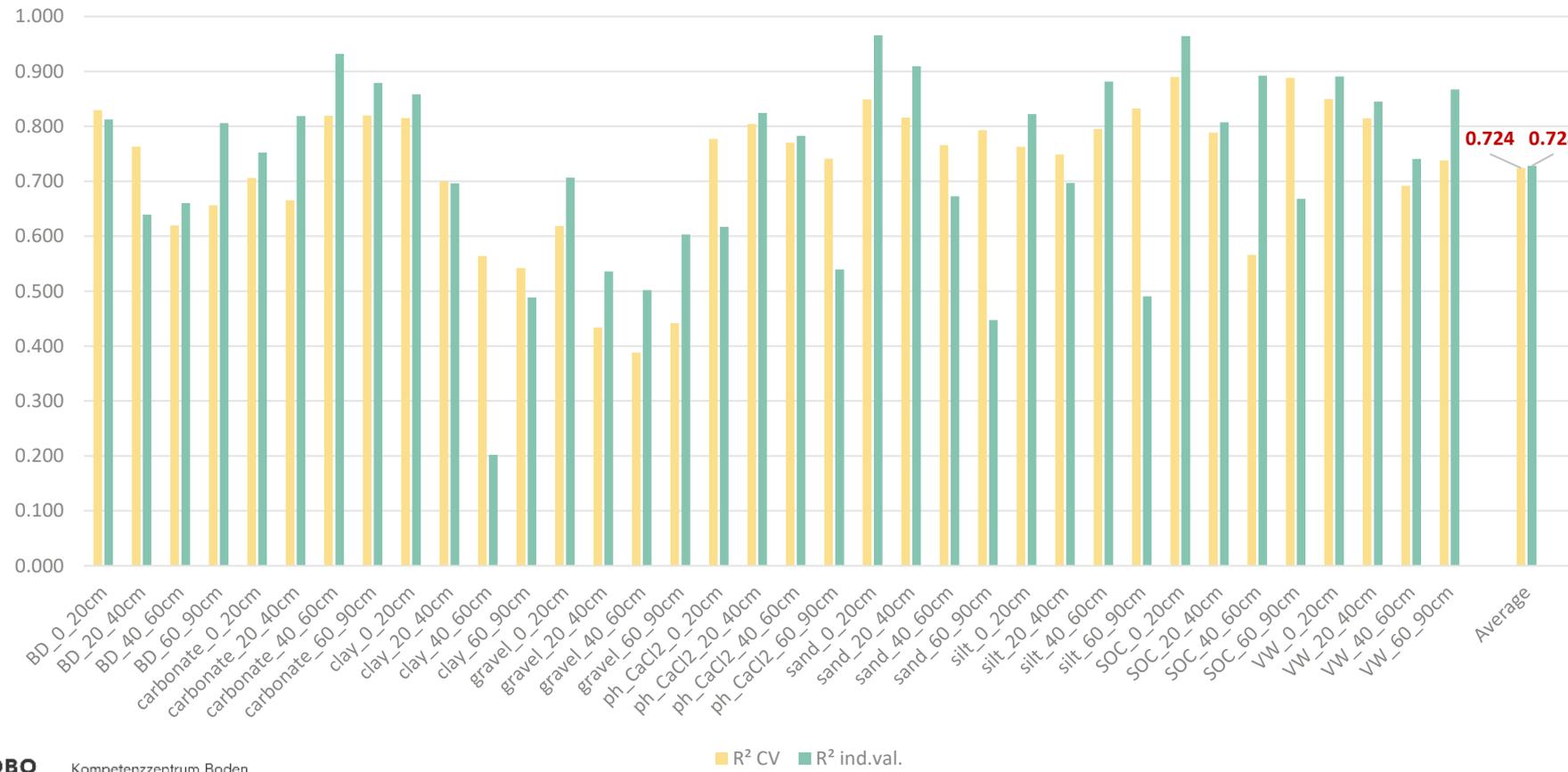
- Räumlich geschichtete Zufallsstichprobe



