Effet de la densité des sondages sur la modélisation de la profondeur utile, exemple du projet Longirod (VD)

Colloque groupe de travail KLABS, KA et Boden 4.0

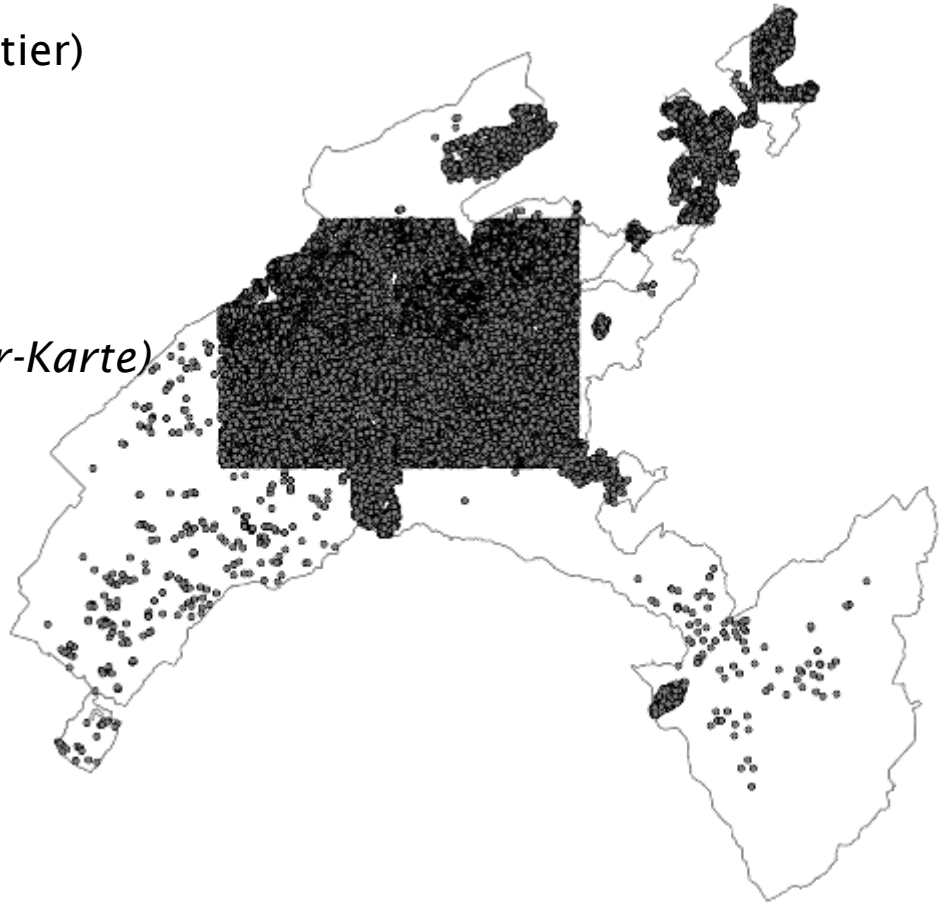
S. Burgos, M Nussbaum, S. Oechslin, T Sprafke, U. Koch

Cadre général….

Valorisation des données existantes

- 30’000 relevés par Michel Gratier (1970-2000)
- Mandats terminés en 2022
- Carte de prédiction de la PU pour tout le canton de Vaud (Carte Gratier)
- Cibler les zones avec un haut potentiel SDA

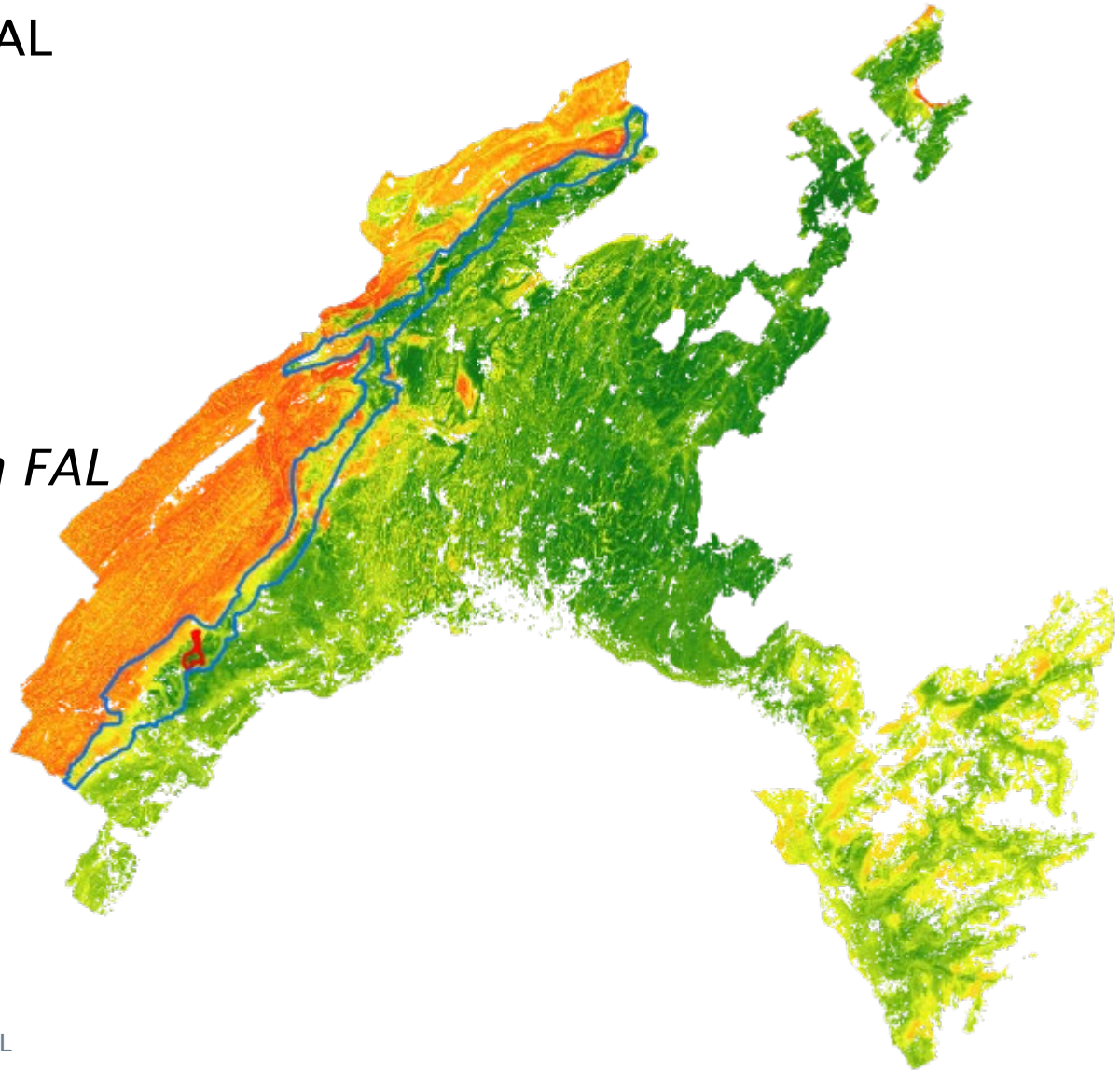
- *30’000 Erhebungen von Michel Gratier (1970-2000)*
- *Mandate 2022 abgeschlossen*
- *Karte der PNG-Vorhersage für den gesamten Kanton Waadt (Gratier-Karte)*
- *Zonen mit hohem FFF-Potenzial gezielt erkennen*



Valorisation des données existantes

- Traduction des 30’000 sondages en FAL
- Calcul de la PU
- Modélisation de la PU
- Terminé fin 2022

- *Übersetzung der 30’000 Bohrungen in FAL*
- *PNG Berechnung*
- *PNG Modellierung*
- *Ende 2022 abgeschlossen*



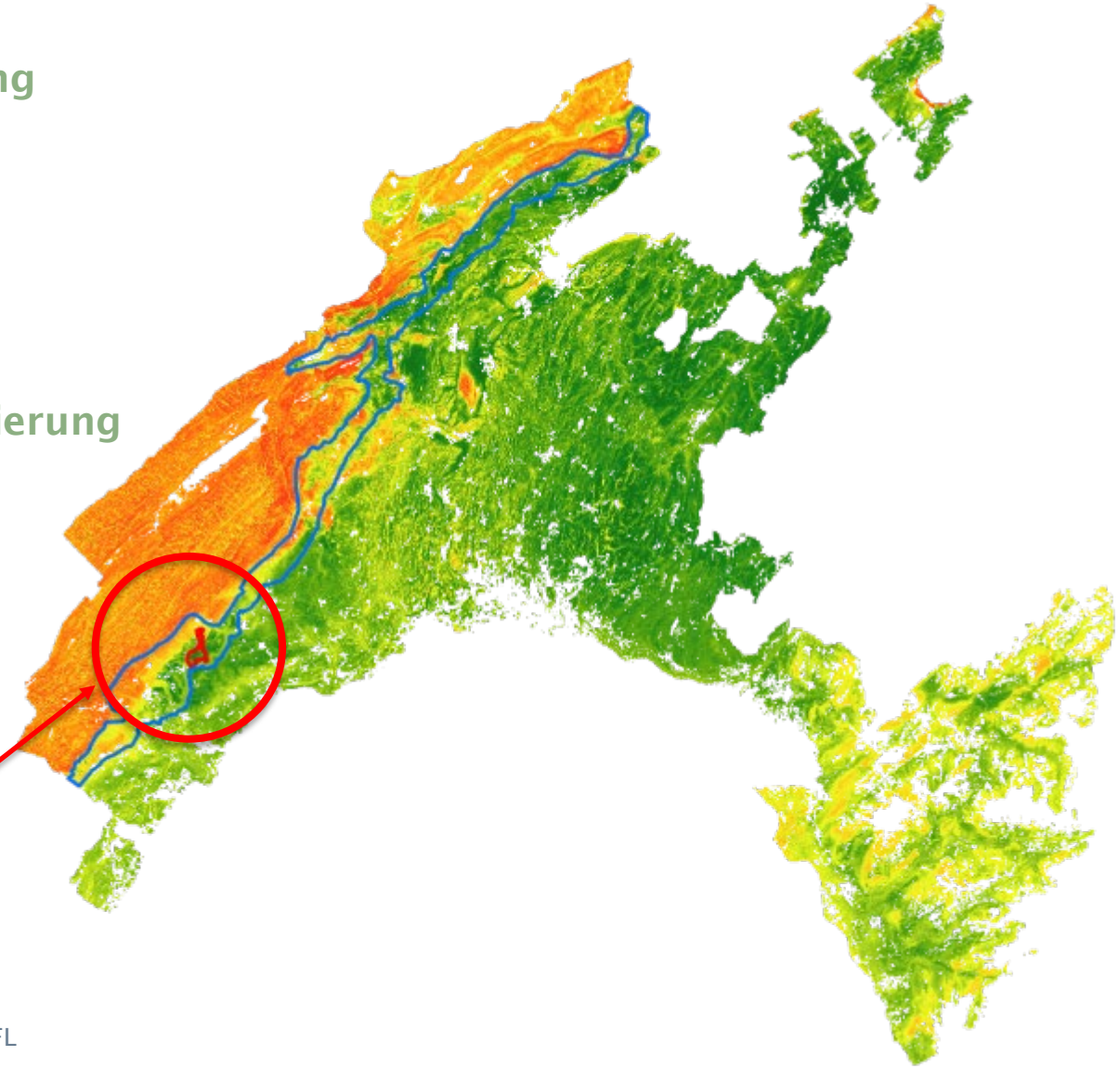
Valorisation des données existantes, projets actuels

Cartographie des sols de détail / Detailbodenkartierung

- En rouge – 200 ha
- Saint-George, Longirod et Marchissy
- Résultats présentés aujourd'hui
- Terminé fin 2024

