High-resolution spatial modelling of total soil depth for France

Marine Lacoste, Vera Leatitia Mulder, Nicolas N. Saby, Dominique D. Arrouays

High-resolution spatial modelling of total soil depth for France

Marine Lacoste – INRA UR 0272 Science du Sol, Orléans, France
Titia Mulder – INRA US Infosol, Orléans, France
Contributors: M. Martin, N. Saby, A. Richer de Forges & D. Arrouays
Project: GlobalSoilMap

6th Global Workshop on Digital Soil Mapping
11-14 November, 2014, Nanjing, China

INTRODUCTION

- Soil depth (SDt):
  - Key soil property for water availability and carbon stocks
  - Exhaustive mapping of total soil depth = requirement of the GlobalSoilMap project

- Difficulties of SDt mapping due to:
  - Soil properties: high spatial variability
  - Soil observation tools: estimation of soil depth for deep soils (> 1.5 m)
  - Discordance about SDt definition

- Evaluate two different modelling approaches to produce a high-resolution soil depth map of France
  - In a regional or global context + high resolution
    - Large data sets
    - Spatial heterogeneity
    - Local, large and nested-scale processes
  - Robust and reproducible
  - Spatial explicit uncertainties

RESEARCH OVERVIEW

Input data

- **Soil sample data** (source: French Soil Monitoring network)
- **Exhaustive covariates** capturing biotic and abiotic conditions
  - Soil type and properties
  - Parent material
  - Relief (SRTM-DEM)
  - Climate
  - Land use

Analysis

1) Data mining

+ Bias correction
+ Ordinary kriging of the residuals

Resolution: 90 m

R packages: *caret*, *gbm*, *qmap*, *gstat*

2) Multi-resolution kriging for large datasets

Fixed trend model + kriging

Resolution: 500 m

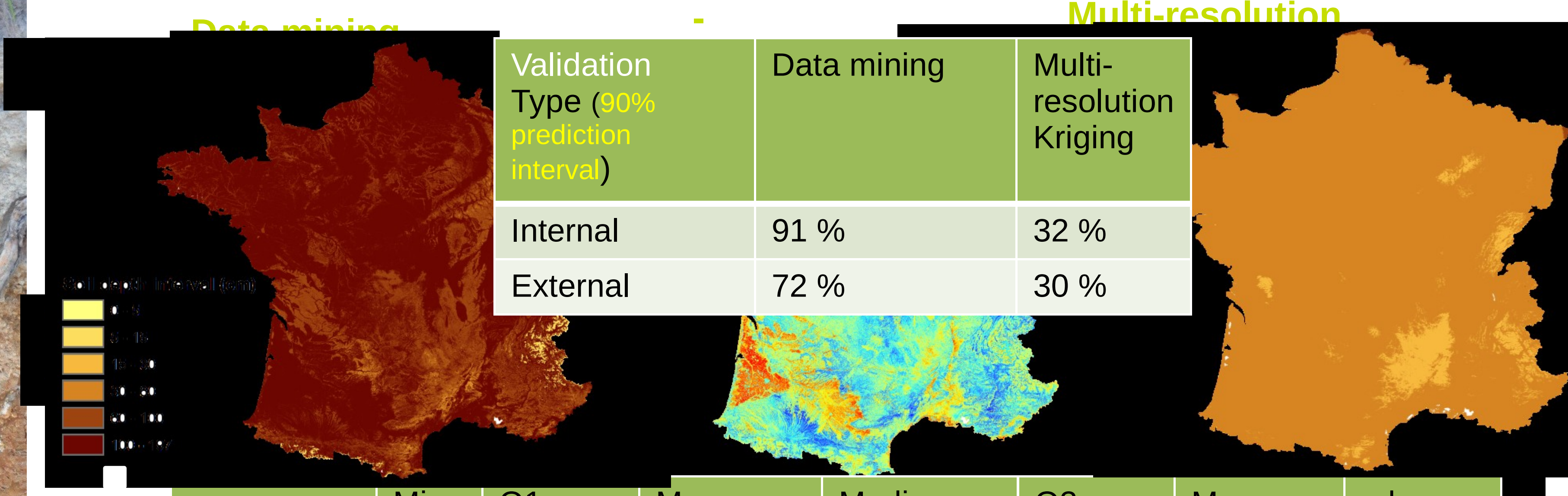
R packages: *LatticeKrig*

Evaluation criteria

1) Map accuracy
   - Internal validation
   - Cross-validation
   - External validation: concordance with previous soil map

2) Prediction and confidence intervals by conditional simulation of kriging model

RESULTS (I/V)



Validation Type (90% prediction interval)	Data mining	Multi-resolution Kriging
Internal	91 %	32 %
External	72 %	30 %