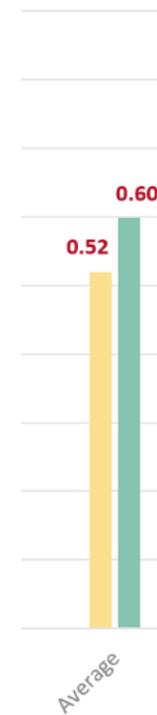
# Unabhängige Validierung

- 10x 10fold Kreuzvalidierung vs. räumlich geschichtete Zufallsstichprobe

LU | Rickenbach | unabhängige Validierung



GR | Mathon-Lohn

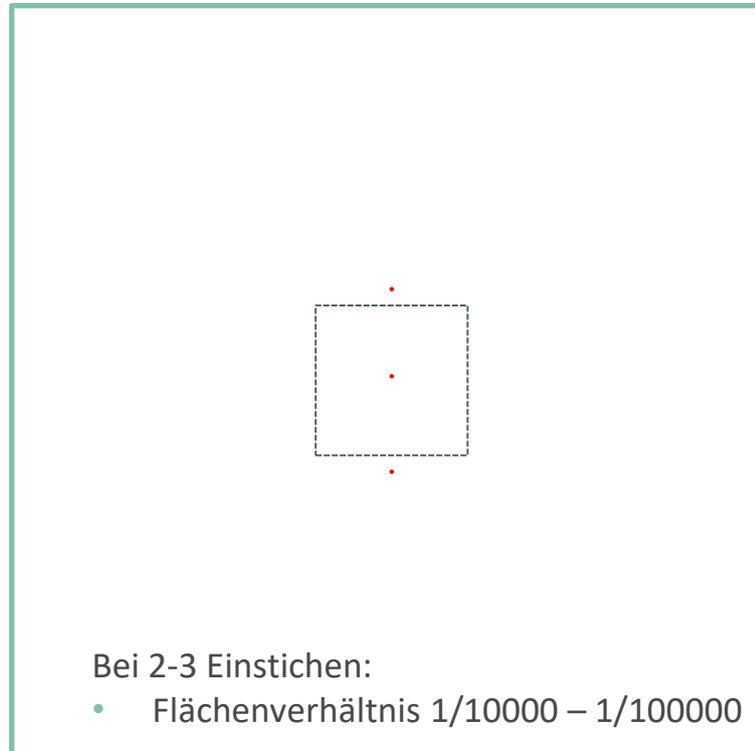


# Generalisierung und Datenfehler

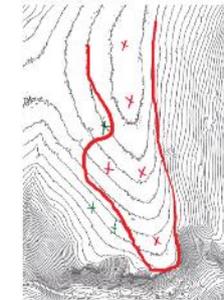
- Nicht direkt vergleichbare Datensätze
- **Generalisierte Karte**
  - Auf Basis bereinigter Daten
- **Validierungspunkte**
  - Kleinräumige Variabilität
  - Nicht bereinigt
  - Zu kleiner Support
- Multi-dimensionale Konfidenzintervallanalyse
  - Systematische Reduktion um (0)1-3 Punkte

Beim Kartieren ist es sehr wichtig, die **Gesetzmässigkeiten zwischen Ausgangsmaterial, Relief und Bodenform zu erkennen** und **mit Analogieschlüssen zu arbeiten** (Abb. 7.3f). **So sollten die Kartierenden versuchen**, in ähnlichen Lagen auf gleichem Ausgangsmaterial die gleichen Böden zu kartieren und **sich nicht durch lokal geringe Bodenunterschiede** oder durch Bohrungen in Übergangsbereichen **verunsichern zu lassen**.

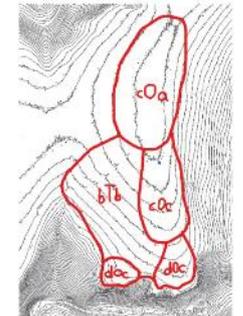
Schriftenreihe der FAL (24), 1997



• KA23



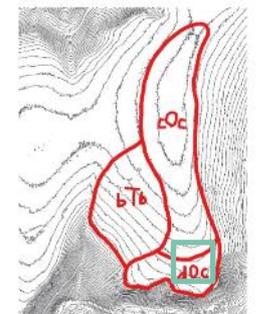
«stur» entlang von Einzelbohrungen gezogene Grenze



