Food and Agriculture Organization of the United Nations

GLOSOLAN
Soil spectroscopy training workshops

Characterization of soil properties using the French national spectral libraries (Vis-NIR, NIR and MIR)

GOMEZ Cécile
cecile.gomez@ird.fr
French National Research Institute for Sustainable Development
Montpellier (France) / Bangalore (India)

Online webinars

GLOBAL SOIL PARTNERSHIP

With the support of

INRAE
la science pour la vie, l'humain, la terre

Dominique Arrouays

INRAE
la science pour la vie, l'humain, la terre

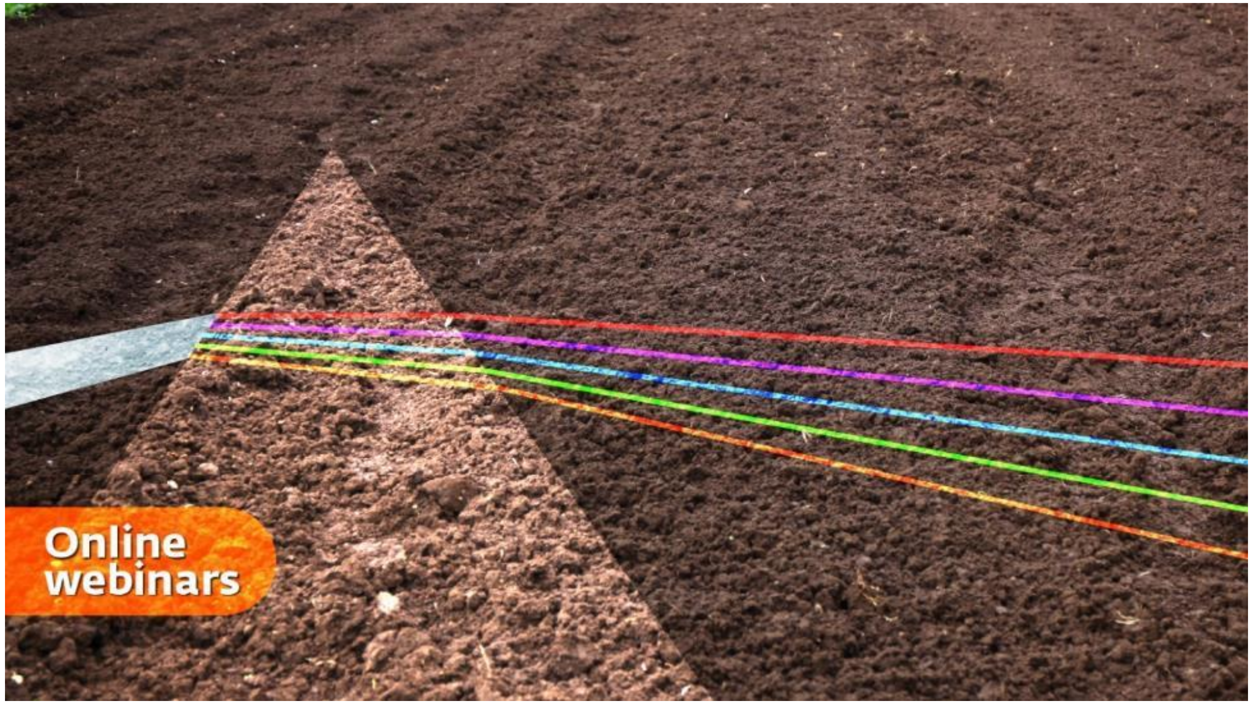
Claudy Jolivet

IRD
Institut de Recherche pour le Développement
FRANCE

Bernard Barthès

IRD
Institut de Recherche pour le Développement
FRANCE

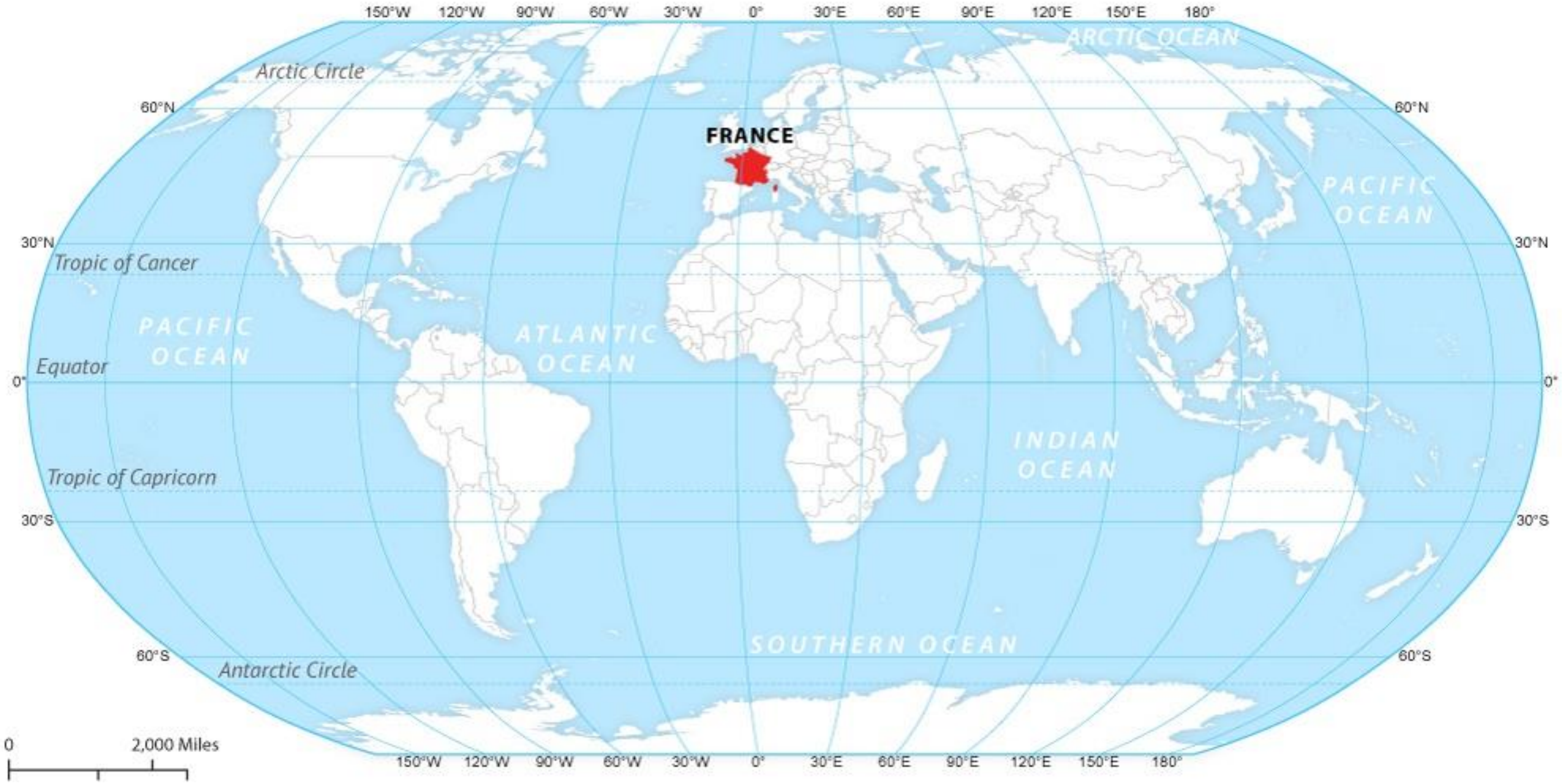
Patricia Moulin



CONTENTS

- The soil quality monitoring network for French soils (RMQS)

- Can we use the French spectral library for soil properties estimation
  - At National Scale (in France)?
  - At Regional Scale (in France)?
  - At Regional Scale (Outside of France)?



The soil quality monitoring network for French soils (RMQS)

A program funded and coordinated by the group of scientific interest SOL (Gis Sol)

MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE
Liberté
Égalité
Fraternité