	Min	Q1	Mean	Median	Q3	Max	sd
Data Mining	0	97	111	113	127		197
MR Kriging	13	35	38	38	42	72	6
Difference	-154	-18	4	1	24	186	33

As discussed: the validation is incorrect for MR Kriging

Discussion

Data mining | **Multi-resolution Kriging**

Predictive map of soil depth

Consistent spatial pattern

Good prediction of the mean values

Data mining

- Prediction of extremes values

Multi-resolution Kriging

- Ongoing: increasing the resolution/levels

90% Confidence int

Data mining

- Large (high uncertainties)
- "Consistent" with observed values

Multi-resolution Kriging

- Narrow (low uncertainties)
- Ongoing: test lower confidence intervals

Implementatior

Data mining

- Multisteps/multitools approach
- No direct estimation of uncertainties
- Flexible for large datasets, high resolution

Multi-resolution Kriging

- Straight forward modelling approach
- Flexible in delivering spatial explicit uncertainty measures

Outlook

Data mining

- Promising prediction of soil depth class instead

Multi-resolution Kriging

- Potential for modelling beyond the country level, at high resolution as demonstrated in other global environmental models

THANK YOU ALL!

Essentially, all life depends upon the soil.
There can be no life without soil and no soil without life;
they have evolved together.

American naturalist Charles Kellogg, 1938.

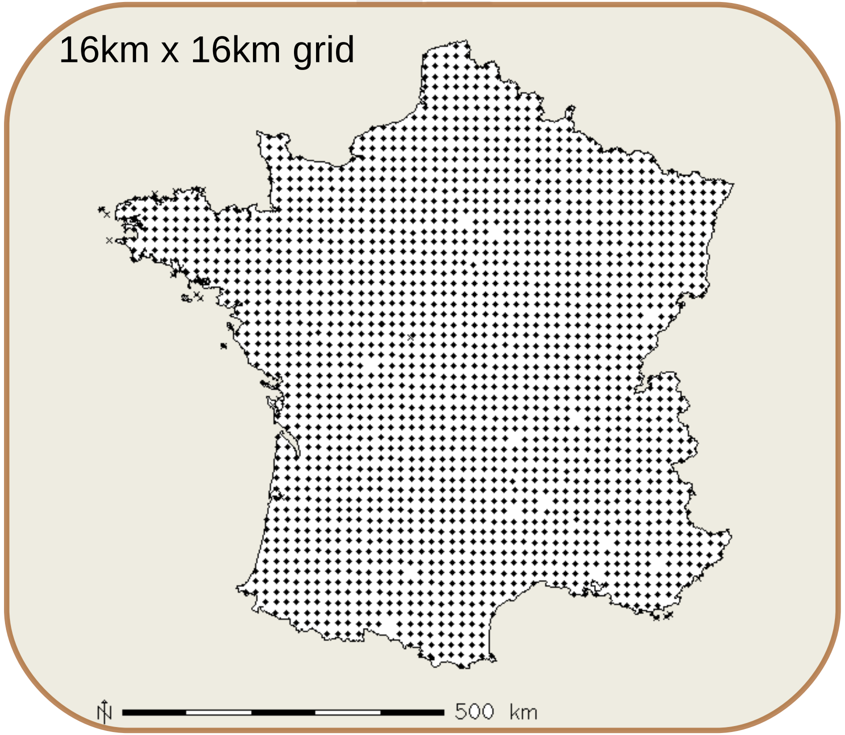
FINANCIAL SUPPORT:

AgreenSkills
Pathways for inventive researchers

Inventaire, gestion
et conservation
des sols
Sol

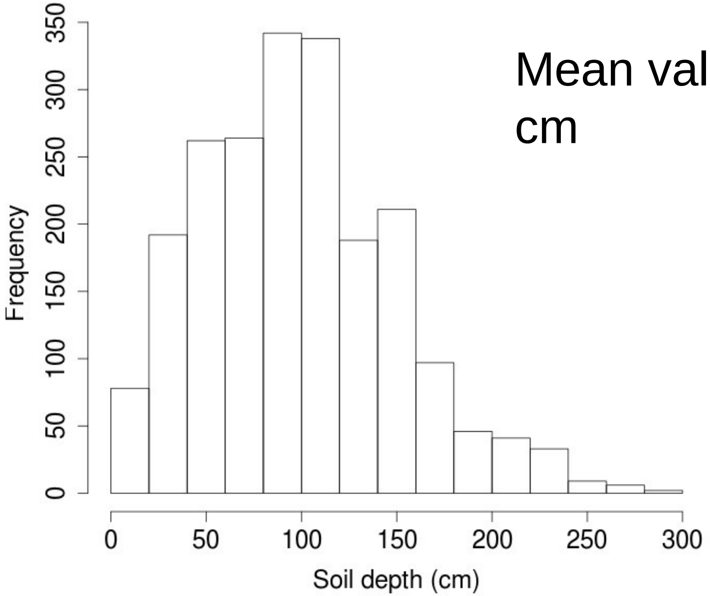
Région Centre

STUDY AREA : France (~ 540K km2)