Cartographie des sols à faible résolution / Grobe Kartierung

- En bleu – 17'000 ha
- Entier du Pied du Jura
- En cours
- Terminé fin 2025

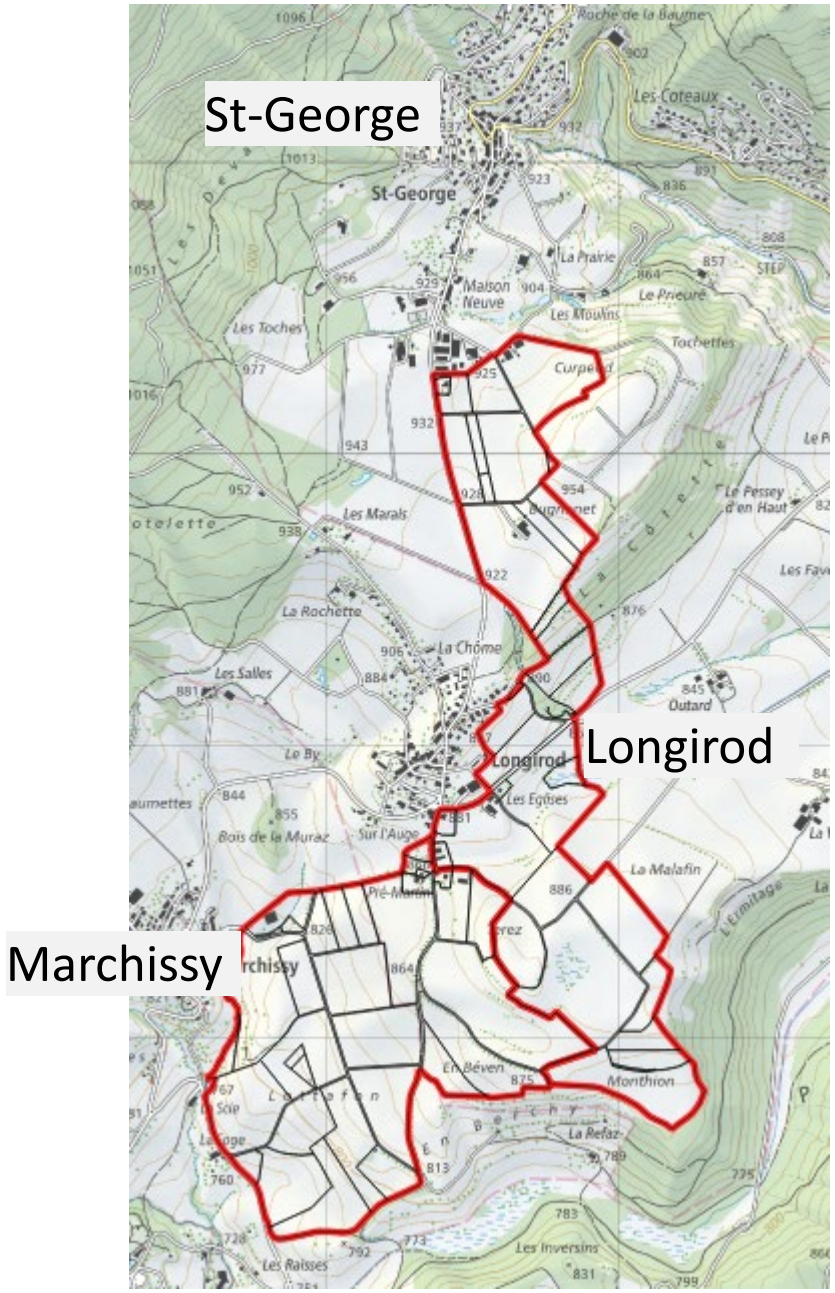


Haute probabilité d'avoir SDA
grosse FFF-Wahrscheinlichkeit

Objectifs de la cartographie de détail

1. Evaluer l'effet de la densité de sondage sur la qualité de la carte
2. Identifier des surfaces SDA
3. Harmoniser les pratiques
4. Valoriser les données obtenues pour les agriculteurs

1. *Den Effekt der Stichprobendichte auf die Kartenqualität bewerten*
2. *FFF-Flächen identifizieren*
3. *Harmonisierung der Bodenkundler*innen*
4. *Die gewonnenen Daten für die Landwirt*innen aufwerten*



Méthodes….

Méthode: prospection et creuse des profils

Expérience préalable (on part rarement de zéro) :

- 270 sondages sur tout le canton / *270 Bohrungen im ganzen Kanton*
- 100 + 50 sondages dans 2 régions proches (2 km) / *100 + 50 Bohrungen in der Umgebung*
- Analyses des informations Gratier / *Analyse der Gratier-Daten*

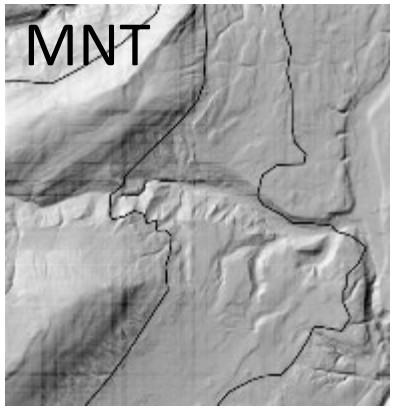
→ **Bonne compréhension de la région**
Gutes Verständnis der Region

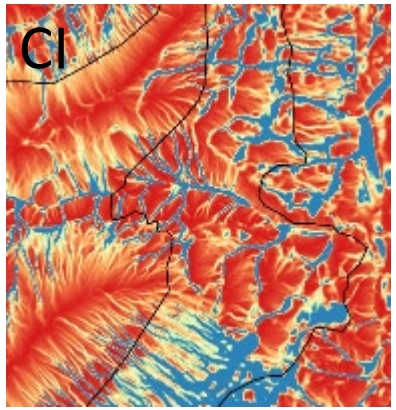
Prospection :

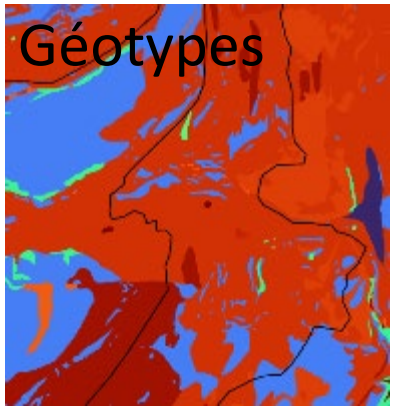
- 12 sondages pour comprendre spécifiquement la zone / *12 Bohrungen für das genauere Verständnis der Zone*
- 70 sondages de prospection avec les bureaux / *70 Explorationsbohrungen mit Büros*
- Choix des 10 profils en fonction des processus pédogénétique / *Wahl von 10 Profilen anhand von Bodenbildungsfaktoren*

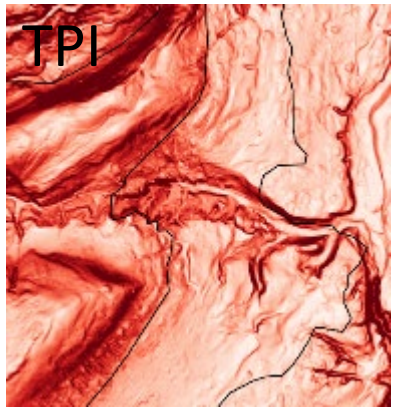
→ **Choix des variables SCORPAN les plus discriminantes (variables et résolutions)**
Auswahl der SCORPAN-Variablen mit dem höchsten Erklärungsgrad (Variablen & Auflösung)

Méthode, clustering









Variables de départ

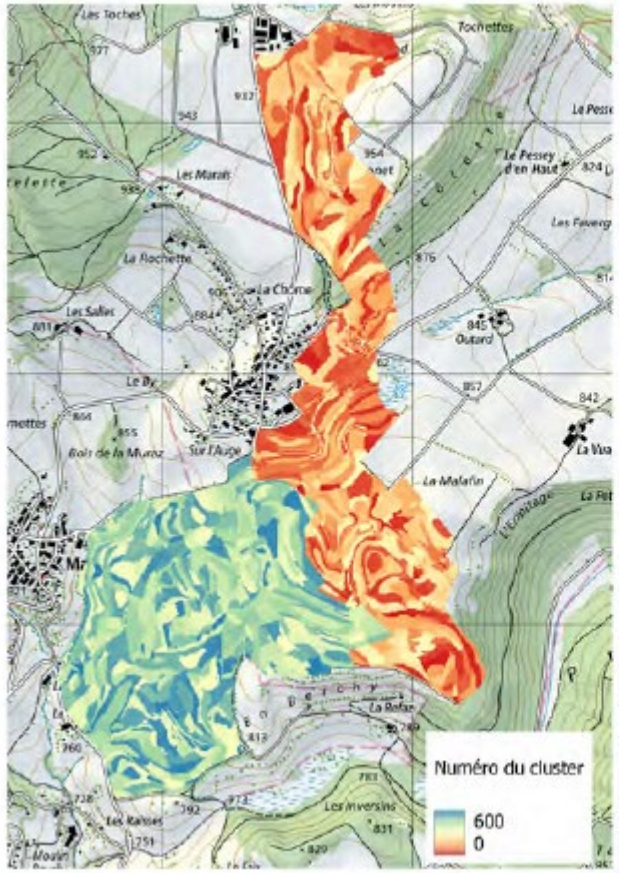
Clustering (après ACP et visite sur le terrain):

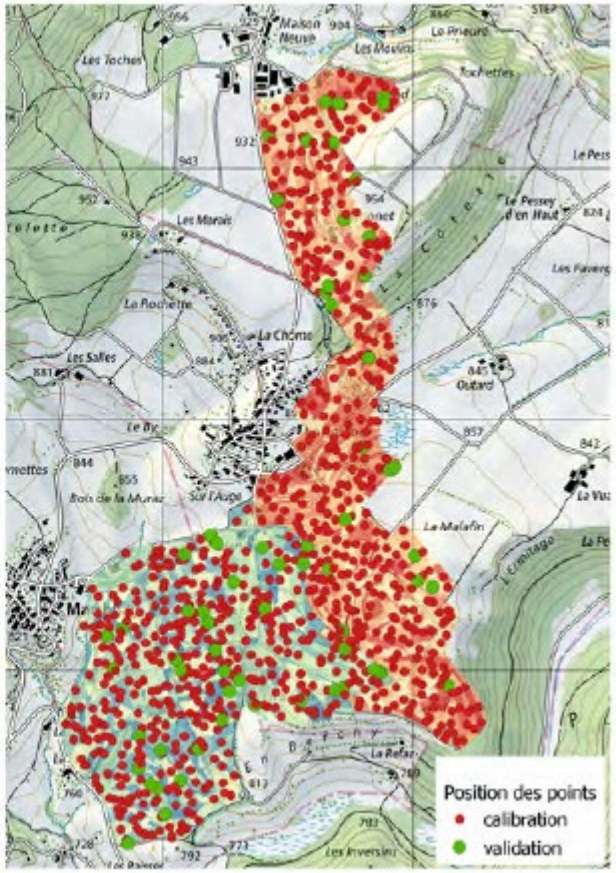
- Indice de position topographique à 50, 200 et 500 m de résolution *(Topographic Position Index, TPI)*
- Indice d'humidité et catchment area à 10 m *(SAGA Wetness Index, SWI)*
- Pentes avec une résolution de 5 et 25 m *(Neigung)*
- Modèle numérique de terrain (MNT) à 5 m *(Höhenmodell)*
- Indice d'appartenance à une vallée à 10 m *(Multiresolution Index of Valley Bottom Flatness, MRVBF)*
- Indice de convergence des flux d'eau à 10 m *(Convergence Index CI, SAGA morphometry)*
- Carte des géotypes *(Geotypen-Karte)*
- Zone fluvioglaciaire décrite dans une étude hydrologique *(Schotterzone aus alter Studie)*