Das Landschaftselement Kuppe wird durch ungünstige Grenzziehung der Kartiereinheiten «zerbrochen»



Generalisierte Grenze mit Rücksicht auf die Kuppenlage

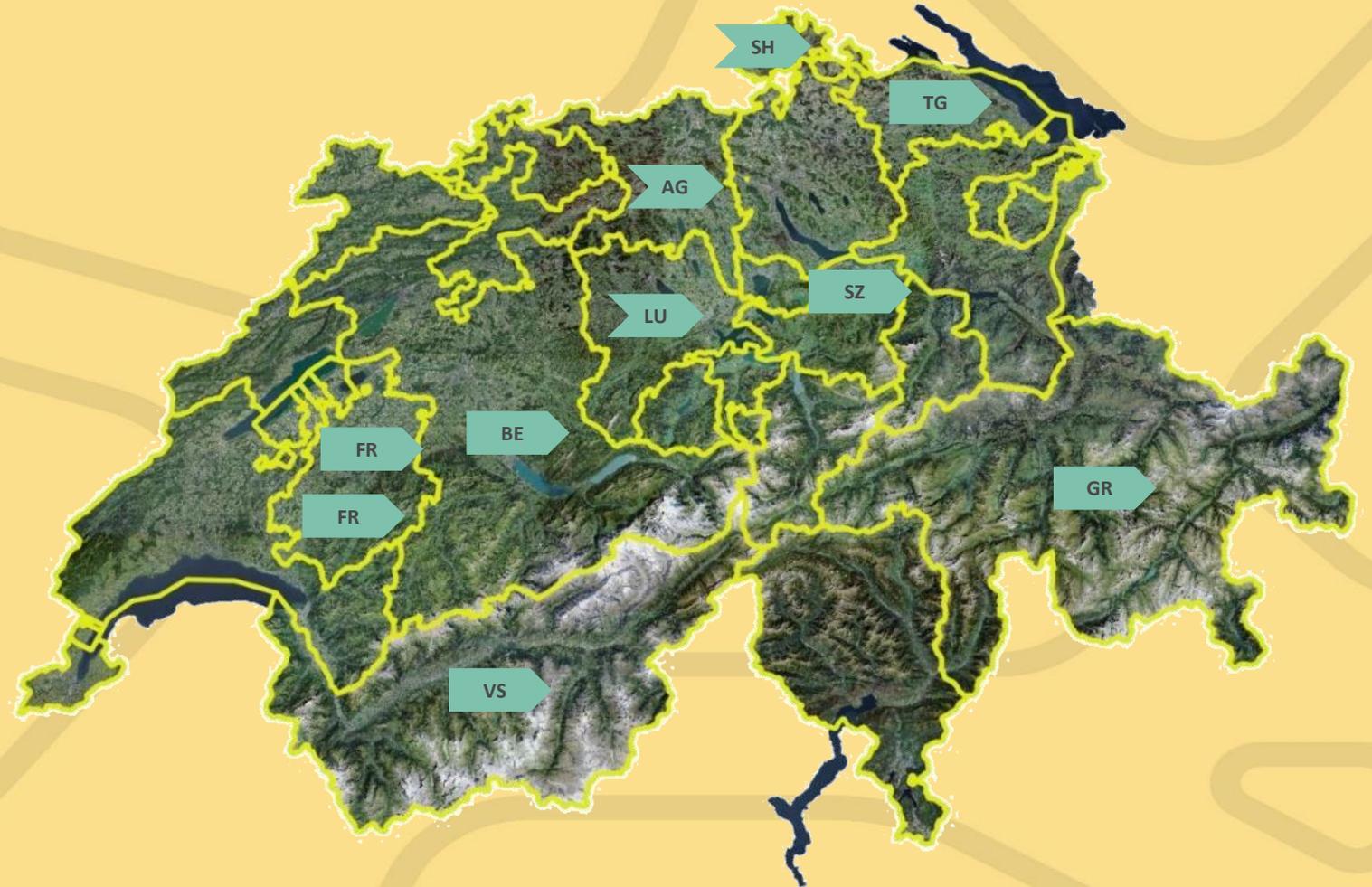