MINISTÈRE
DE L'AGRICULTURE
ET DE L'ALIMENTATION
Liberté
Égalité
Fraternité

ADEME
AGENCE DE LA
TRANSITION
ÉCOLOGIQUE

OFB
OFFICE FRANÇAIS
DE LA BIODIVERSITÉ

INRAE

IRD

IGN

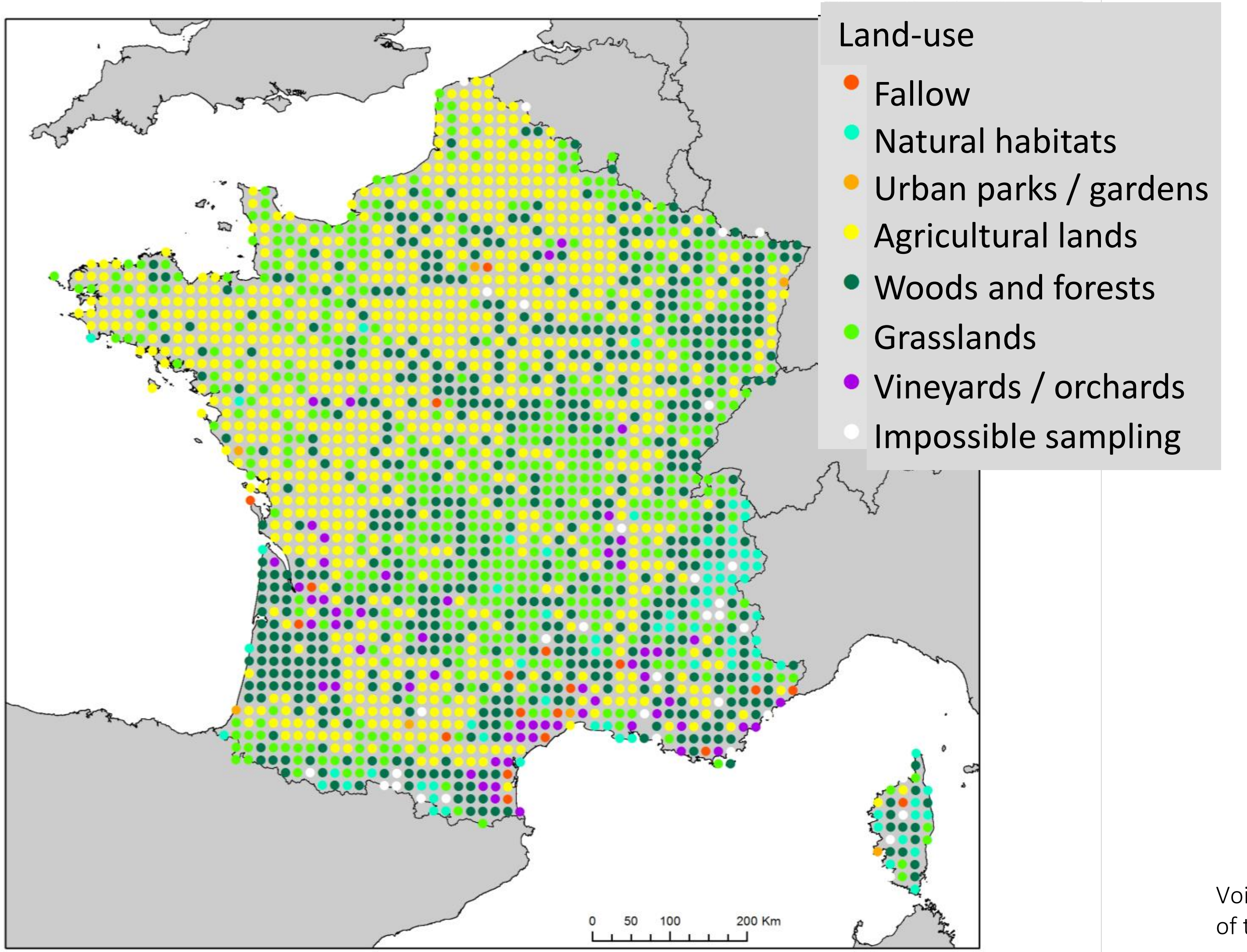
Géosciences pour une Terre durable
brgm

Toward a long-term monitoring of soil quality

- **National statement** (global statistic report on soil parameters evolution)
- **Mapping** (to get an instant picture of soil quality and detect gradients)
- **Warning** (early detection of unsuspected evolution)
- **Archiving** (to constitute a bank of soil samples)

Arrouays D., Jolivet C., Boulonne L., Bodineau G., Saby N. et Grolleau E., 2002 - Une initiative nouvelle en France : la mise en place d'un réseau multi-institutionnel de la mesure de la qualité des sols (RMQS). Comptes rendus de l'Académie d'Agriculture de Paris, 88, n° 5, pp. 93-103

A systematic network representative of French soils and land-uses



- 2200 sites
- located along a 16 x 16 km grid
- on continental France and overseas territories
- representative of French soils and land-uses
- resampled with 15 years interval

Voisin E., Boulonne L., Jolivet C., Ratié C., Arrouays D. – 2012. Analysis of the representativeness of land use in France by the French soil monitoring network, Poster, Eurosoil congress, Bari, Italy

A new sampling campaign each 15 years

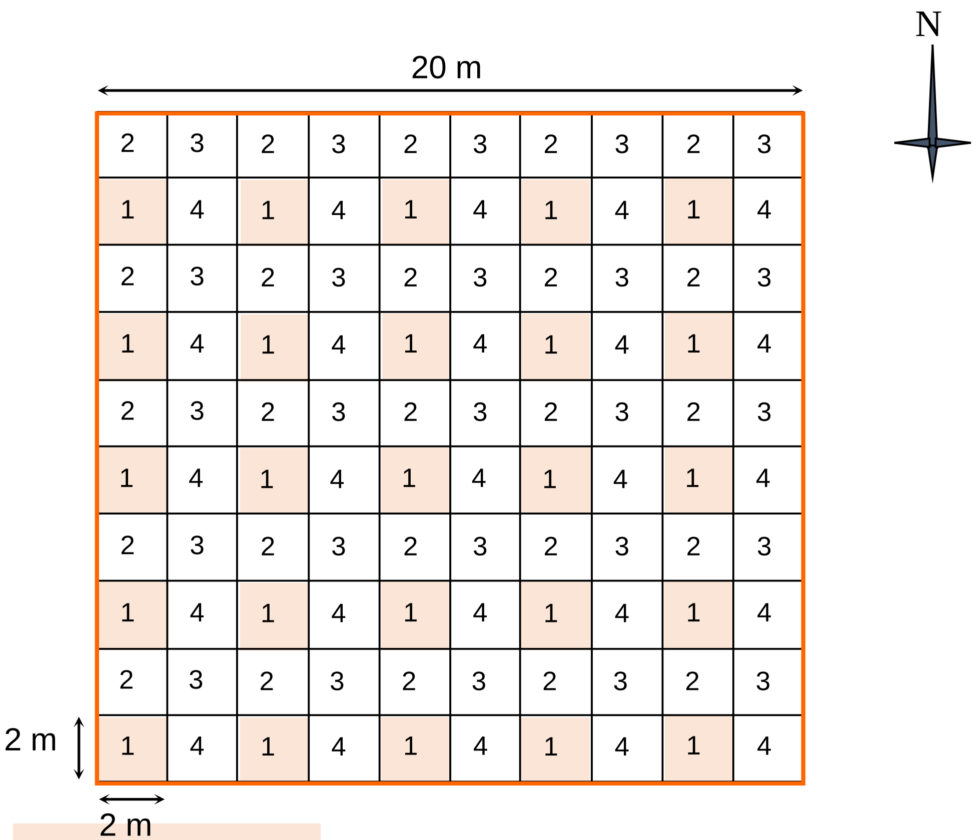
RMQS1 : 2000-2015

- Continental France : 2000-2009
- Overseas territories : 2006-2015
  - 2006 Guadeloupe
  - 2007 Martinique
  - 2012 Réunion & Mayotte
  - 2014-2015 French Guyana (coast)