Validation:

- Carte des unités de sol Gratier (stratified random design, independent validation)

Méthode: clustering en niveaux hiérarchiques





2 Lots, 600 clusters, points au centroïdes des clusters

- niveau hiérarchique 1 = densité 1 sondage/ha (150 sondages)
- niveau hiérarchique 2 = densité 2 sondages/ha (150 sondages)
- niveau hiérarchique 3 = densité 4 sondages/ha (300 sondages)

600 points (C3, 300 zone 1 et 300 zone 2)

↓ Clustering des 600

300 points (C2, 150 zone 1 et 150 zone 2)

↓ Clustering des 300

150 points (C1, 75 zone 1 et 75 zone 2)

↓

Sondages utilisés	C1	C2	C3	Z	Nb points
Densité 1/ha	150			50	200
Densité 2/ha	150	150		100	400
Densité 4/ha	150	150	300	200	800

Z = points complémentaires par les bureaux / *zusätzliche Punkte durch die Büros*

Méthode, terrain

Système de description standard

- Profils: complets, digitalisés sur soildat, tous les paramètres + Analyses
- Sondages: description par horizon, digitalisés sur soildat / *horizontweise, soildat*
- Éléments grossiers -> quelques calibrations / *Skelett Eichung*

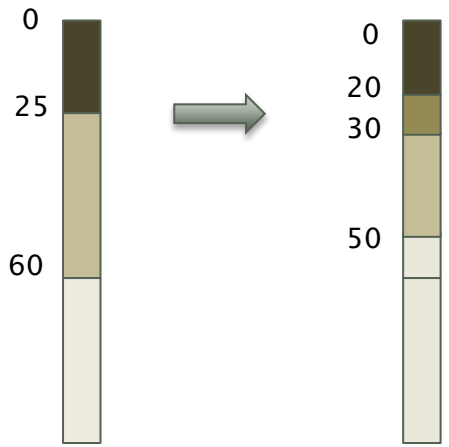
Faces hétérogènes, lentille de sable, gros blocks… où prélever ?

Heterogene Flächen, Sandlinse, große Blöcke...wo beproben ?

Méthode, modélisation

Propriétés simples:

- Teneurs en argile, silt, sable et matière organique (0-20, 0-30, 30-50, 50-100) valeurs des couches dérivées avec pondération des horizons
- Profondeur d'apparition d'un horizon gg (minPgg) ou d'un horizon g (minPg) :
- Profondeur d'apparition d'un horizon r (minPr) :
- Profondeur maximale atteinte (Pmax)



Basiseigenschaften:

- *Ton-, Schluff-, Sand- und organischer Gehalt (0-20, 0-30, 30-50, 50-100) Werte der abgeleiteten Schichten mit Gewichtung der Horizonte.*
- *Tiefe des Auftretens eines gg-Horizonts (minPgg) oder eines g-Horizonts (minPg)*
- *Tiefe des Auftretens eines r-Horizonts (minPr)*
- *Maximal erreichte Tiefe (Pmax)*

Méthode, modélisation

Propriétés composites:

- Profondeur utile calculée (PU)
- Compatibilité SDA (FFF) (carte raster et vectorielle avec polygones) [classes de profondeur ou oui/non].
- Quantification de l'incertitude associée à chaque pixel pour chaque paramètre prédit. Intervalles de prédiction (70%, 80% ou 90 %) ou indice d'incertitude (coefficient de variation) par Quantile Regression Forest.
- NEK2023
- Types et sous-types de sols
- Réserve en eau utile maximale selon KA5, USDA, GEPPA

Zusammengesetzte Eigenschaften:

- *Berechnete pflanzennutzbare Gründigkeit (PNG).*
- *FFF-Kompatibilität (Raster- und Vektorkarten mit Polygonen) [Tiefenklassen oder ja/nein].*
- *Quantifizierung der Unsicherheit, die mit jedem Pixel für jeden vorhergesagten Parameter verbunden ist. Vorhersageintervalle (70%, 80% oder 90%) oder Unsicherheitsindex (Variationskoeffizient) durch Quantile Regression Forest.*
- *NEK2023*
- *Bodentyp / UT*
- *Maximaler nutzbarer Wasservorrat nach KA5, USDA, GEPPA*

Assurance qualité

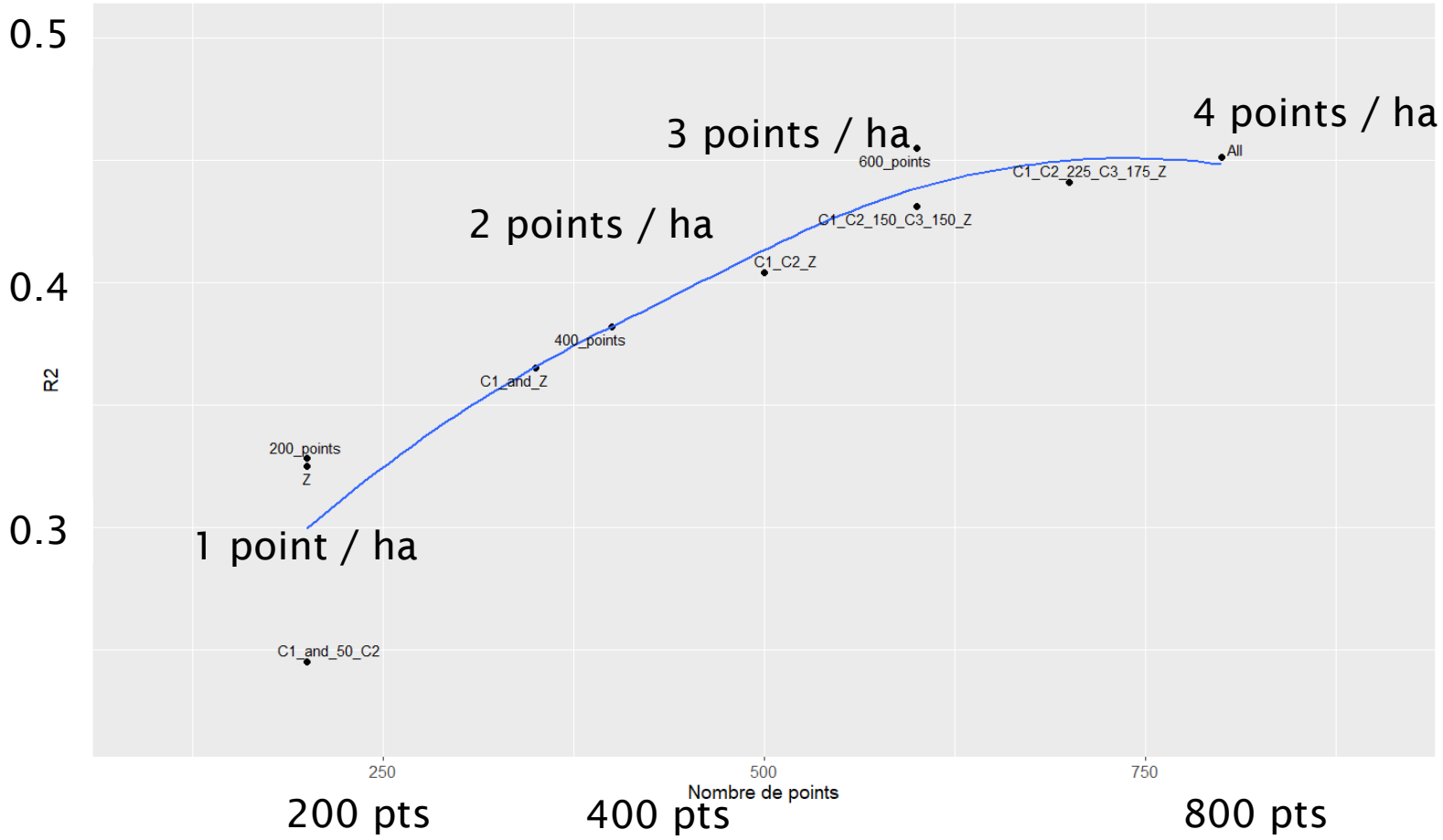
Les étapes suivantes ont été effectuées afin d'assurer la réalisation du projet et le contrôle de qualité des données :

- Description des profils sur le terrain par les bureaux d'études avec prélèvement des échantillons.
- Mise au propre des données des profils et préparation de la campagne de terrain.
- Suivi régulier des responsables cantonaux et de la HAFL sur des points techniques.
- Choix de pédologues expérimentés pour réaliser les relevés de terrain.
- Utilisation de la plateforme Soildat qui permet de collecter les données de façon efficace et exhaustive ainsi que de rendre les données facilement consultables pour les différents acteurs du projet.
- Prise de conscience des bureaux d'études que la qualité et l'homogénéité des relevés pédologiques est le point clé pour que la modélisation des cartes soit réussie.
- Prise de connaissance du terrain pour les bureaux d'études avec les experts de la HAFL (description du paysage, des principales unités géomorphologiques, et d'un premier étalonnage autour de quelques sondages).
- Description de 70 sondages de prospection par les bureaux permettant de définir l'emplacement des profils. Les discussions autour du choix des profils ont permis d'intégrer les bureaux d'études dès le début et d'avoir une première calibration des différents pédologues.
- Description de tous les profils par deux expertes indépendantes (Nathalie Dakhel et Anina Schmidhauser).
- Journée de calibration autour des profils avec les deux expertes et avec un compte rendu qui résumait les principales décisions pour la suite (par ex. Quelles faces décrire ? Comment nommer le matériel parental (moraine remaniée) ?). Lors de cette journée, un sondage à la tarière a été effectué à côté de chaque profil afin de bien faire le lien entre le relevé à la tarière et le profil.
- Établissement d'un rapport automatique sur la qualité des données relevées sur Soildat et d'une correction des éventuelles erreurs par les bureaux d'études.

Les résultats….