Bei der Grenzziehung wurde das Landschaftselement Kuppe berücksichtigt

Sigrist und Marugg, 2023

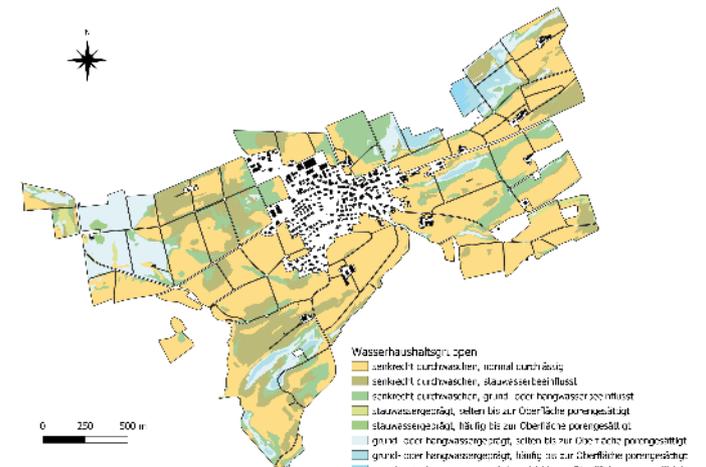
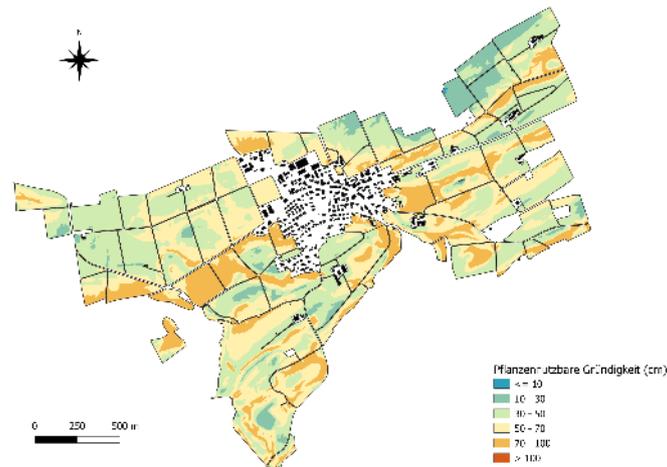