RMQS2 : 2016-2030

…

A sampling design dedicated to soil monitoring



2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4

1st campaign

A sampling design dedicated to soil monitoring

20 m

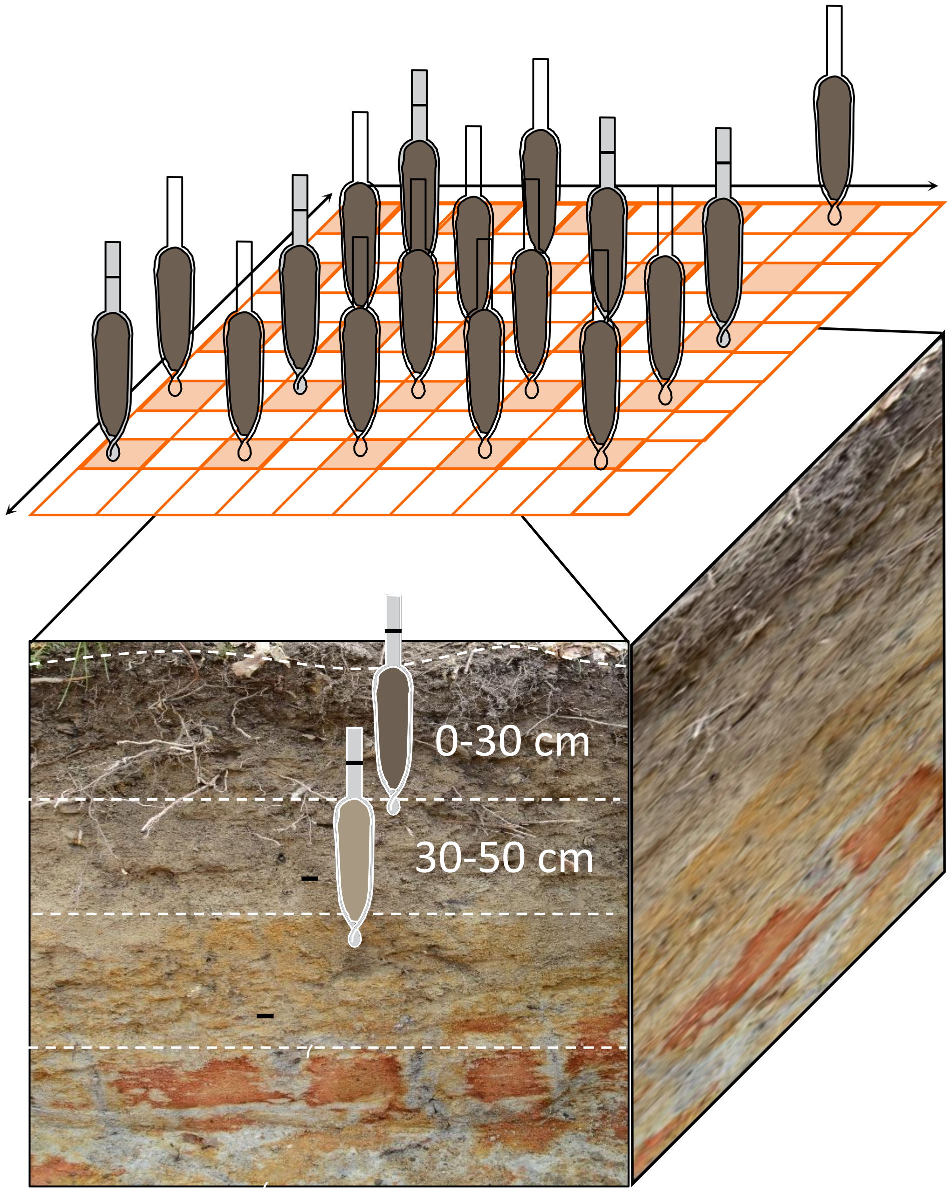
N

2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4

2 m

2 m

1st campaign



A sampling design dedicated to soil monitoring

2nd campaign

20 m

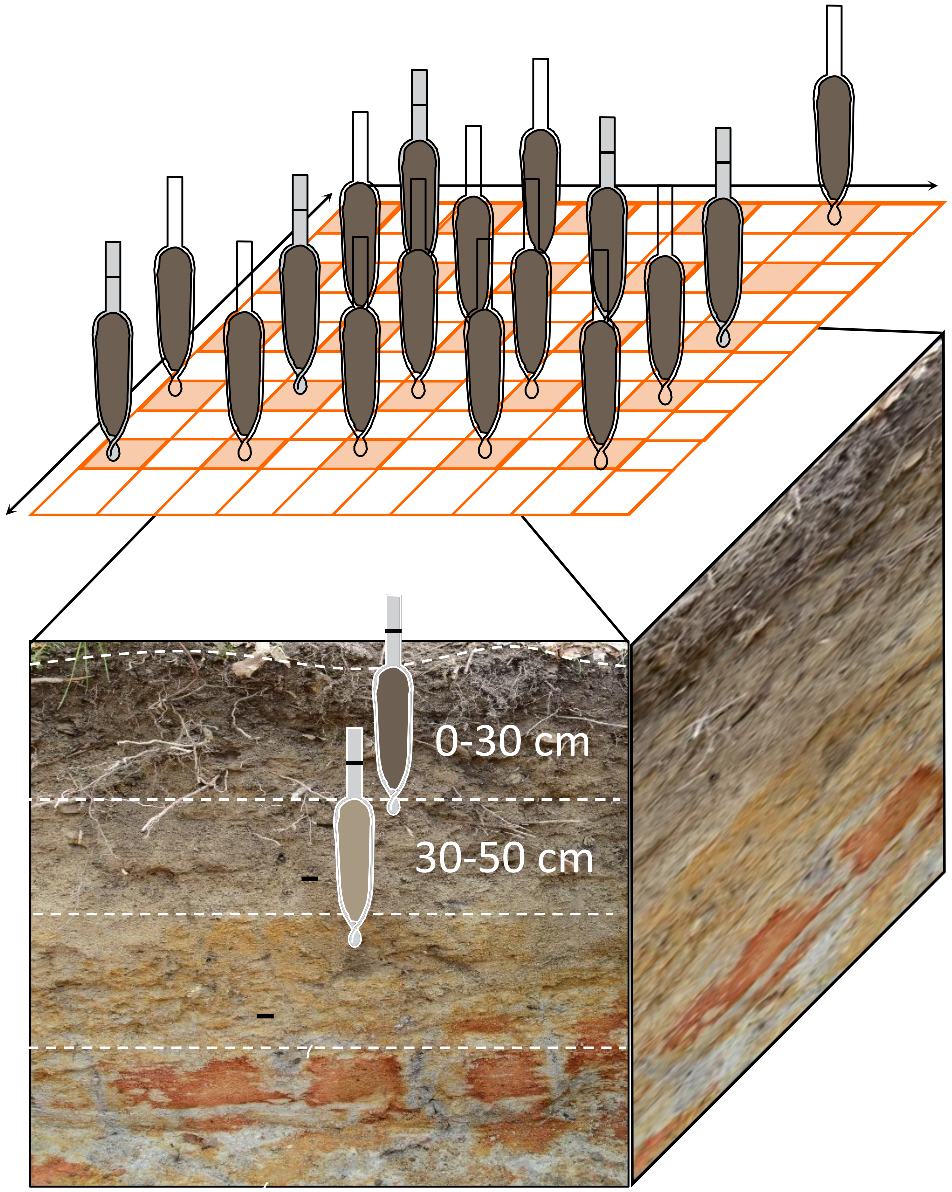
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4

2 m

2 m

1st campaign

N



A sampling design dedicated to soil monitoring

2nd campaign

3rd campaign

20 m

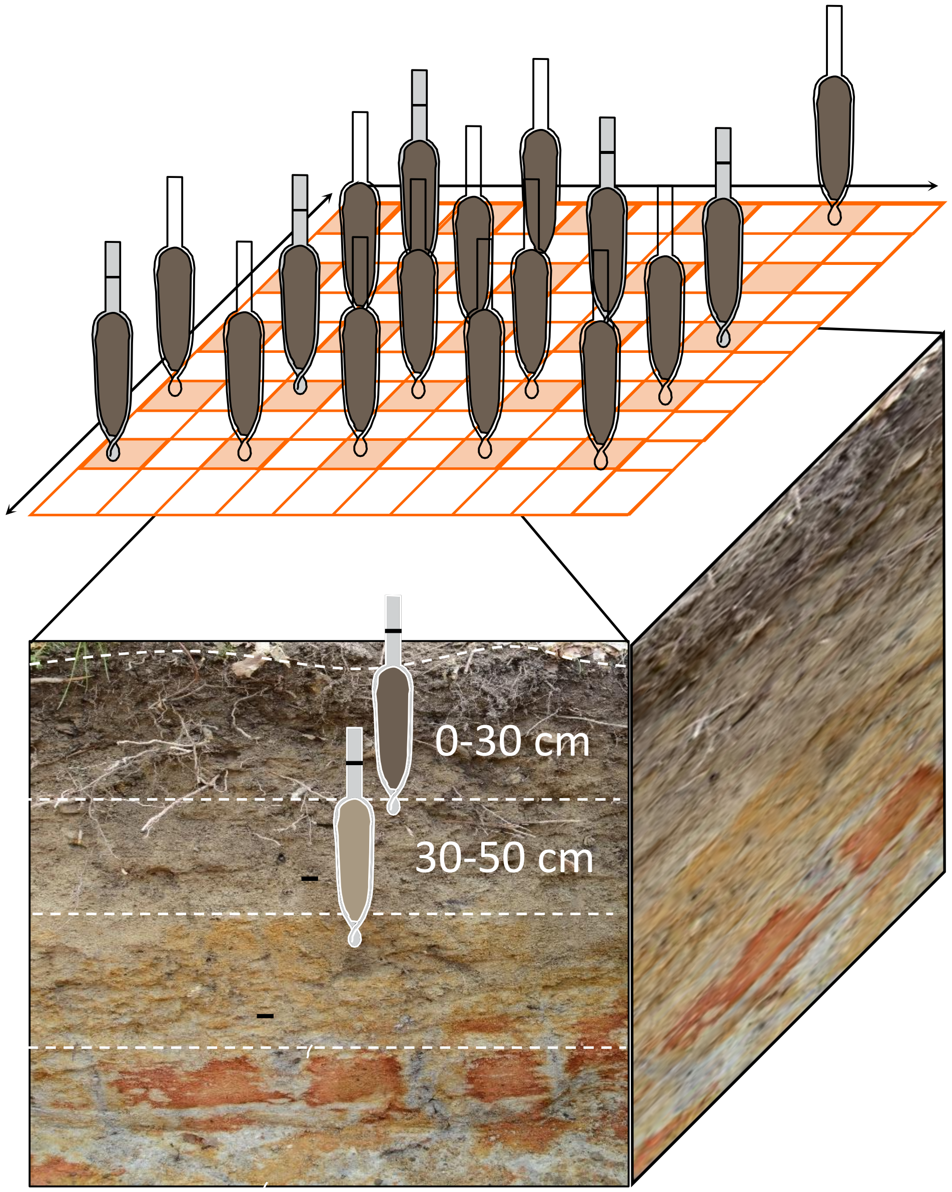
N

2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4
2	3	2	3	2	3	2	3	2	3
1	4	1	4	1	4	1	4	1	4