Objectif I : Evaluation de l'effet de la densité de sondages (PU / PNG)

R^2 de la validation indépendante, n=65 / unabhängige Validierung

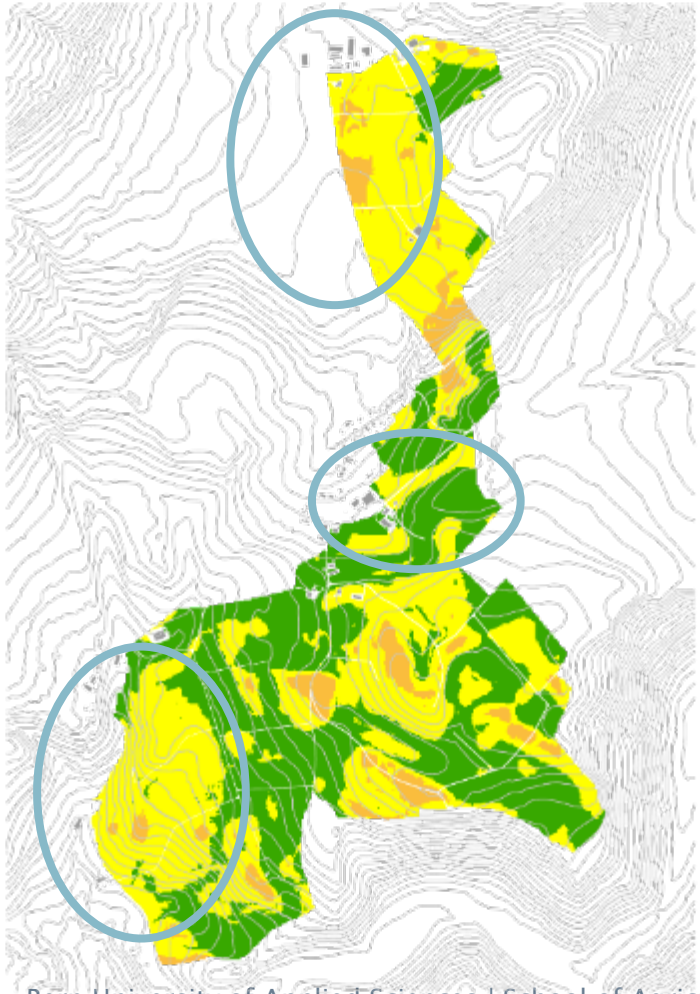


Perte d'informations/
Informationsverluste

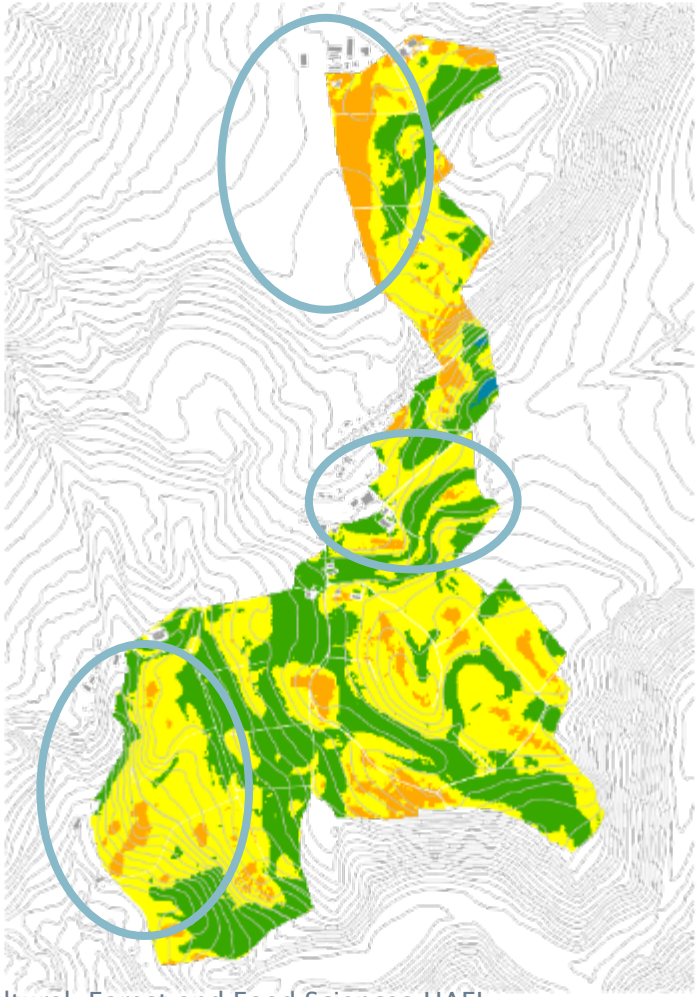
Qu'est-ce que l'on perd
et pourquoi ?
Was verlieren wir
und warum ?

Objectif I : Evaluation de l'effet de la densité de sondages (PU / PNG)

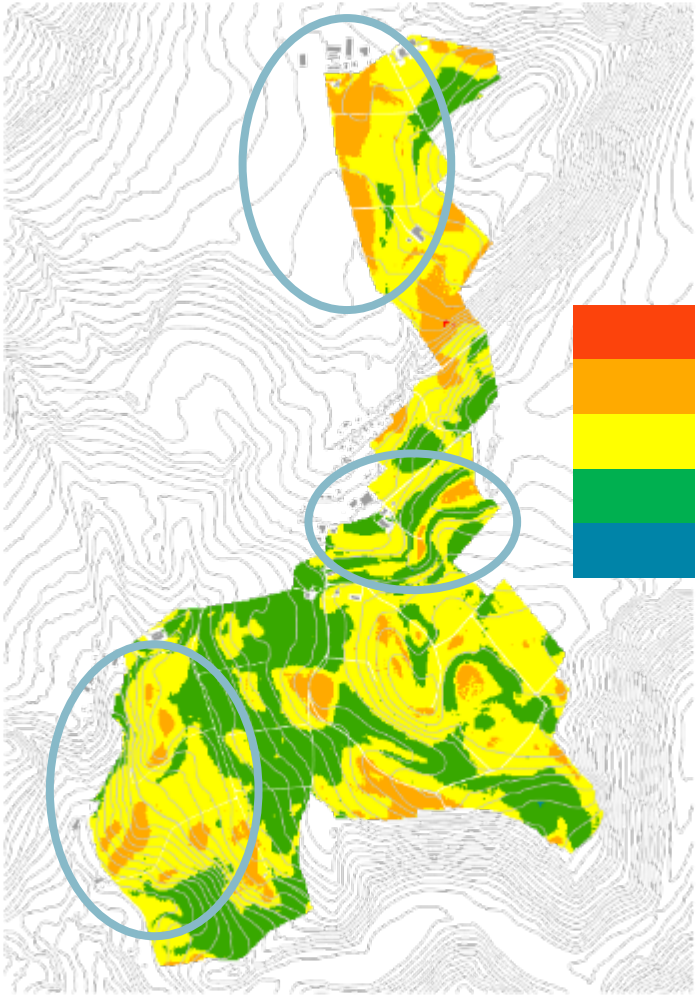
1 sondage/ha



2 sondages/ha



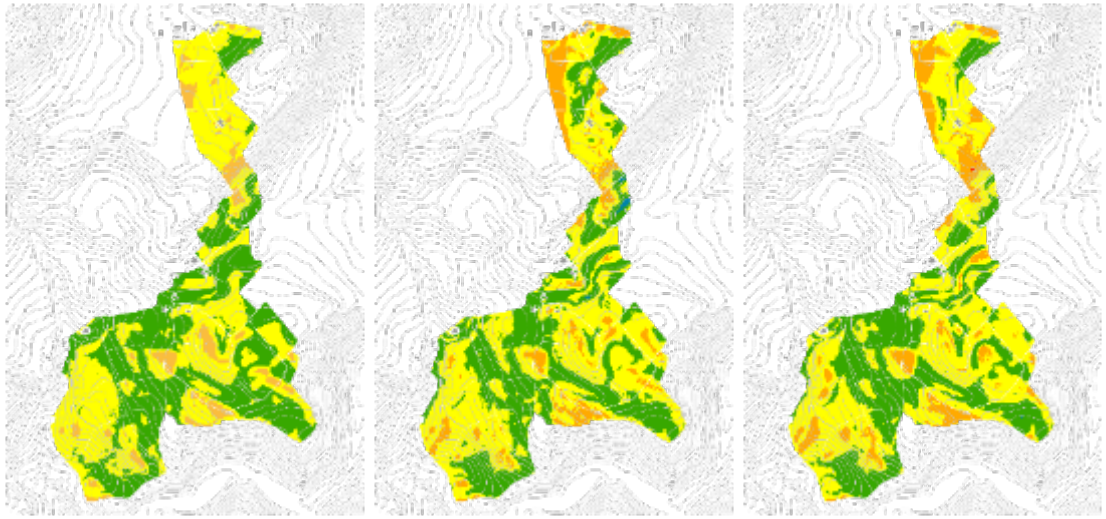
4 sondages/ha



PU / *PNG*

- <30 cm
- 31 à 50 cm
- 51 à 70 cm
- 71 à 100 cm
- >101 cm

Objectif I : Evaluation de l'effet de la densité de sondages (PU / PNG)



Dans quelle situation on perd de l'information /
Wo geht Information verloren?

Lorsqu'il y a de petites variations du relief qui induisent un changement des processus pédologiques

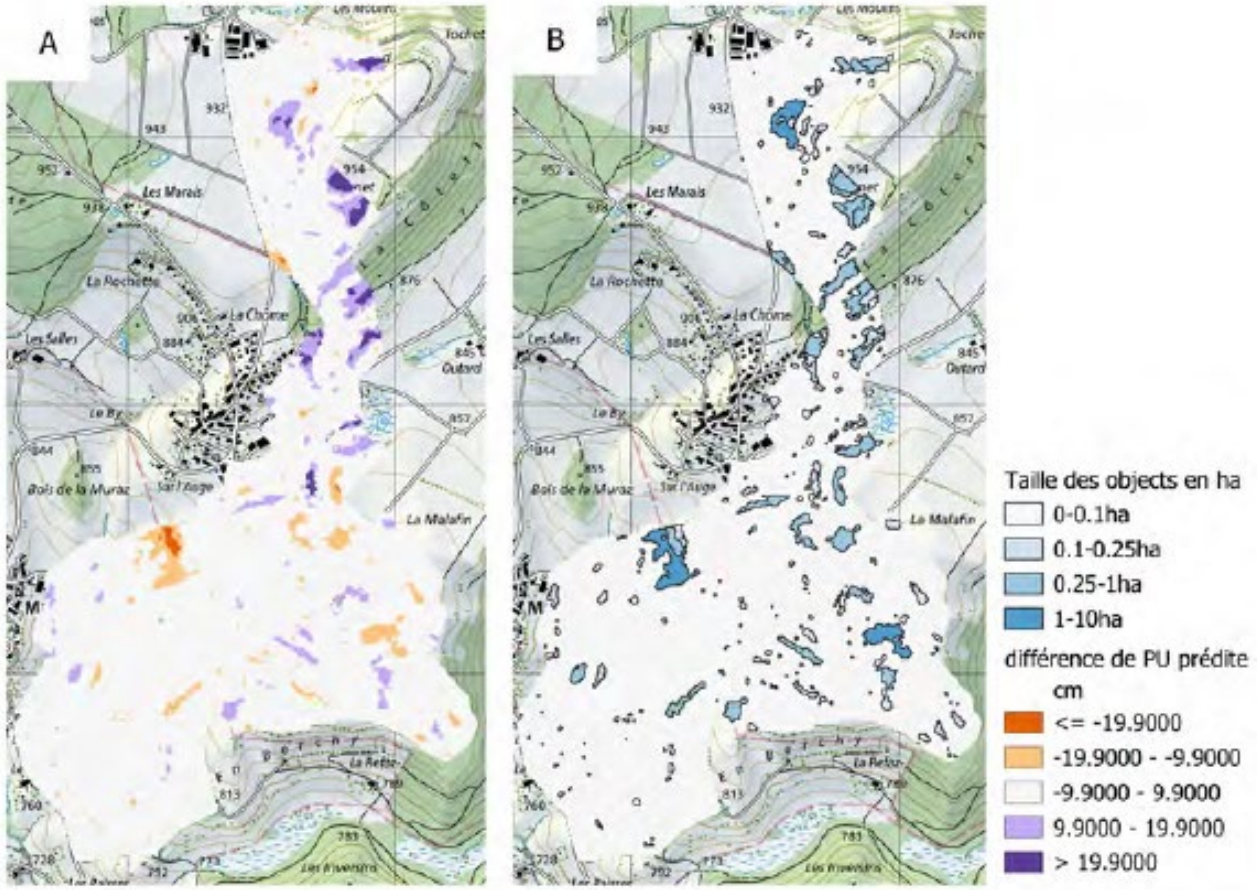
Wenn es kleine Reliefveränderungen gibt, die eine Veränderung der Bodenprozesse bewirken

Comparaison quantitative avec la carte 4 sondages/ha

- 2 sondages/ha: 12% d'erreur*
- 1 sondage/ha: 20% d'erreur* → objets < 2-3ha pas identifiés / 2-3 ha Objekte fehlen

**Erreur = différence de plus de 10 cm de PU sur un pixel par rapport à la carte 4 sondages/ha*
Fehler wenn > 10cm PNG Unterschied im Vergleich zur Karte 4pkt/ha besteht

Objectif I : Evaluation de l'effet de la densité de sondages (PU / PNG)



Dans quelle situation on perd de l'information /
Wo geht Information verloren?