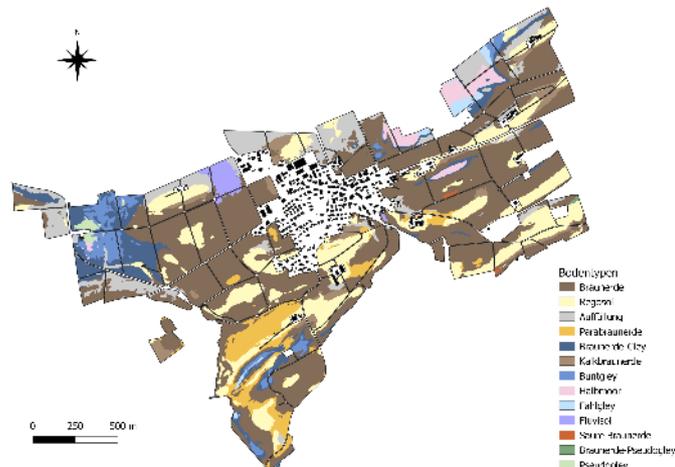
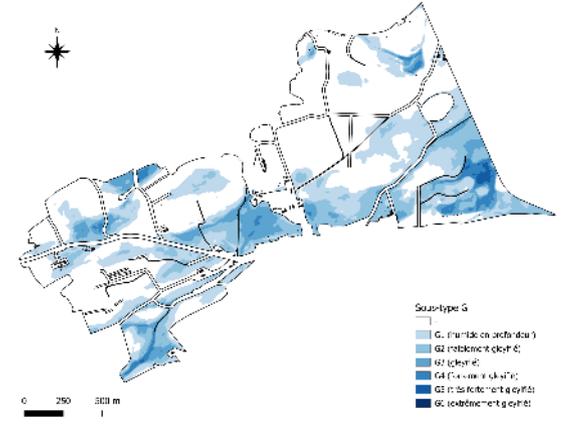
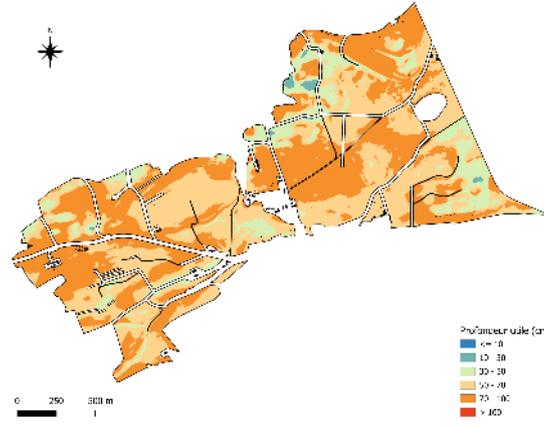
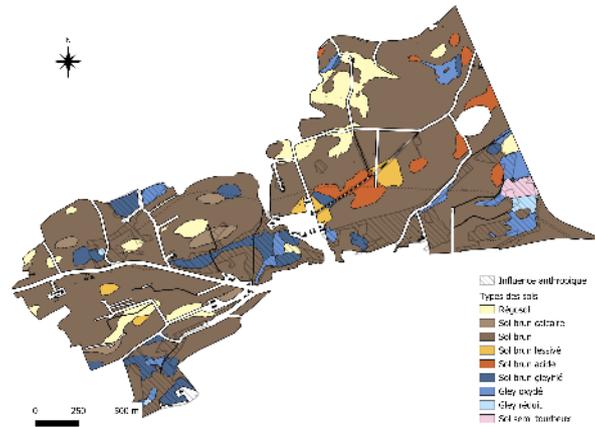
## Ergebnisse