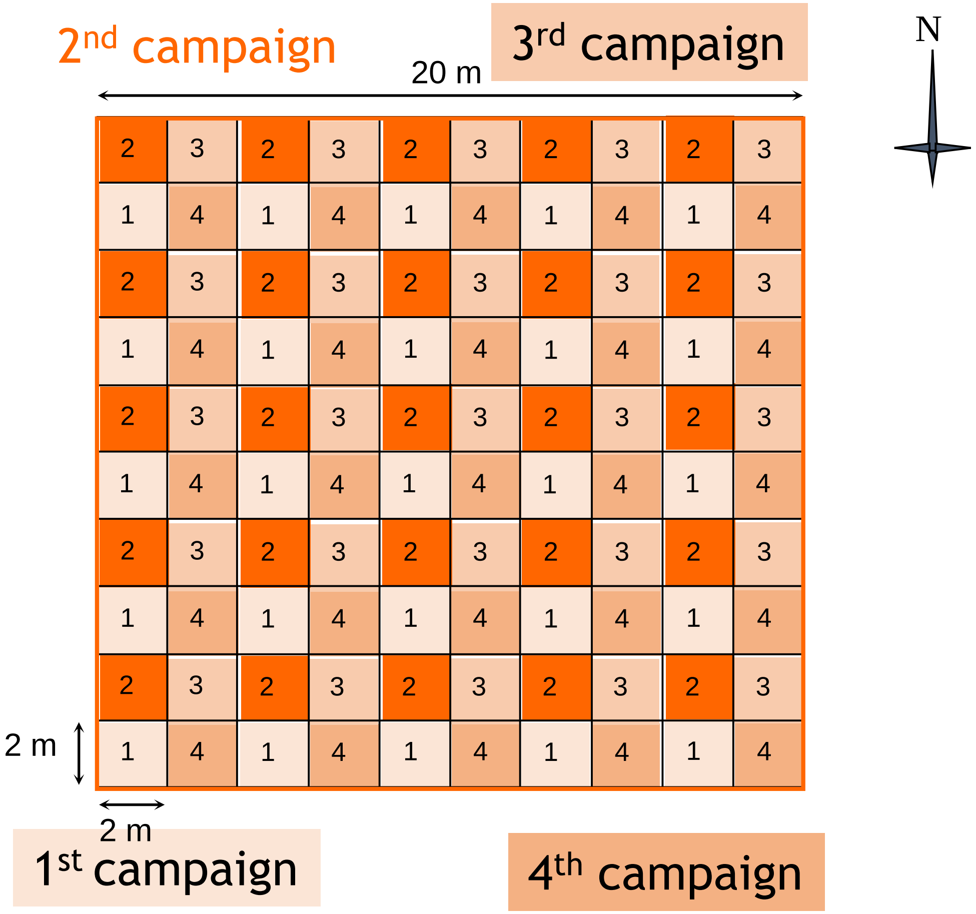
2 m

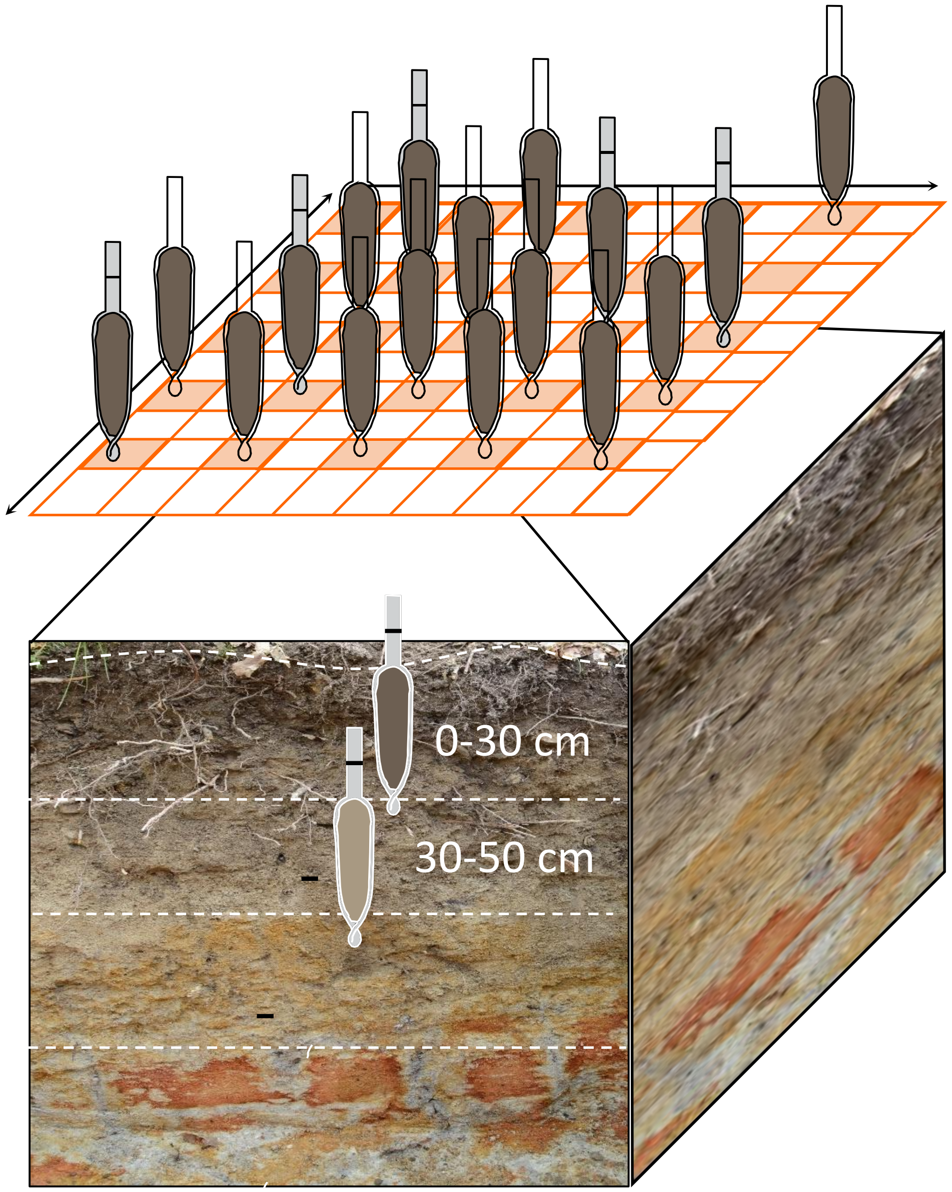
2 m

1st campaign

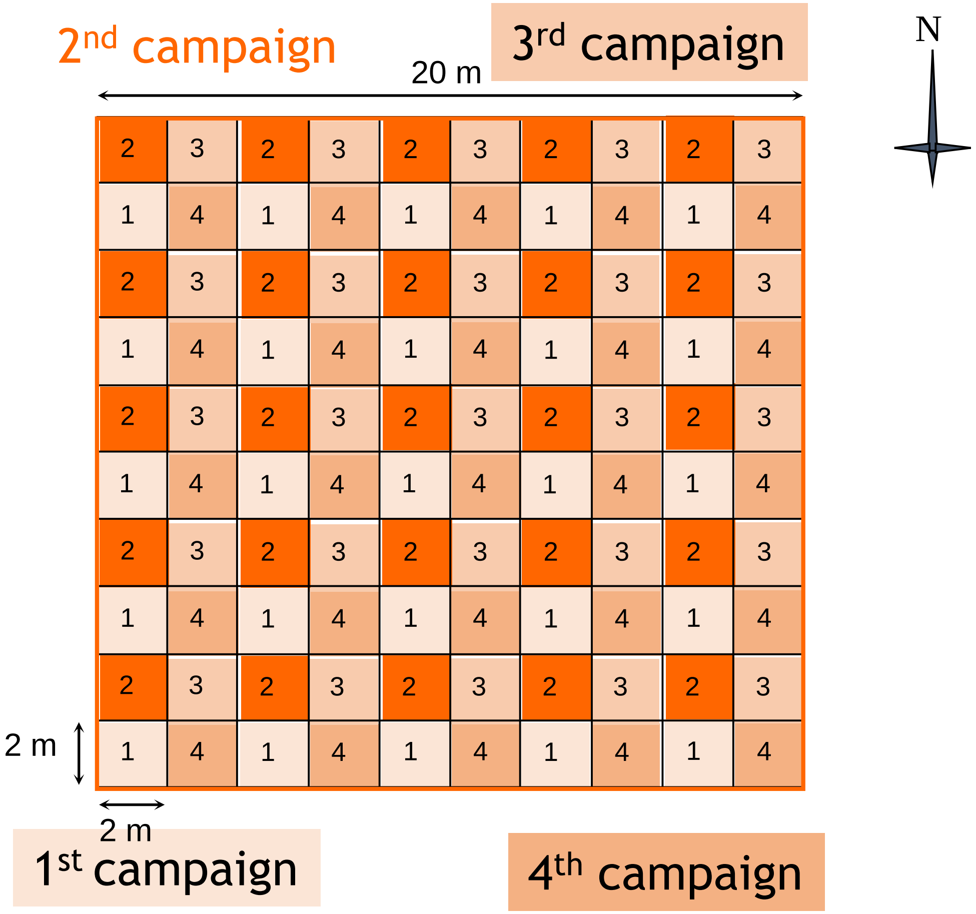


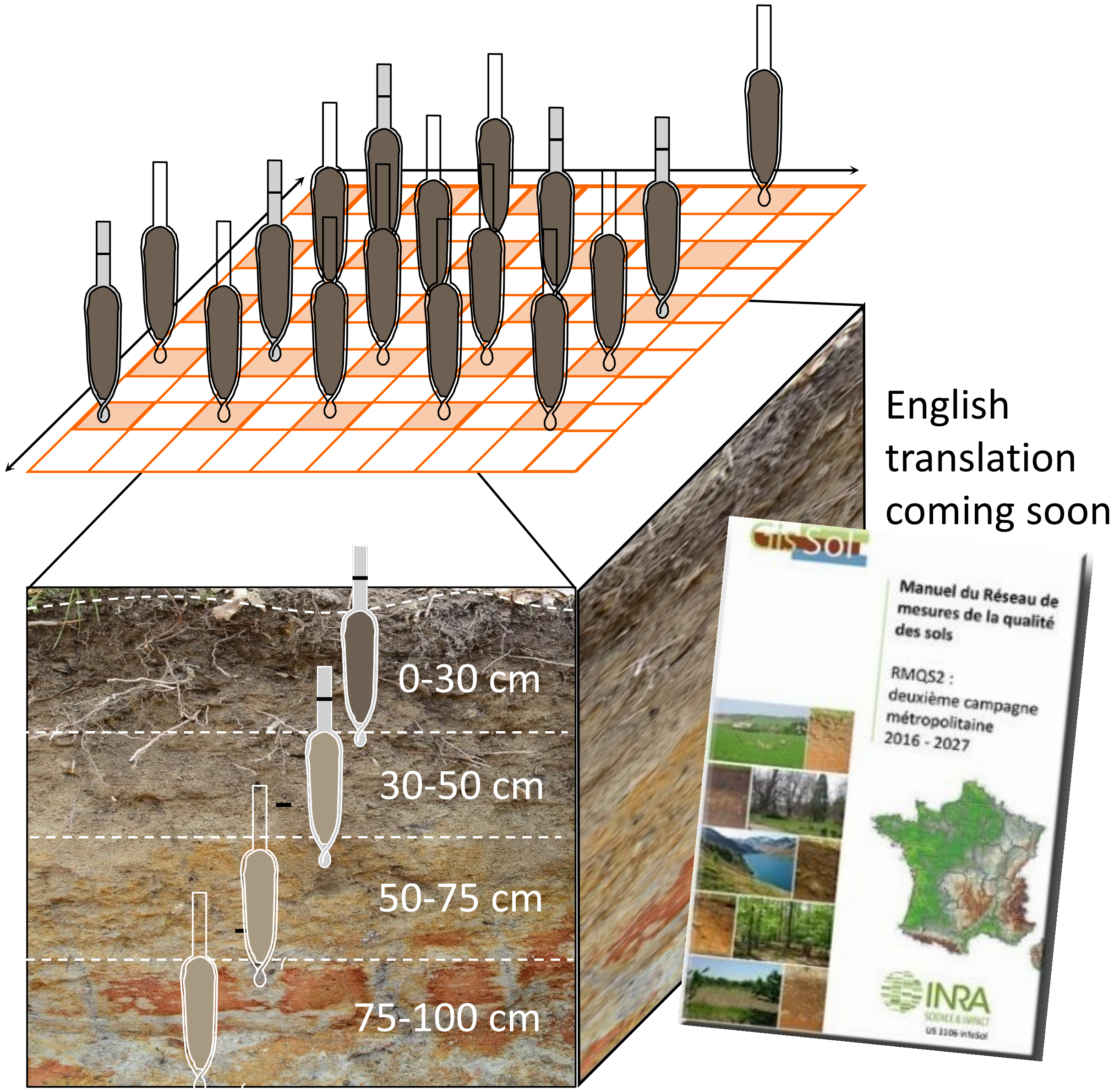
A sampling design dedicated to soil monitoring