Les zones différant de plus de 10cm pour la PU ne sont pas très étendues et souvent inférieures à 1ha

Gebiete, die sich bei PNG um mehr als 10 cm unterscheiden, sind nicht sehr groß und oft kleiner als 1 ha

Objectif I : Evaluer la densité de sondages

Fréquence de la surface totale occupée par des zones possédant une classe différente de celle prédite par la carte avec la densité de sondage maximale (800 points).

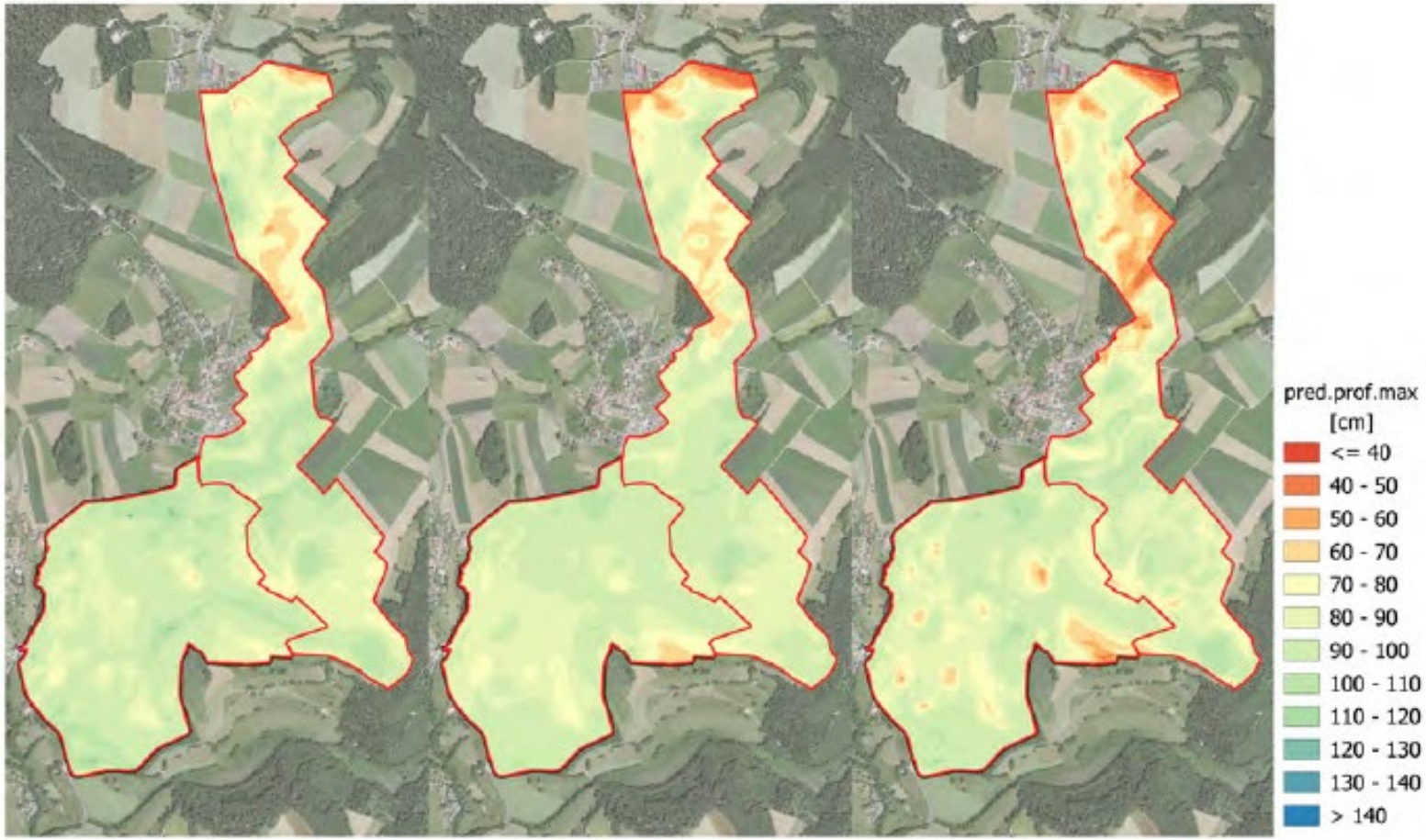
Häufigkeit der Gesamtfläche, die von Gebieten eingenommen wird, die eine andere Klasse besitzen als die, die von der Karte mit der höchsten Stichprobendichte (800 Punkte) vorhergesagt wird.

0 à 0.1ha	0.1 à 0.25ha	0.25 à 1ha	1 à 10ha	nb de sondages	
0.02	0.02	0.06	0.07	200 points	
0.02	0.02	0.05	0.02	400 points	
0.01	0.00	0.00	0.00	600 points	
0.00	0.00	0.00	0.00	800 points	référence
0.02	0.03	0.08	0.06	C1_and_50_C2	
0.02	0.01	0.03	0.01	C1_and_Z	
0.01	0.01	0.02	0.01	C1_C2_150_C3_150_Z	
0.01	0.00	0.00	0.00	C1_C2_225_C3_175_Z	
0.01	0.01	0.02	0.00	C1_C2_Z	
0.02	0.03	0.04	0.10	Z	

N'apparaitrait pas au 1:5000
würde nicht in 1:5000 erscheinen

Max 14 % de différence
Max 14% Unterschied

Objectif I : Evaluer la densité de sondages

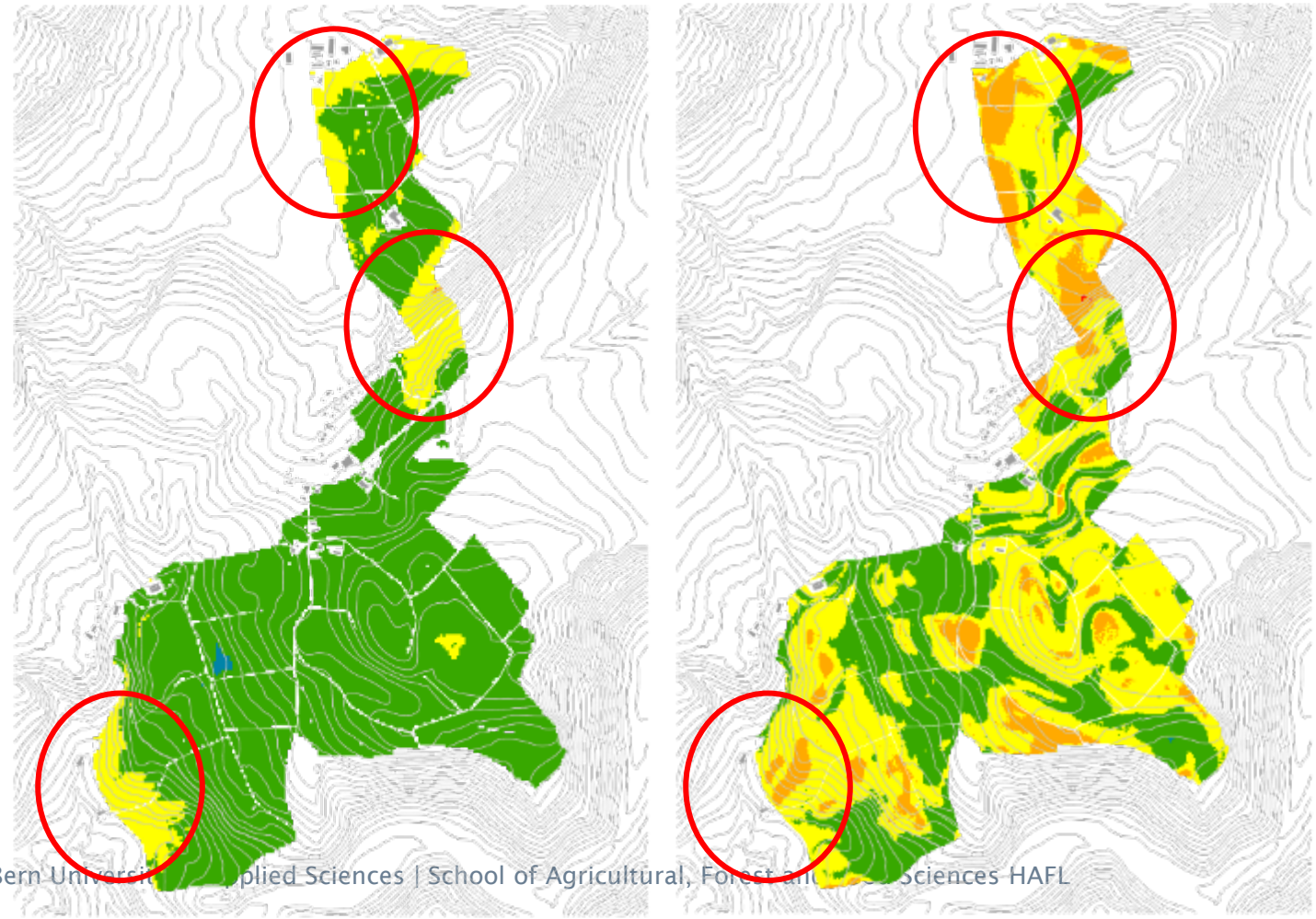


- Même tendance, plus de sondages ont été bloqués par les éléments grossiers
- *Gleicher Trend, mehr Bohrungen wurden durch Skelett blockiert*

Objectif I : Evaluer la densité de sondages

Carte Gratier modélisée, 30’000 sondages

4 sondages/ha



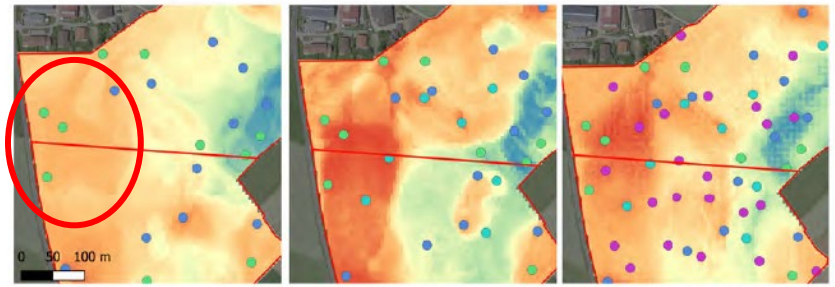
PU
- <30 cm
- 31 à 50 cm
- 51 à 70 cm
- 71 à 100 cm
- >101 cm

- Même tendance, mais surestimation de la PU
- La carte Gratier est une bonne carte conceptuelle pour la PU

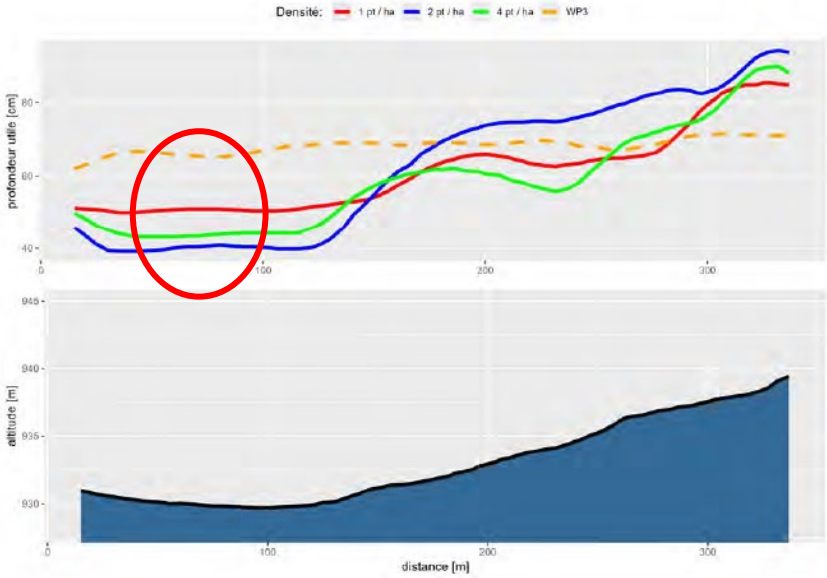
- *Gleiche Tendenz, aber Überschätzung der PNG.*
- *Die Gratier-Karte ist eine gute Konzeptkarte für die PNG*

Objectif I : Evaluer la densité de sondages

- Avec 1 sondage / ha les zones humides sont moins bien discriminées
- *Mit 1 Bohrung / ha werden vernässte Zonen weniger gut beschrieben*



(a) Détails de la carte des profondeurs utiles (PU) de la zone d'anciens marais au nord du périmètre. À gauche, 1 sondage /ha (C1 et 50 Z), au centre 2 sondages /ha (C1 + C2 + 100 points Z, à droite 4 sondage /ha (C1, C2, C3 et 200 points Z). Les combinaisons sont détaillées dans le tableau 1.