- 2020 - 2025



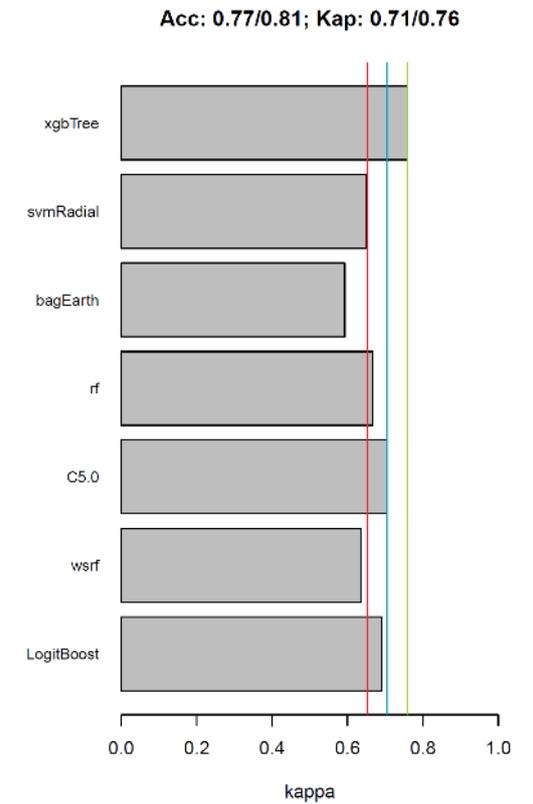
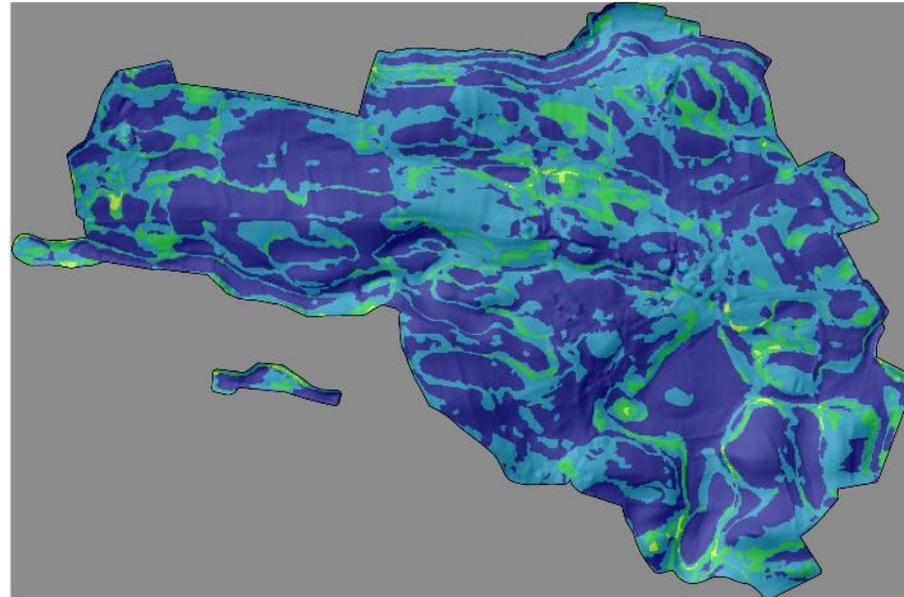
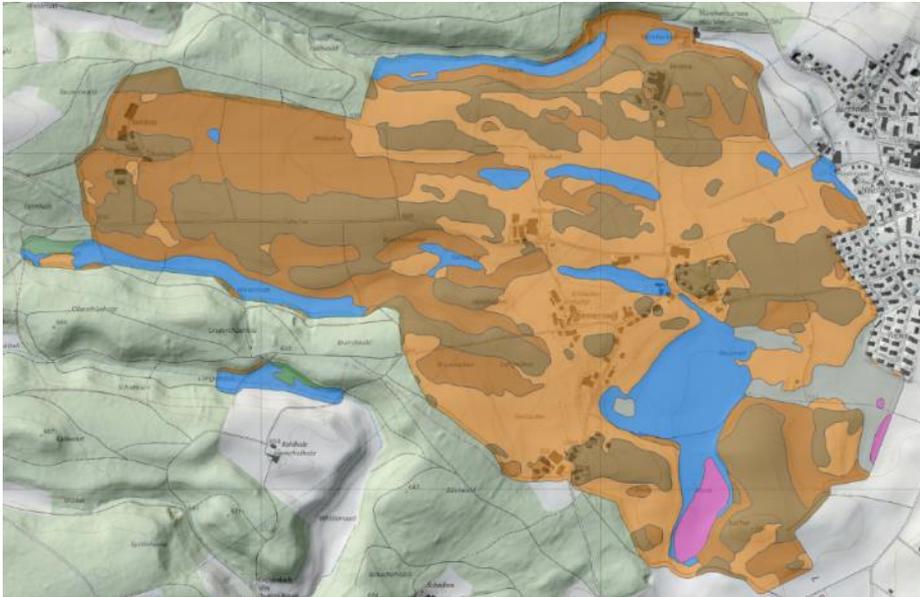
# Ergebnis