A sampling design dedicated to soil monitoring





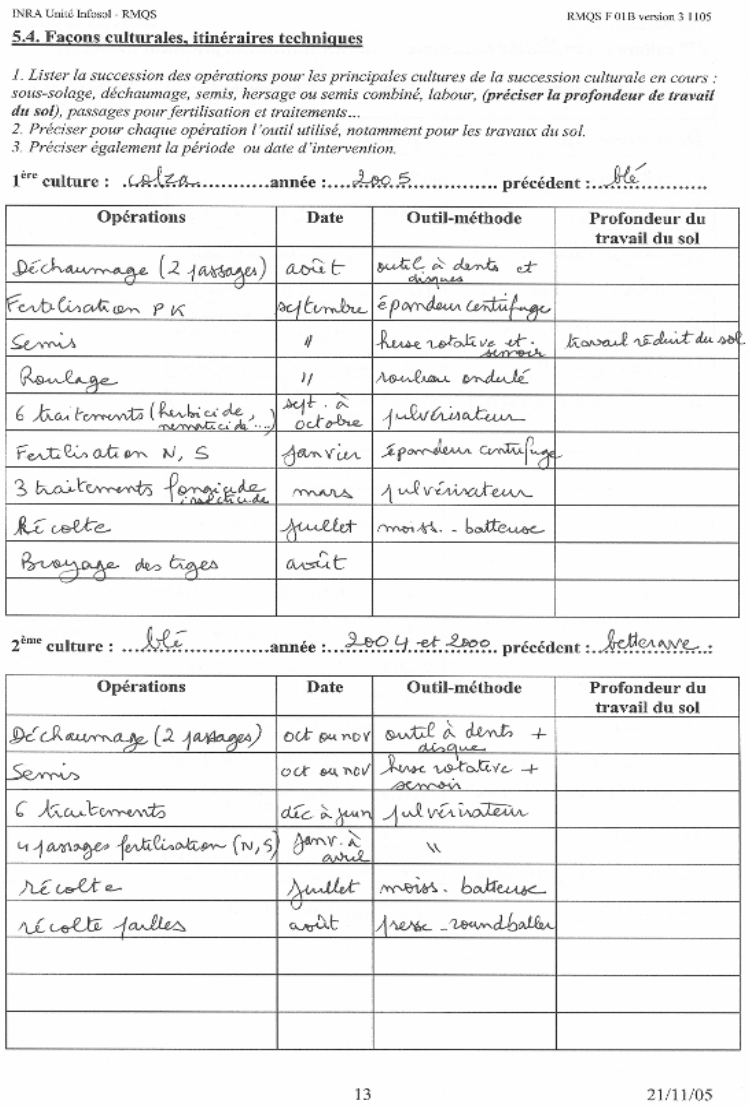
English translation coming soon !

https://www.gissol.fr/le-gis/programmes/rmqs-34

Data collection on the environment and practices

Environment & contamination sources

History & management practices = enquiries



INRA Unité Infosol - RMQS — RMQS F 01B version 3 1105

5.4. Façons culturales, itinéraires techniques

*1. Lister la succession des opérations pour les principales cultures de la succession culturale en cours : sous-solage, déchaumage, semis, hersage ou semis combiné, labour, **(préciser la profondeur de travail du sol)**, passages pour fertilisation et traitements…*
2. Préciser pour chaque opération l'outil utilisé, notamment pour les travaux du sol.
3. Préciser également la période ou date d'intervention.

1ère culture : colza **année :** 2005 **précédent :** blé

Opérations	Date	Outil-méthode	Profondeur du travail du sol
Déchaumage (2 passages)	août	outil à dents et disques	
Fertilisation PK	septembre	épandeur centrifuge	
Semis	"	herse rotative et semoir	travail réduit du sol
Roulage	"	rouleau ondulé	
6 traitements (herbicide, nematicide…)	sept. à octobre	pulvérisateur	
Fertilisation N, S	janvier	épandeur centrifuge	
3 traitements fongicide insecticide	mars	pulvérisateur	
Récolte	juillet	moiss. - batteuse	
Broyage des tiges	août		

2ème culture : blé **année :** 2004 et 2000 **précédent :** betterave

Opérations	Date	Outil-méthode	Profondeur du travail du sol
Déchaumage (2 passages)	oct ou nov	outil à dents + disque	
Semis	oct ou nov	herse rotative + semoir	
6 traitements	déc à juin	pulvérisateur	
4 passages fertilisation (N, S)	janv. à avril	"	
récolte	juillet	moiss. batteuse	
récolte pailles	août	presse _roundballer	

13 — 21/11/05

→ essential data rarely collected on soil monitoring networks

A complete analytical menu for soil characterisation

- **Soil attributes**
  - Particle size distribution, pH, C, N, P, CEC, exchangeable cations, major elements, etc.

- **Hydric properties**
  - Soil water retention

- **Contaminants & human health**
  - Traces elements: As, Cd, Co, Cr, Cu, Hg, Mo, Ni, Pb, Tl, Zn
  - organic pollutants: HAP, PCB, dioxines, furanes, OCP, herbicides
  - Pathogenic microorganisms

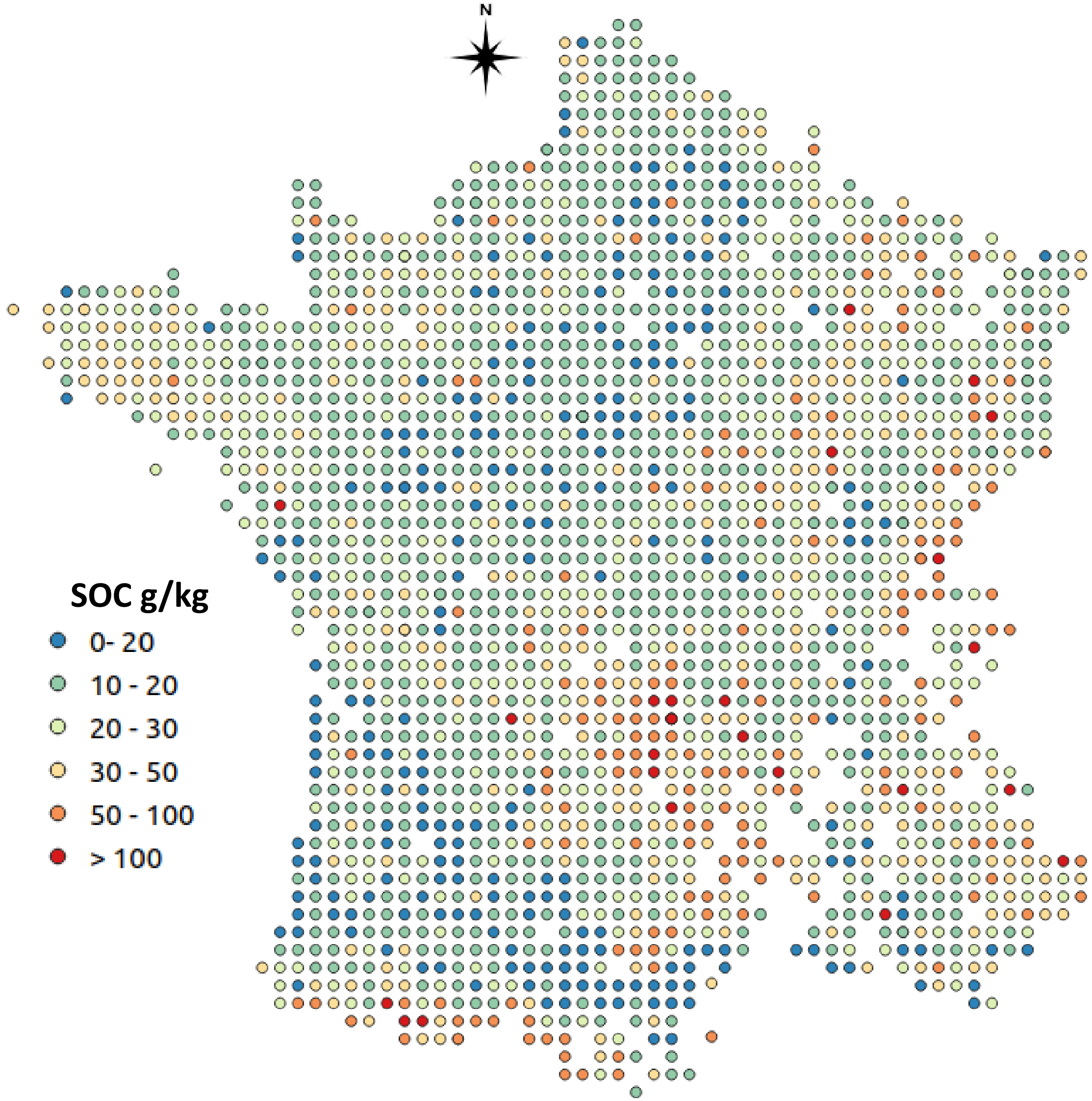
- **Carbon & climate change**
  - Deep organic carbon stocks ≤1m
  - Particulate organic matter
  - Organic matter quality: Black carbon, Glomaline

- **Soil biodiversity**
  - Microbial richness & diversity by DNA extraction (Génosol)
  - Enzymatic activity (BioChemEnv)