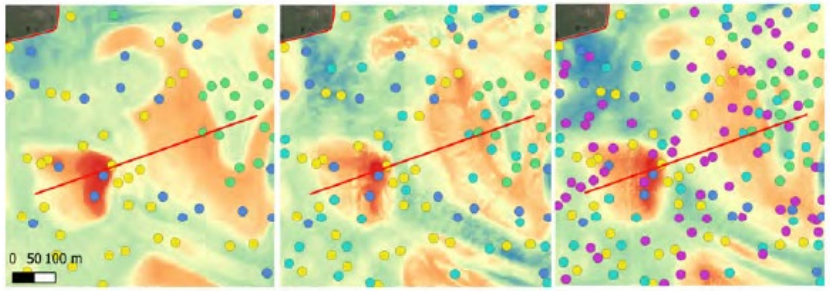


(b) En haut, profil des profondeurs utiles prédites en fonction de la densité de sondages et modélisée avec les 30'000 anciens sondages Gratier (WP3=Workpackage 3 du projet de valorisation des anciennes données Gratier) dans la zone des anciens marais au nord du périmètre. En bas, coupe de l'altitude décrivant la topographie.

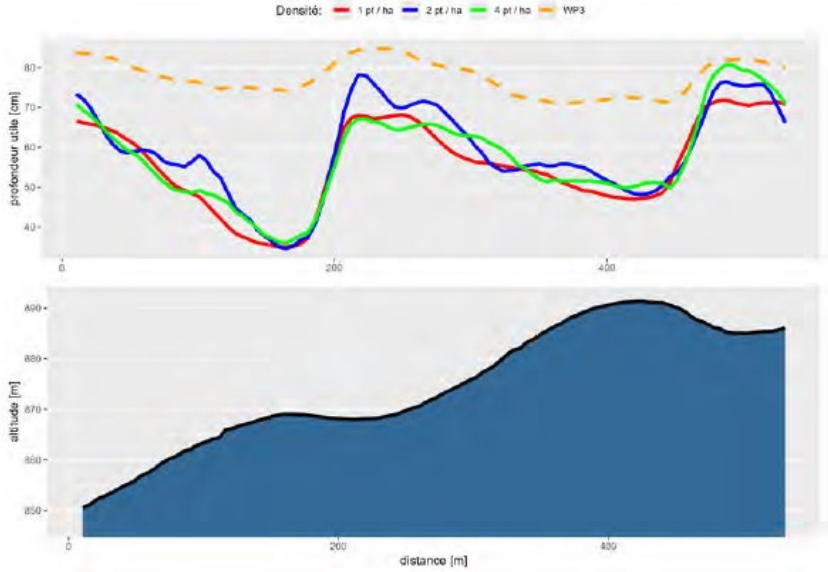
Objectif I : Evaluer la densité de sondages

- Si le relief est le principal facteur et la géologie similaire la densité des sondages est moins importante

- *Wenn das Relief der Hauptfaktor ist und die Geologie ähnlich die Bohrungdichte ist weniger wichtig*



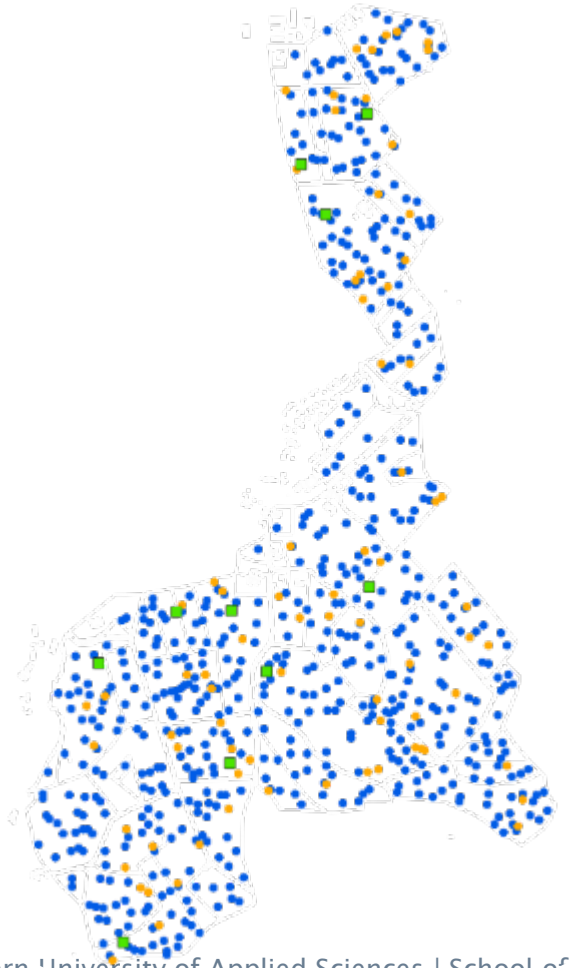
(a) Détail de la carte des profondeurs utiles (PU) de la colline (Drumlin) au sud est du périmètre. À gauche, 1 sondage /ha (C1 et 50 Z), au centre 2 sondages /ha (C1 + C2 + 100 points Z, à droite 4 sondage /ha (C1, C2, C3 et 200 points Z). Les combinaisons sont détaillées dans le tableau 1.



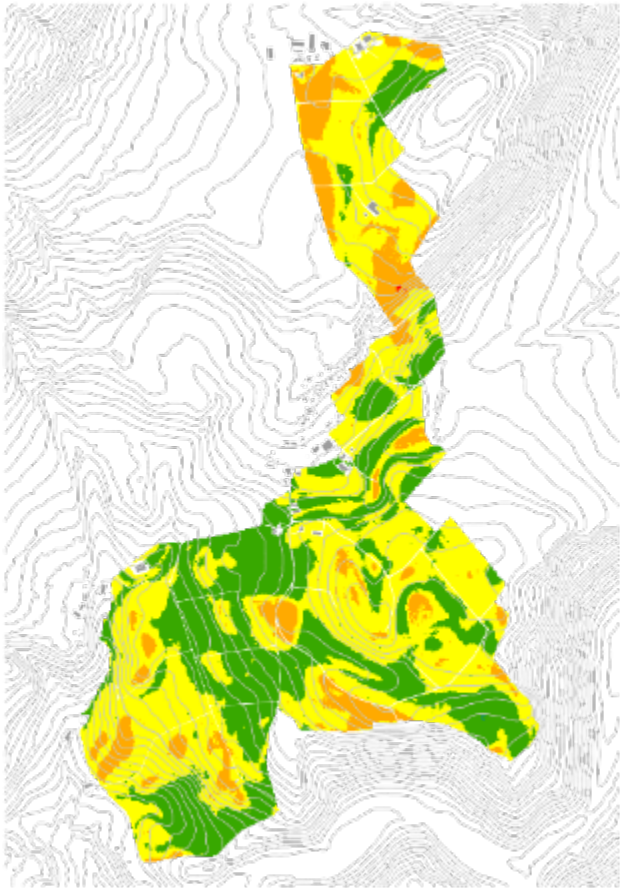
(b) En haut : profil des profondeurs utiles prédites en fonction de la densité de sondages et modélisée avec les 30'000 anciens sondages Gratier (WP3=Workpackage 3 du projet de valorisation des anciennes données Gratier) dans de la colline (Drumlin) au sud-est du périmètre. En bas : coupe de l'altitude décrivant la topographie.

Objectif II : Identifier des SDA

FAL24 (4 sond/ha)



1. Carte PU cm



2. Incertitude

Objectif II : Identifier des SDA

Le modèle permet de

- Quantifier une incertitude
- Localiser l'incertitude

Pour l'instant

- Approche prudente
- Valeur seuil en réflexion
- Réflexion sur extrapolation à d'autres zones

Das Modell ermöglicht Folgendes

- *Eine Unsicherheit zu quantifizieren*
- *Die Unsicherheit zu lokalisieren*

Zum jetzigen Zeitpunkt

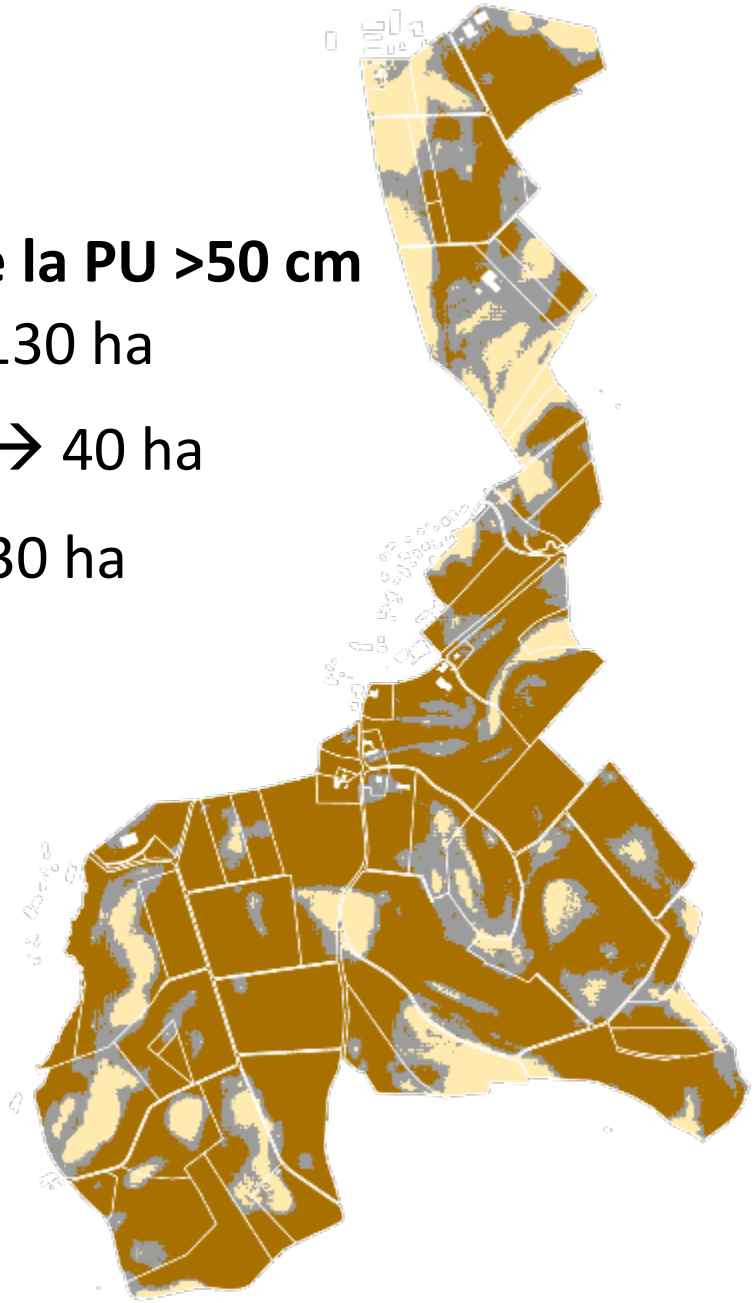
- *Vorsichtiger Ansatz*
- *Reflexion über Schwellenwert*
- *Überlegungen zur Extrapolation auf andere Gebiete*