SDt determined for 2116 sites

- French Soil Monitoring network (RMQS)



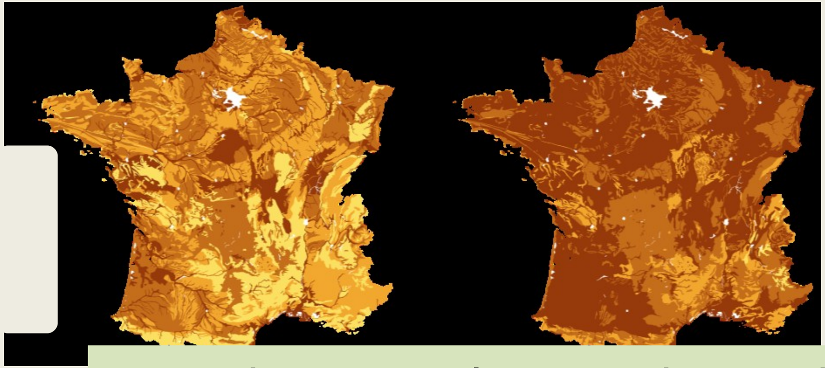
Mean value: 102 cm

STUDY AREA : France (~ 540K km2)

Existing soil depth maps

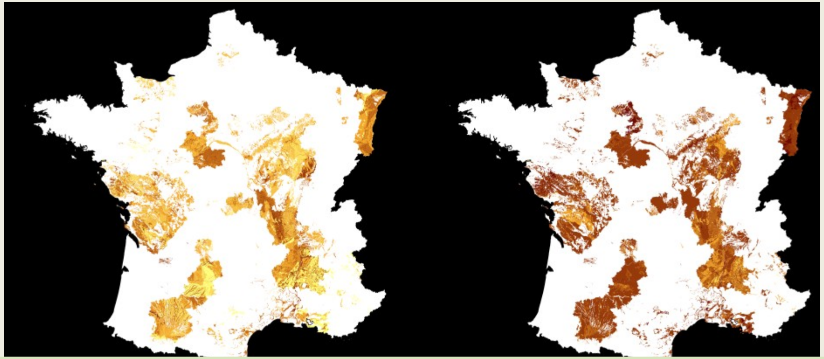
Scale: 1/1 000 000

Lower limit | Upper limit



Scale: 1/250 000

Lower limit | Upper limit



Note: these are classes. The spatial distribution of these classes is the same
As the soil type classes used for the data mining model: this introduces bias
In the following validation results – I have my questions about that approach…

METHODS

Continue soil depth prediction

Data mining

Multi-resolution Kriging

- Estimation of covariance matrix using multi-resolution radial basis functions
- Covariance model can approximate the Matern covariance family
- Developed for handling large datasets
- *R package LatticeKrig*
- Resolution: 500m – for me it doesn't make sense to go to 90m because it is not supported by the data we use….also, the model cannot be calibrated because there is no variability below this level
- Fixed linear trend model: elevation, slope, precipitation, gravimetry, bed rock resistance and NPP
- **Kriging error obtained by conditional Gaussian simulation (1000 times) – this is really a pro!**

RESULTS

Importance of the covariates

Data mining

Variable	Importance (%)
SRTM (elevation)	14
Maximal annual temperature (mean)	9
Parent material	8
Aspect	7
Mean annual precipitation	7
Climate type	7
Roughness	7
Land use for forest areas	6
Wetness index	6
Soil type	6
Drainage network	6
Slope position	6
Slope	6
Bare rock areas	5

Multi-resolution Kriging

- Fixed linear trend model: elevation, slope, precipitation, gravimetry, bed rock resistance and NPP – what are the coefficients?

RESULTS

Models accuracy

Interesting to see the multi-resolution kriging improves with a higher resolution soil class map. The good validation results for data mining relate to the previous mentioned bias. The classes have been very important for the data mining – this data does not have the spatial variability compared to eg SRTM. Matching the soil depth class with modelled soil depth thus shows high agreement + The variogram of the residuals did not show high spatial variability. Concluding – a soil class map is not the best type of validation here. BTW the internal validation of the MR kriging is cross validation – so not too bad compared to the cross-validation of the data mining technique.

The histograms should be changed to relative frequency due to the different resolution – or make 2 separate histograms (difference in resolution = different total). The kriging, as expected, shows a smoothing of values (no extremes). What about the validation with the independent IGCS soil depth data? Still impossible because of the inaccuracy of that dataset? Maybe Anne knows how to select the most accurate samples – perhaps a specific year, institute or sampling programm which was consistent over the years?