- **Vis-NIRS, NIRS, MIRS**

SOC

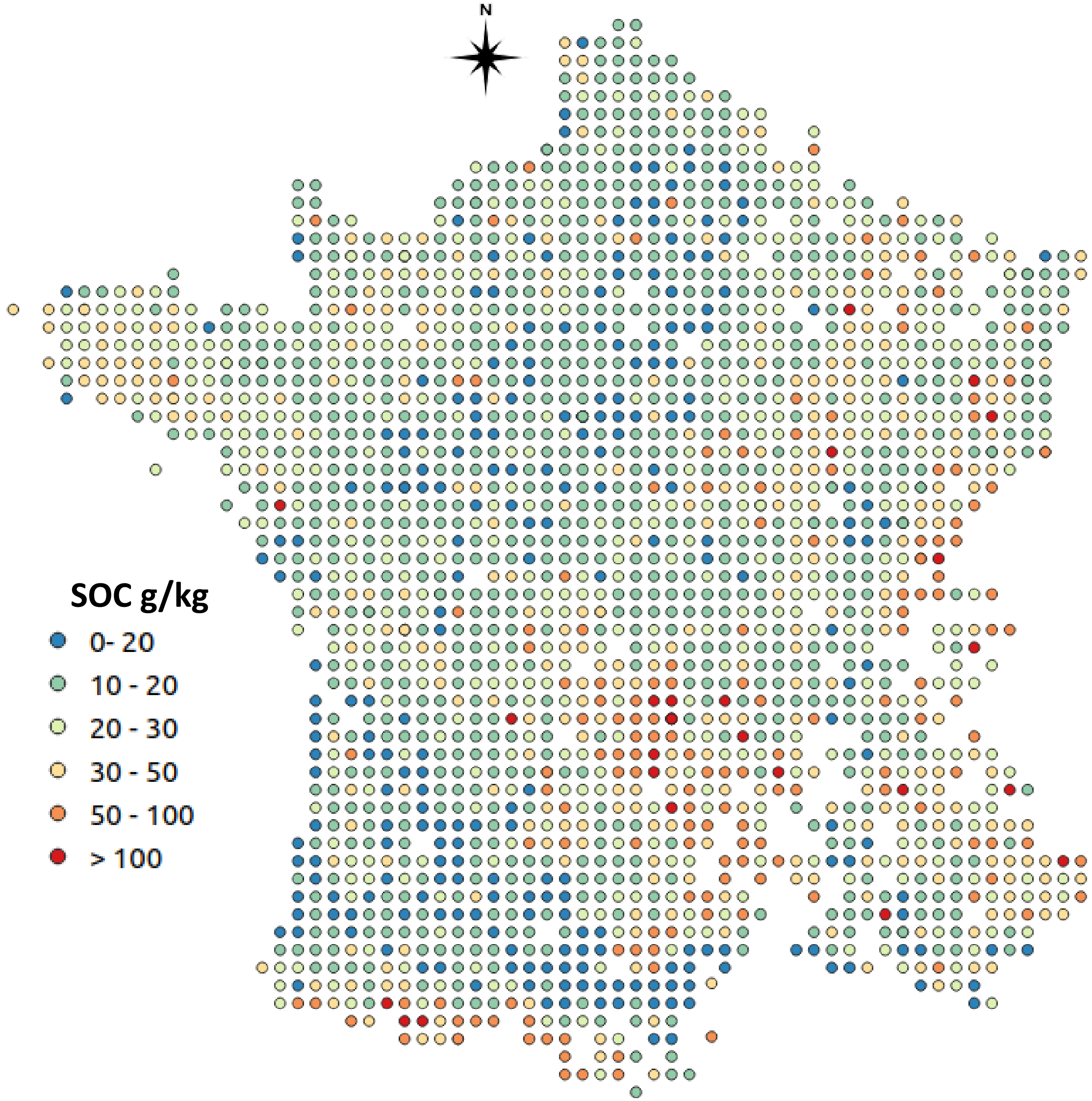
(Collected at 0-30 cm depth)



The soil quality monitoring network for French soils (RMQS)

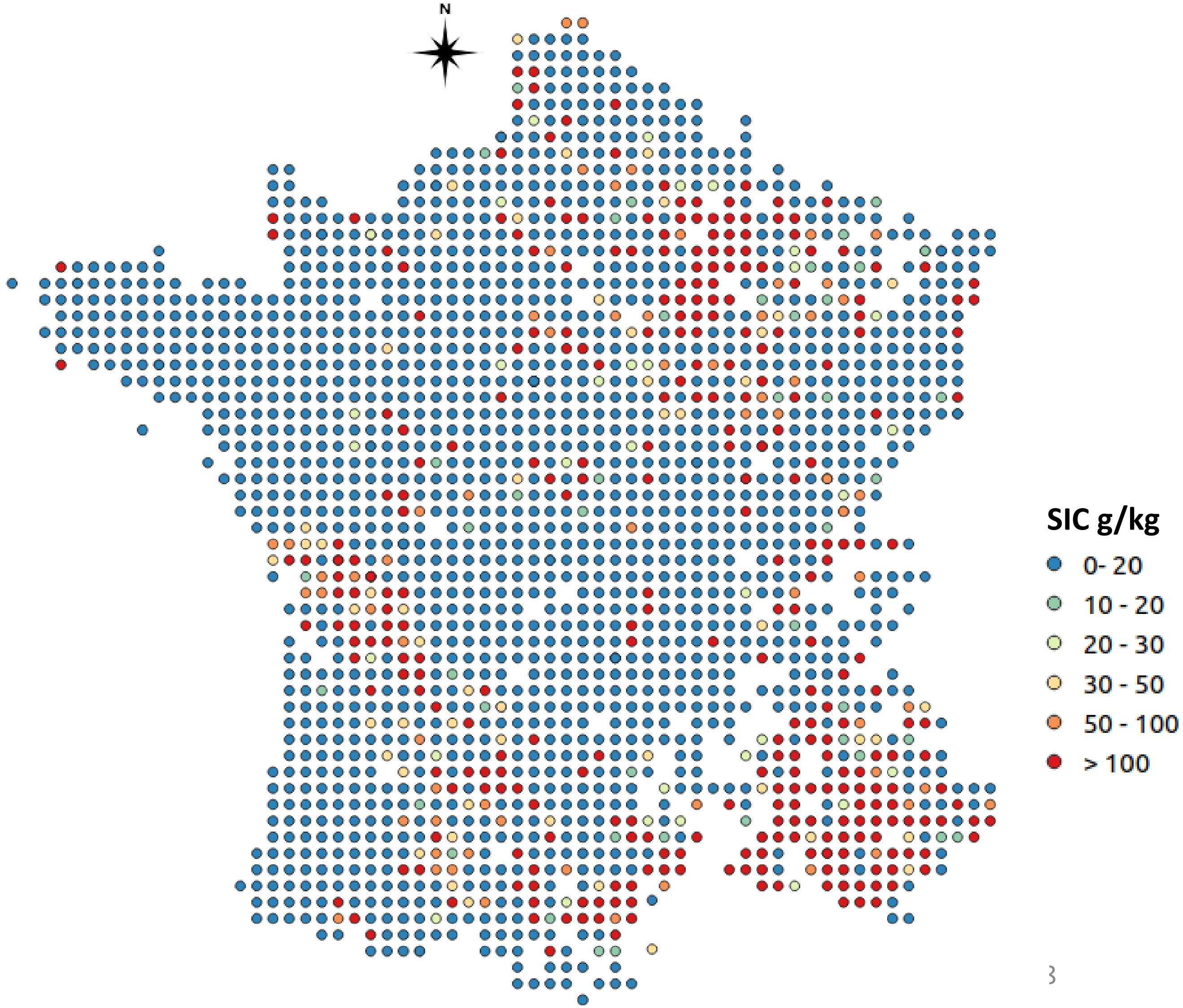
SOC

(Collected at 0-30 cm depth)



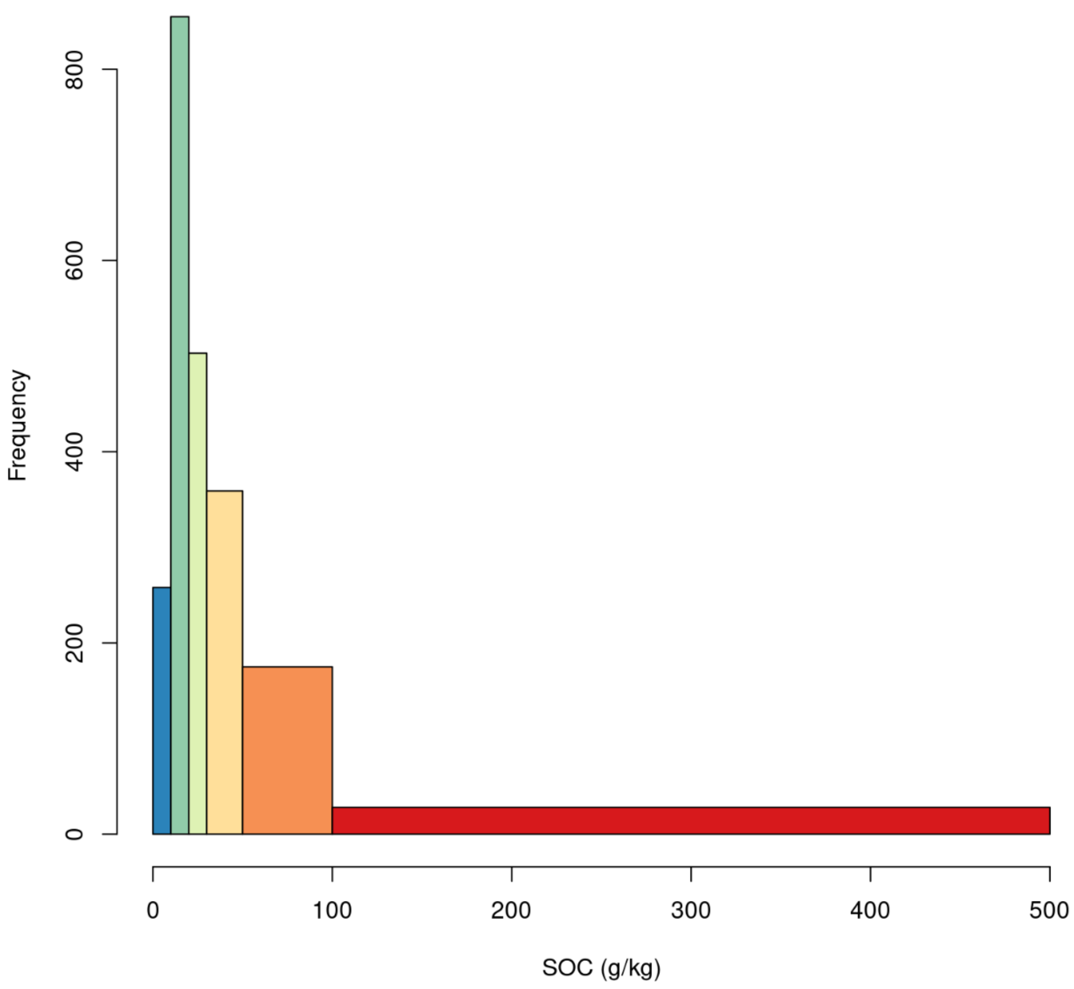
SIC

(Collected at 0-30 cm depth)