Probabilité que la PU >50 cm

> 70% → 130 ha

50 à 70% → 40 ha

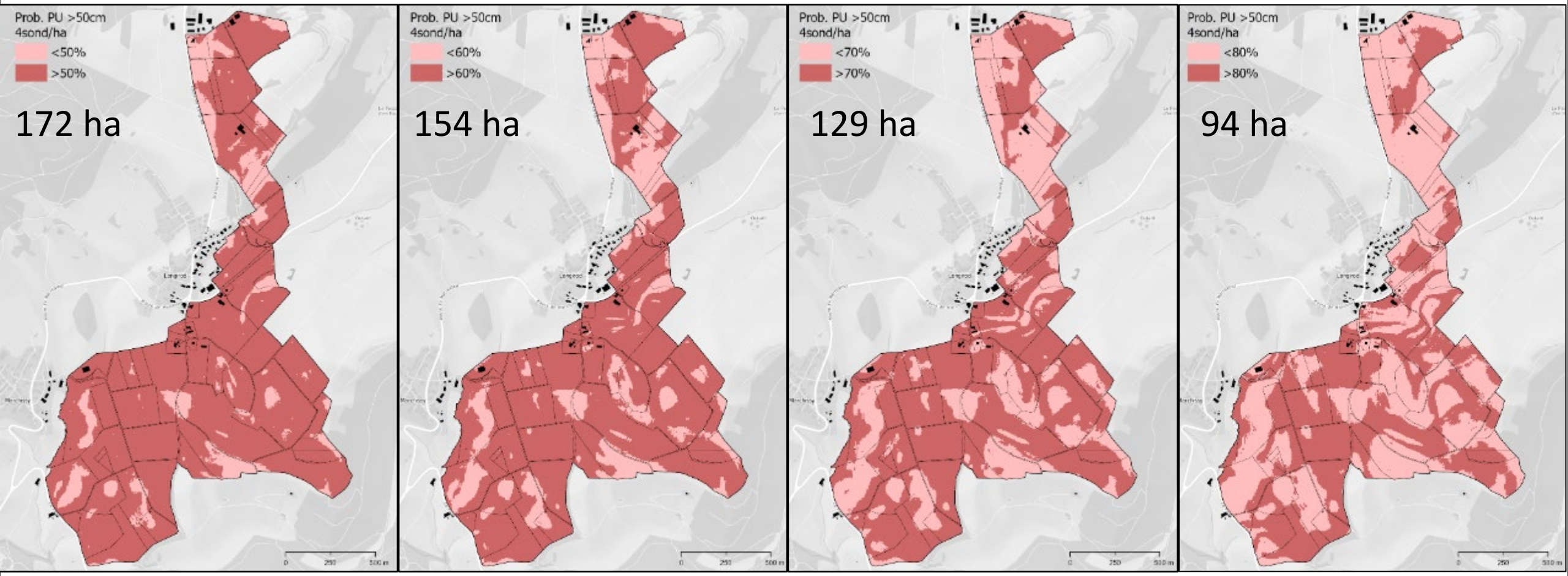
< 50% → 30 ha



Probabilité PU > 50cm – 4 sondages/ha

Une grande probabilité est plus sûre mais réduit les surfaces

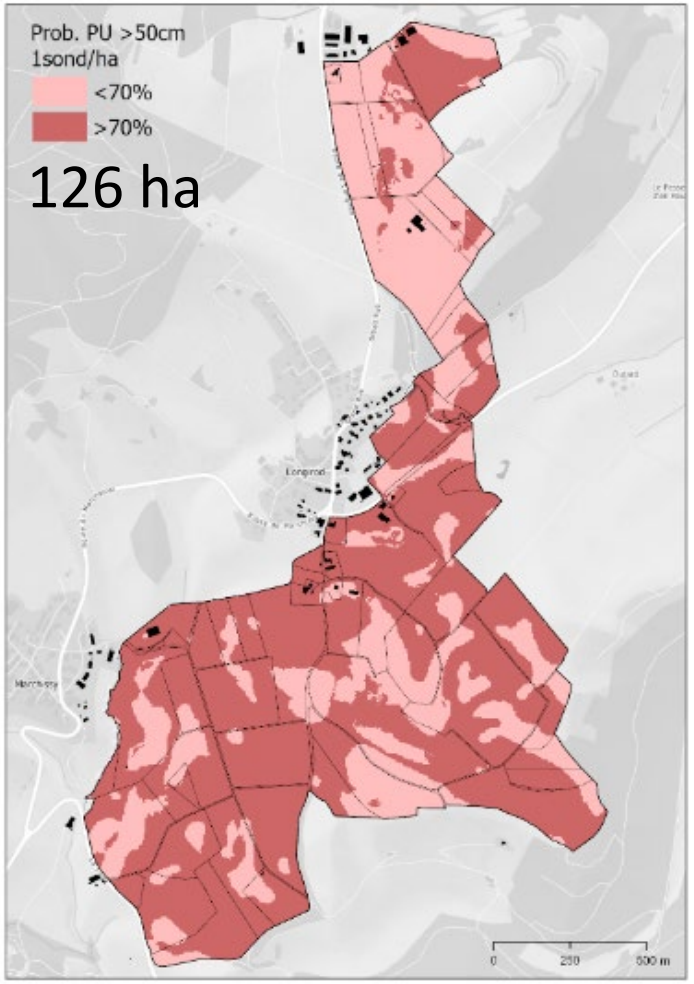
Eine hohe Wahrscheinlichkeit ist sicherer, verringert aber die Fläche

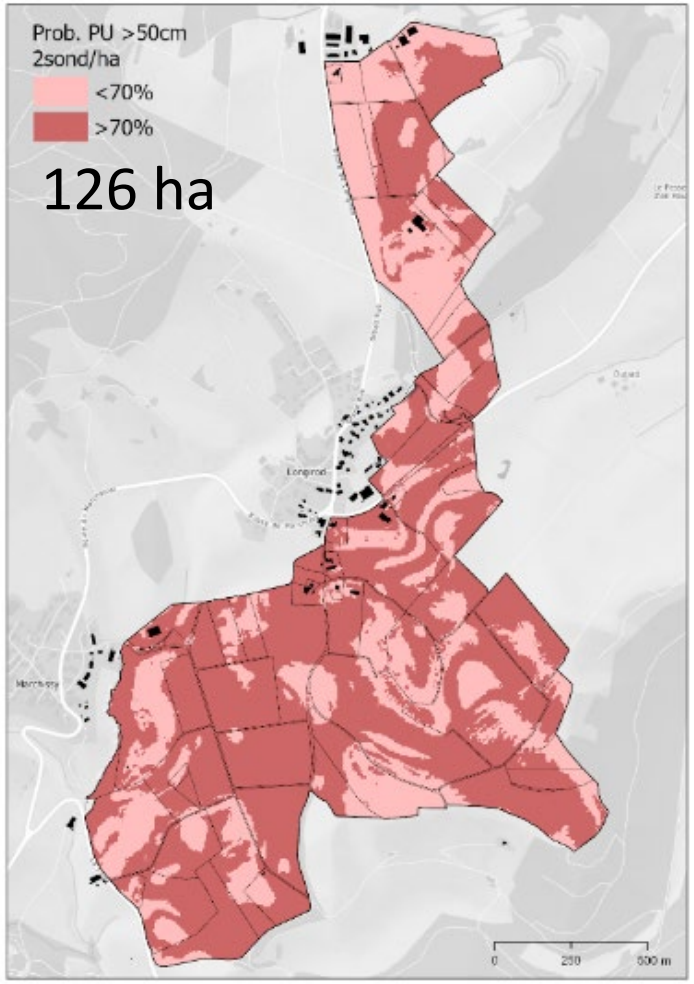


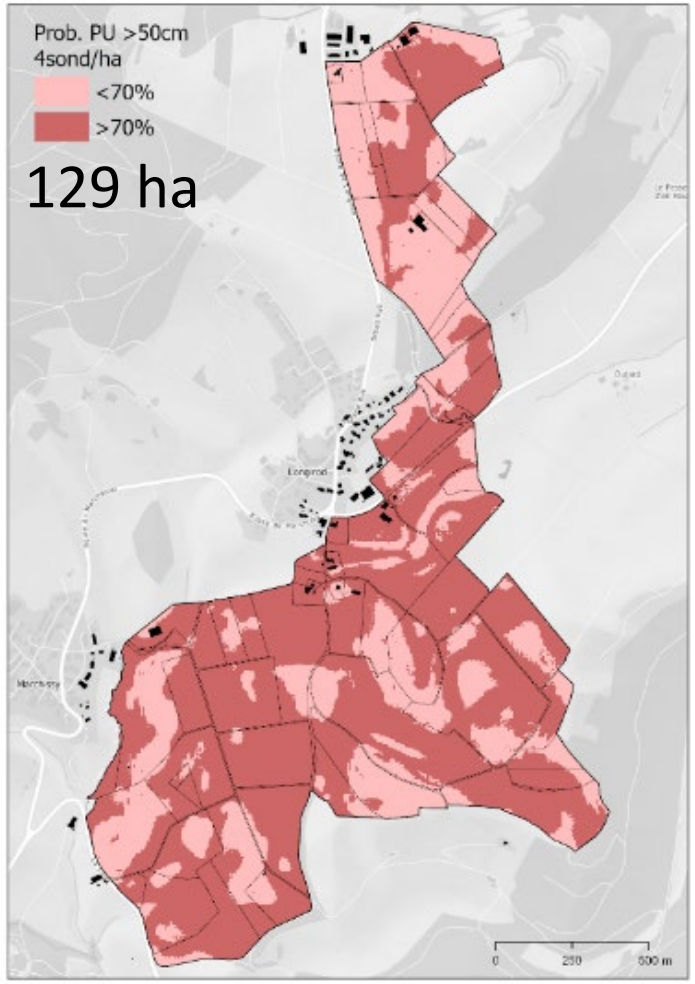
Probabilité PU > 50cm – 70 %

La densité des sondages n'a pas un effet dominant sur le total de la surface potentielle pour une même incertitude

Die Dichte der Bohrungen hat keinen dominierenden Effekt auf die potenzielle Gesamtfläche für gleiche Unsicherheit