- Raster- und Polygonkarten
- Bodenkarten aus H2



# Ergebnis

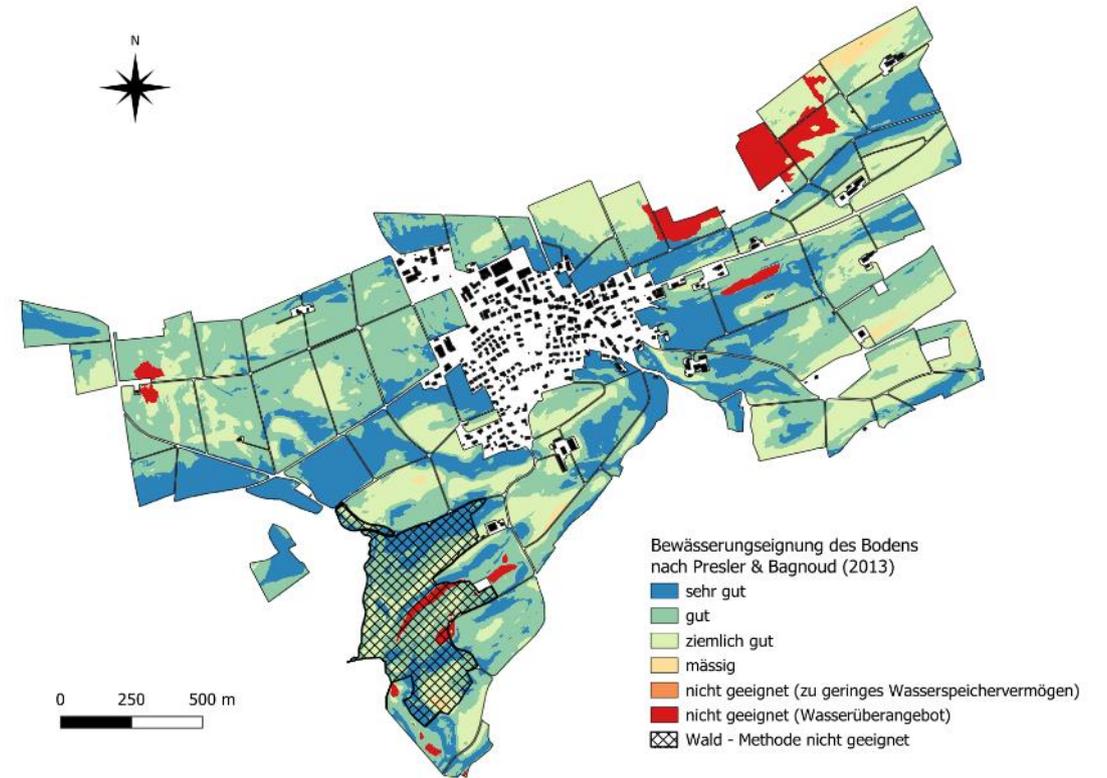
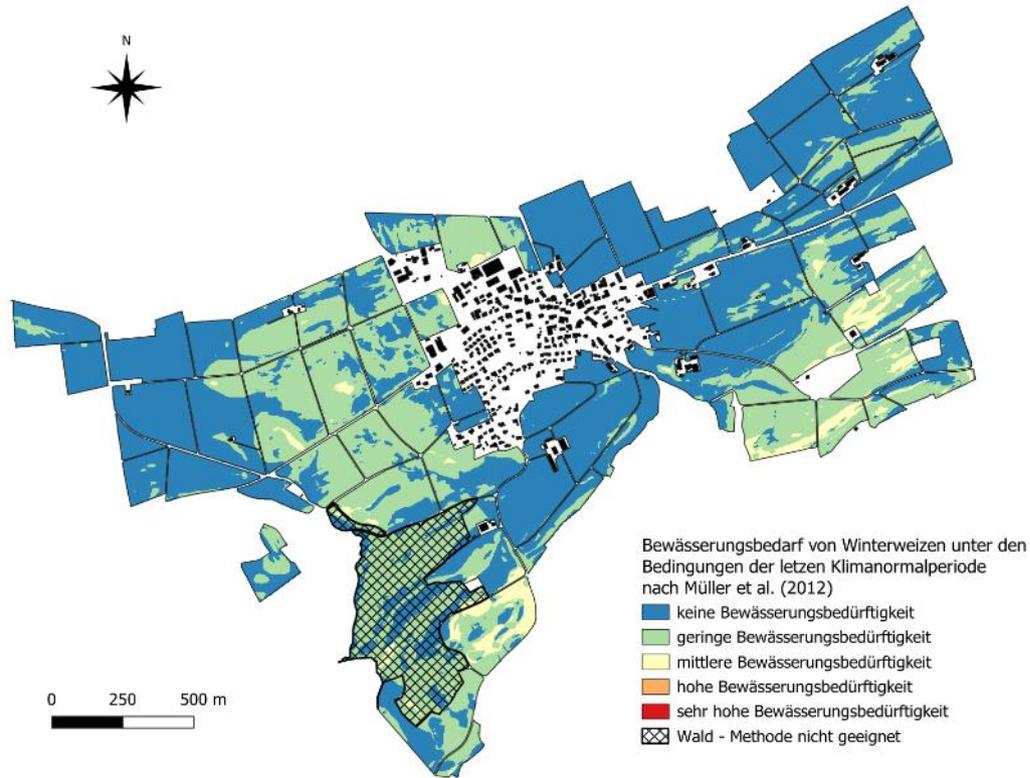
- Raster- und Polygonkarten
  - Unsicherheiten und Modellkombinationen



# Ergebnis

- Raster- und Polygonkarten
- Themenkarten

Bild: Bodentyp und Unsicherheiten



## Bei all der Digitalisierung, Skalierung und Systematisierung: Im Zentrum stehen der Mensch und der Boden