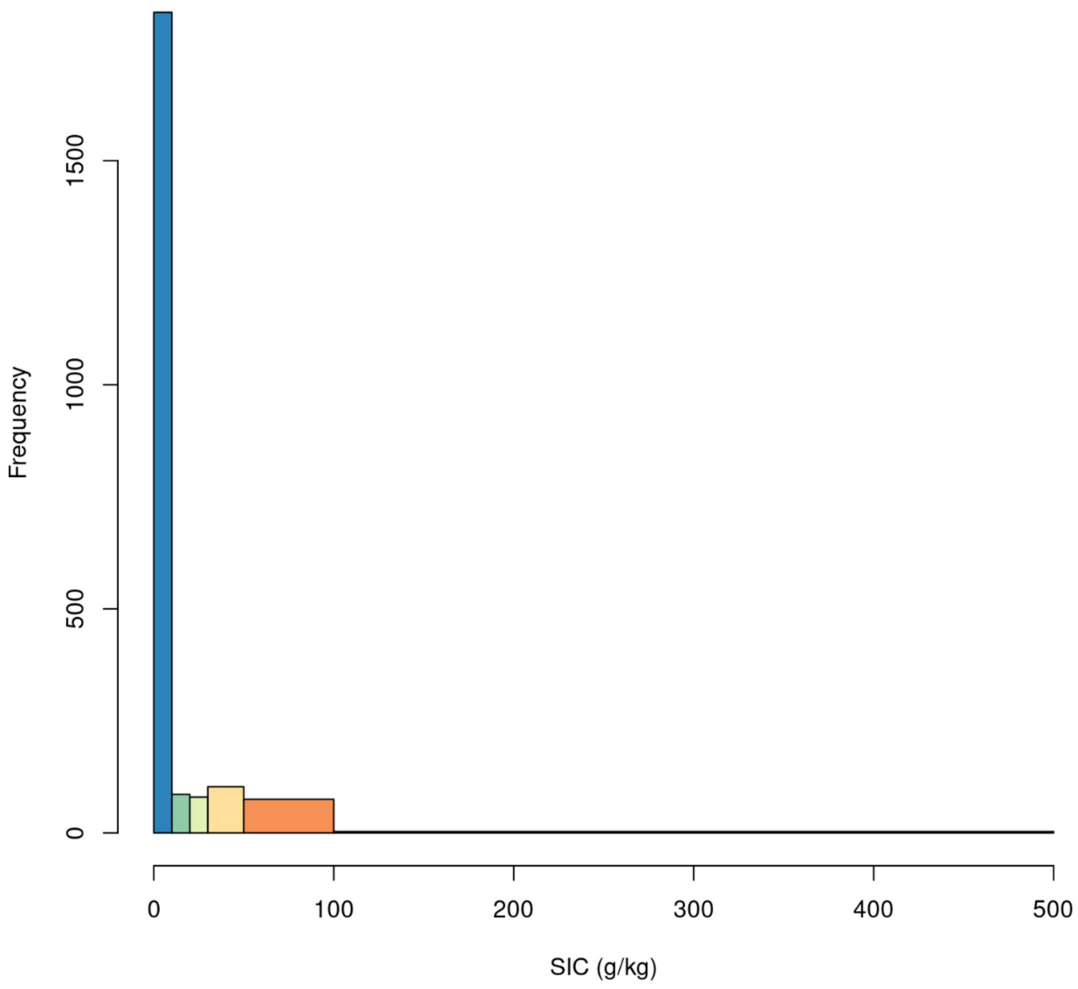
SOC

(Collected at 0-30 cm depth)



SIC

(Collected at 0-30 cm depth)



The Vis-NIR spectral measurements

Visible Near-infrared reflectance spectroscopy
(LabSpec 2500, ASD)

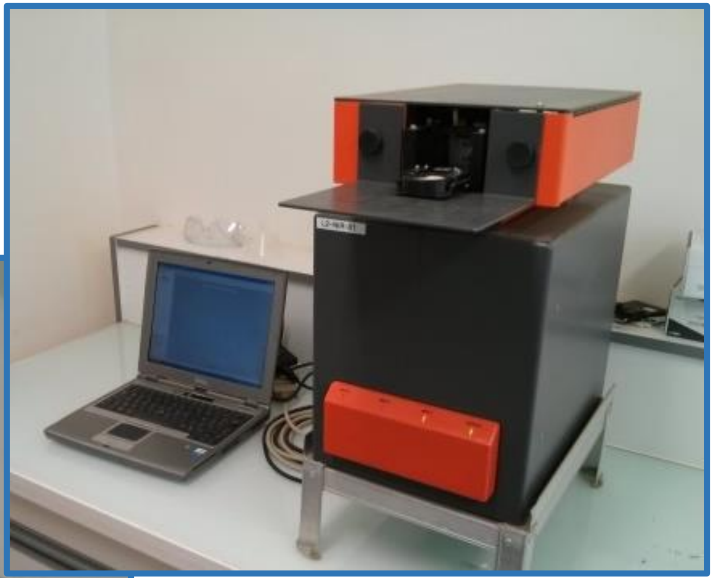
- 350-2500 nm range
- 80 mm² area measured
- 32 co-added scans recorded as Absorbance

Soil Sample preparation:
- 0.2 mm ground
- oven-dried at 40°C

The NIR spectral measurements

Near-infrared reflectance spectroscopy
(FOSS NIRSystems 5000)

- 1100-2500 nm range
- 42 mm² area measured
- 5 g subsample
- 32 co-added scans recorded as Absorbance



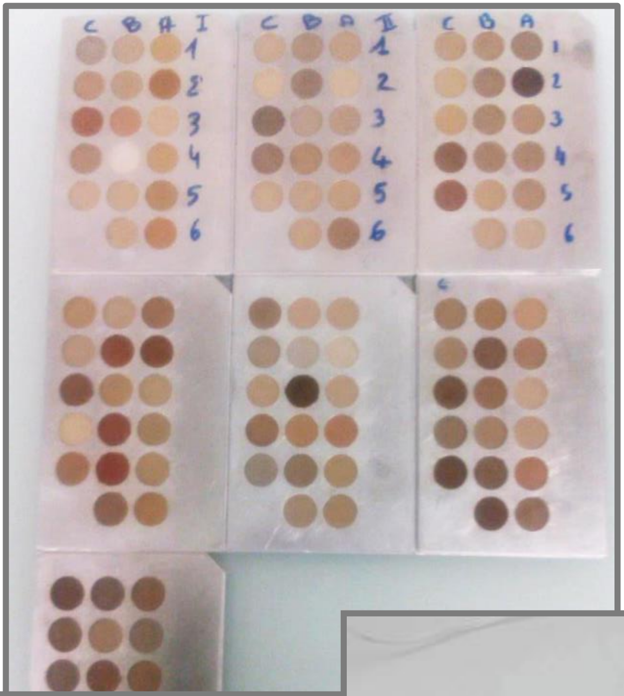
Soil Sample preparation:
- 0.2 mm ground
- oven-dried at 40°C

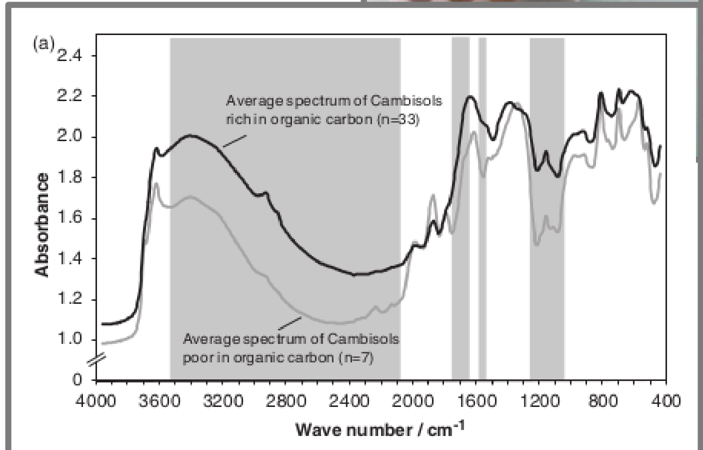
The MIR spectral measurements

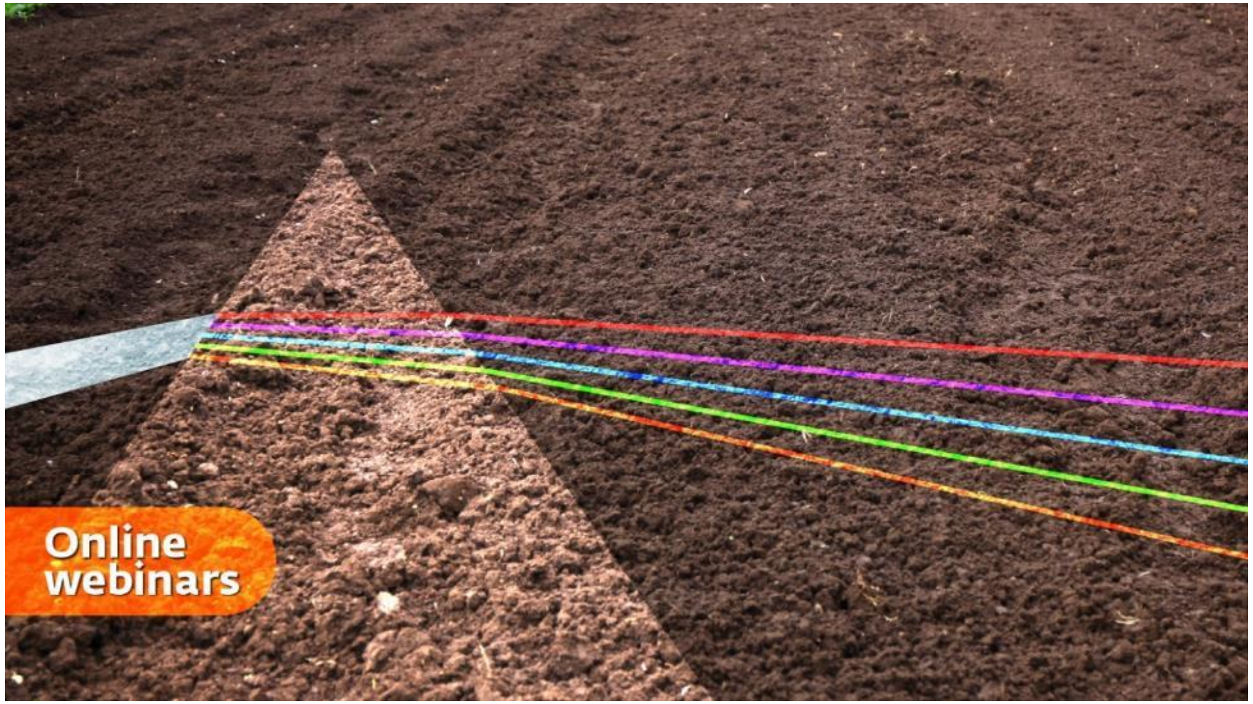
Mid-infrared reflectance spectroscopy
Thermo Nicolet 6700 FTIR