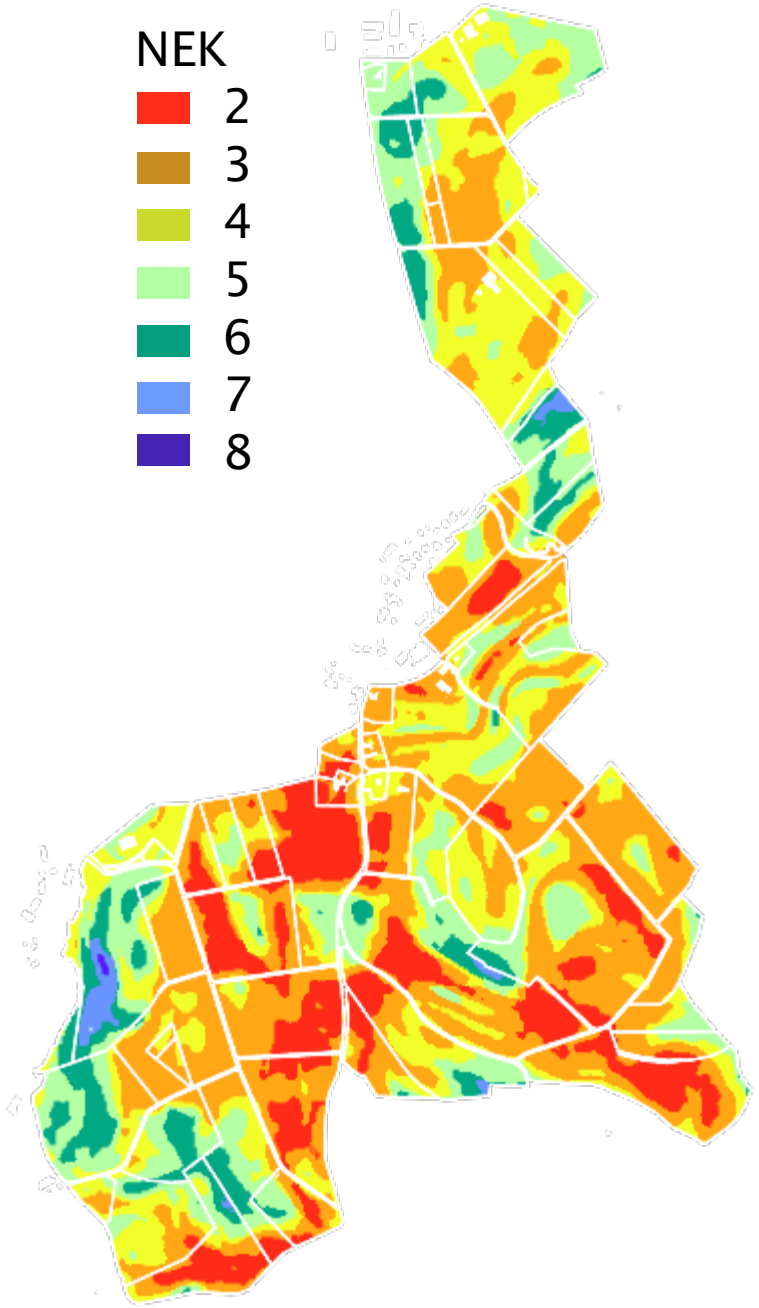
Objectif II : Identifier de SDA

Enjeux : Comment traiter cette nouvelle information?

- une logique binaire ?
- Prise en compte de l'incertitude?
- Jouer avec les % des seuils?
- Intégrer d'autres variables (ex. NEK) ?

Herausforderungen: Wie soll diese neue Information verarbeitet werden?

- *Wie sieht eine binäre Logik aus?*
- *Berücksichtigung der Unsicherheit?*
- *Mit den %-Schwellenwerten spielen?*
- *Integration anderer Variablen (z.B. NEK)?*



Objectif III : Harmoniser les pratiques

Actions mise en place

- Journée de calibration avec des experts
- Vérification et corrections par des experts
- Contrôle qualité «automatique»
- Coordination et homogénéisation des données collectées

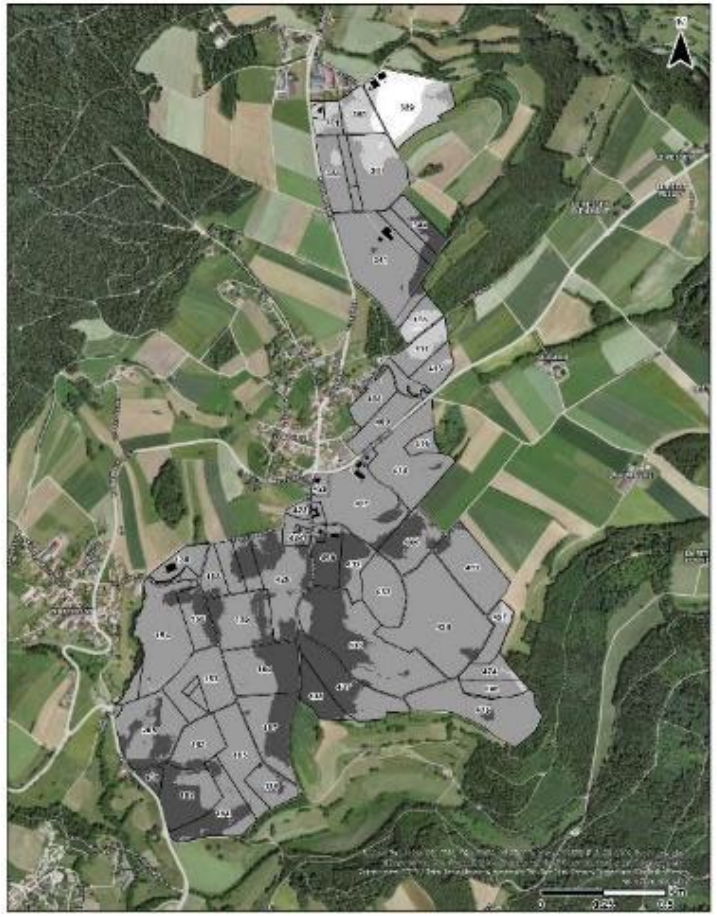
Eingesetzte Aktionen

- *Eichtag-Tag mit Experten*
- *Überprüfung und Korrekturen durch Experten*
- *«Automatische» Qualitätskontrolle*
- *Koordination und Homogenisierung der gesammelten Daten*

Objectif IV : Valoriser les données pour les agriculteurs

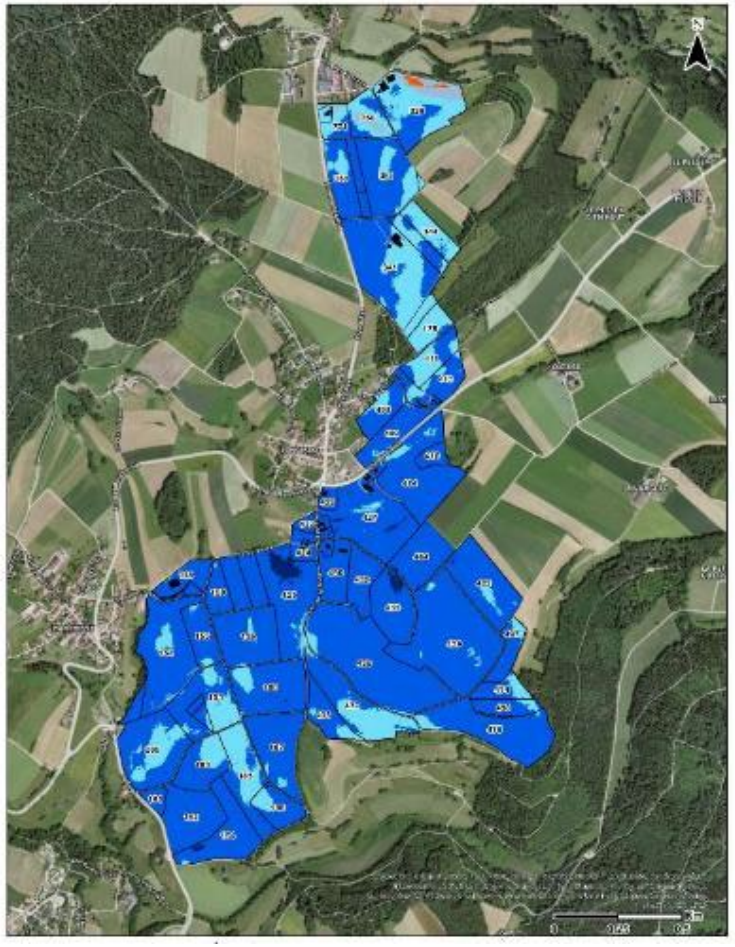
La carte avec 4 sondages / ha a été considérée comme réaliste

Die Karte mit 4 Bohrungen / ha wurde als realistisch angesehen



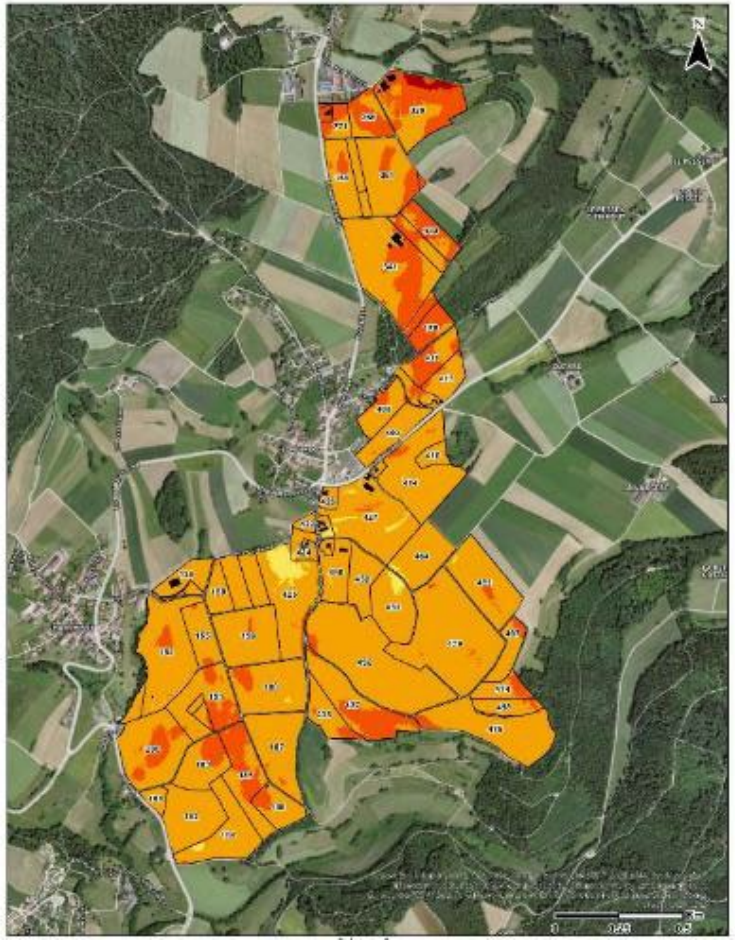
Teneur en argiles sur les 0 - 30 premiers cm (en %)

< 12 | 12 - 15 | 15 - 20 | > 20



Réserve utile en eau du sol selon le calcul GEPPA (en mm d'eau par cm de sol)

< 50 | 50 - 70 | 70 - 100 | 100 - 150 | > 150



Nombre de jours avant la sécheresse pour un taux d'évapotranspiration de 7 mm par jour

≤ 8 | 8 - 14 | 14 - 21 | > 21

Enseignements…

Enseignements:

Constats

- Importance de la qualité des relevés
- Rien n'est possible sans des données homogènes et cohérentes
- Les points placés par les bureaux améliorent le modèle

Feststellungen

- Bedeutung der Qualität der Erhebungen
- Ohne homogene und konsistente Daten ist nichts möglich.
- Die von den Büros gesetzten Punkte verbessern das Modell

Enseignements:

Des données adaptées aux besoins avec une incertitude

- Une densité de 4 sondages/ha n'est pas nécessaire: entre 1.7 et 2.5 suffisent pour identifier des SDA
- Le calcul de l'incertitude est un outil pour procéder par étape: densification progressive et ciblée
- Une connaissance de la zone permet valider les résultats et facilite le choix des covariables

Bedarfsgerechte Daten mit Unsicherheit

- *Eine Dichte von 4 Bohrungen/ha ist nicht notwendig: zwischen 1.7 und 2.5 reichen aus, um FFF zu identifizieren.*
- *Die Unsicherheitsrechnung ist ein Instrument für schrittweises Vorgehen: schrittweise und gezielte Verdichtung*
- *Die Kenntnis des Gebiets ermöglicht die Validierung der Ergebnisse und erleichtert die Auswahl der Kovariablen.*

Merci pour votre attention !