- 2500-25000 nm range
- 13 mm² area measured
- 0.5 g subsample
- 32 co-added scans recorded as Absorbance

Soil Sample preparation:
- 0.2 mm ground
- oven-dried at 40°C







Can we use the French spectral library for soil properties estimation *at National Scale (in French Territory)*?

Can we use the French spectral library for soil properties estimation *at National Scale (in French territory)*?

=> Scenario 1

	1st scenario
Area coverage	Calib DB = Test DB
Database sizes	Calib DB > Test DB

Questions

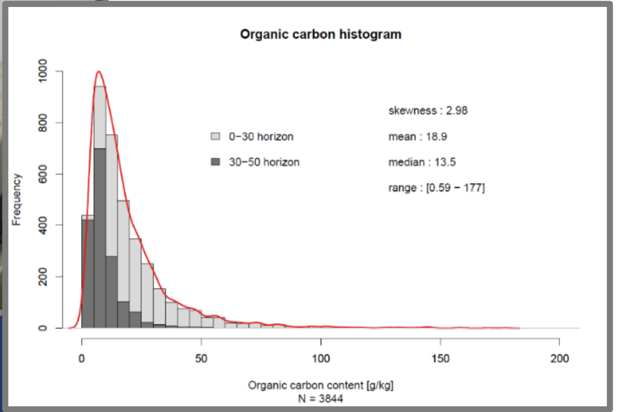
- Which wavelength range performs the best? NIR, MIR?
- What is the optimum number of Calibration samples?
- How to select Calibration data?

Clairotte et al., 2016. Geoderma

Clairotte et al., 2016. Geoderma

Questions

- Which wavelength range performs the best? NIR, MIR?
- What is the optimum number of Calibration samples?
- How to select Calibration data?



Data

- ✓ ~3800 Soil samples of RMQS
Collected at 0-30 cm and 30-50 cm depth
- ✓ NIR and MIRS spectra
- ✓ Soil Organic Carbon (SOC)

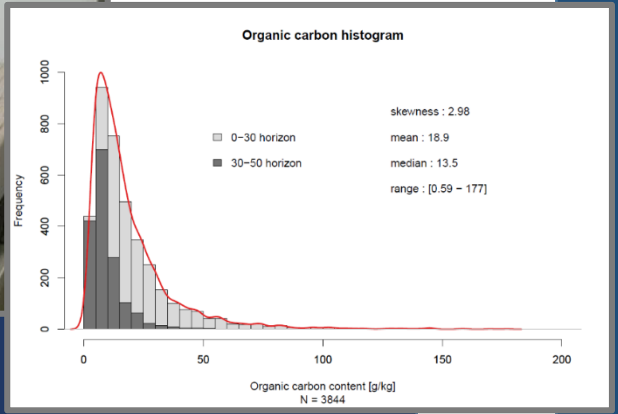
Clairotte et al., 2016. Geoderma

Questions

- Which wavelength range performs the best? NIR, MIR?
- What is the optimum number of Calibration samples?
- How to select Calibration data?

Methods

- ✓ PLSR to build regression models
- ✓ 10 % of samples (= 380) used for Independent Validation
- ✓ 10 % of samples (= 380) used for model tuning
- ✓ Calibration Data selected among remaining samples (~3040 samples):
  - randomly (with 10 replicates)
  - by Kennard-Stone algorithm (with 10 replicates)
  - By spectral neighbours
- ✓ N_{cal} = *20 - 100%* of the remaining samples (~3040 samples)

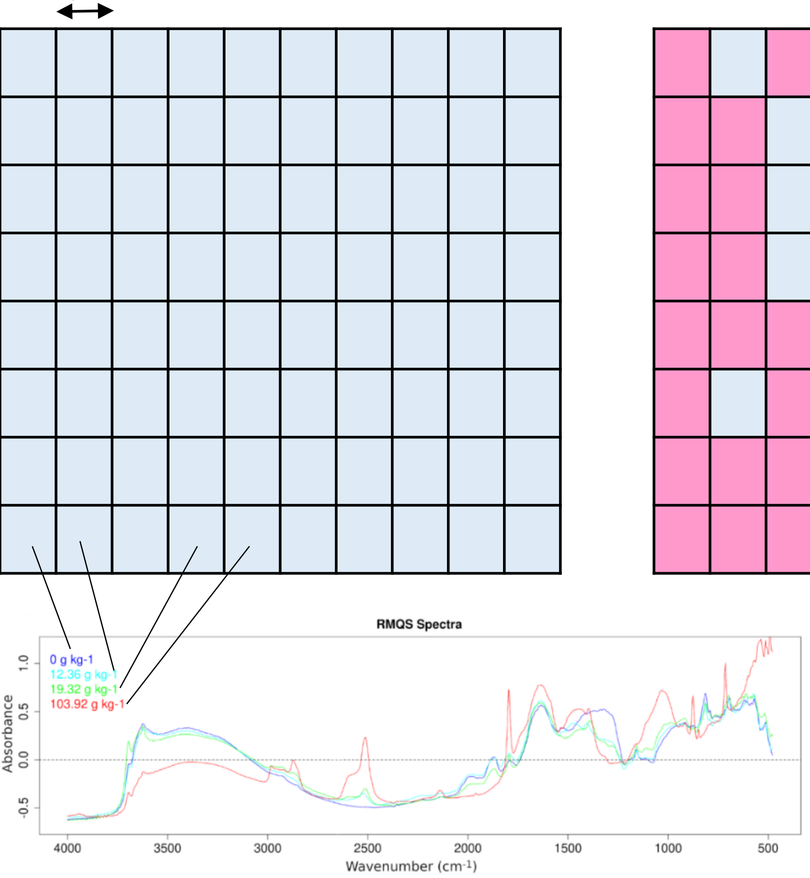


Data

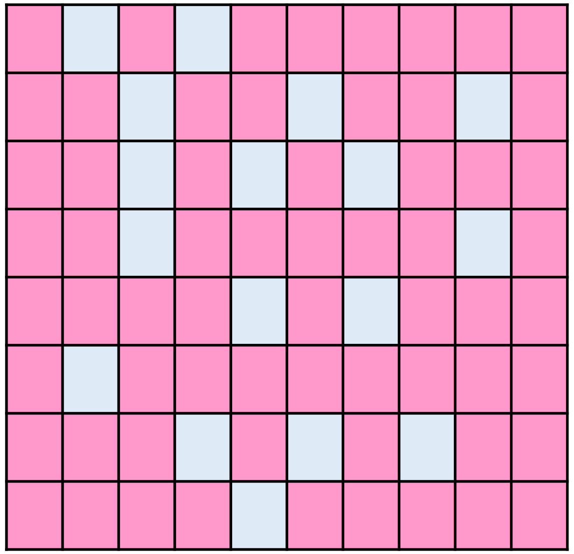
- ✓ ~3800 Soil samples of RMQS Collected at 0-30 cm and 30-50 cm depth
- ✓ NIR and MIRS spectra
- ✓ Soil Organic Carbon (SOC)

3 ways for the Calibration Data selection

RMQS dataset



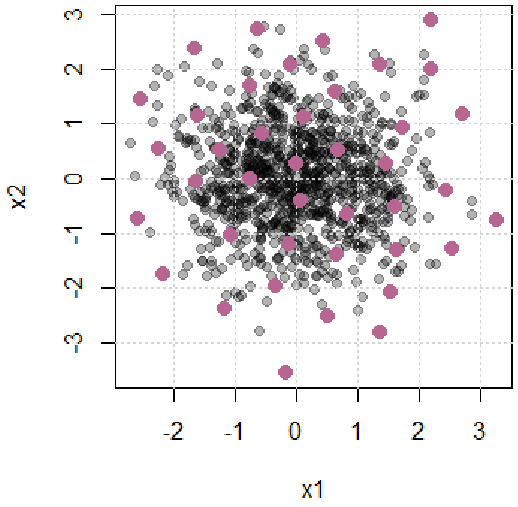
Random selection



Kennard-Stone

Based on spectral representativeness

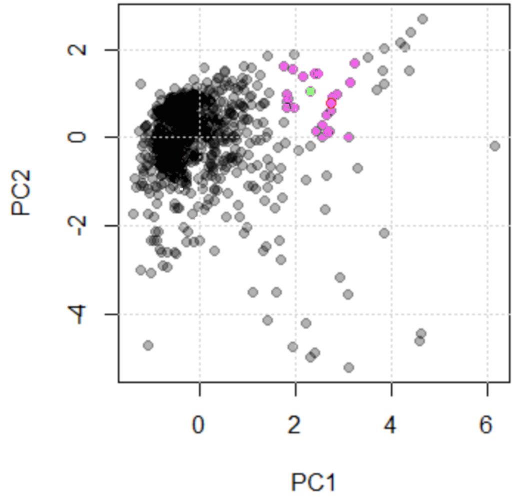
(Kennard and Stone, 1969)



Local calibration samples

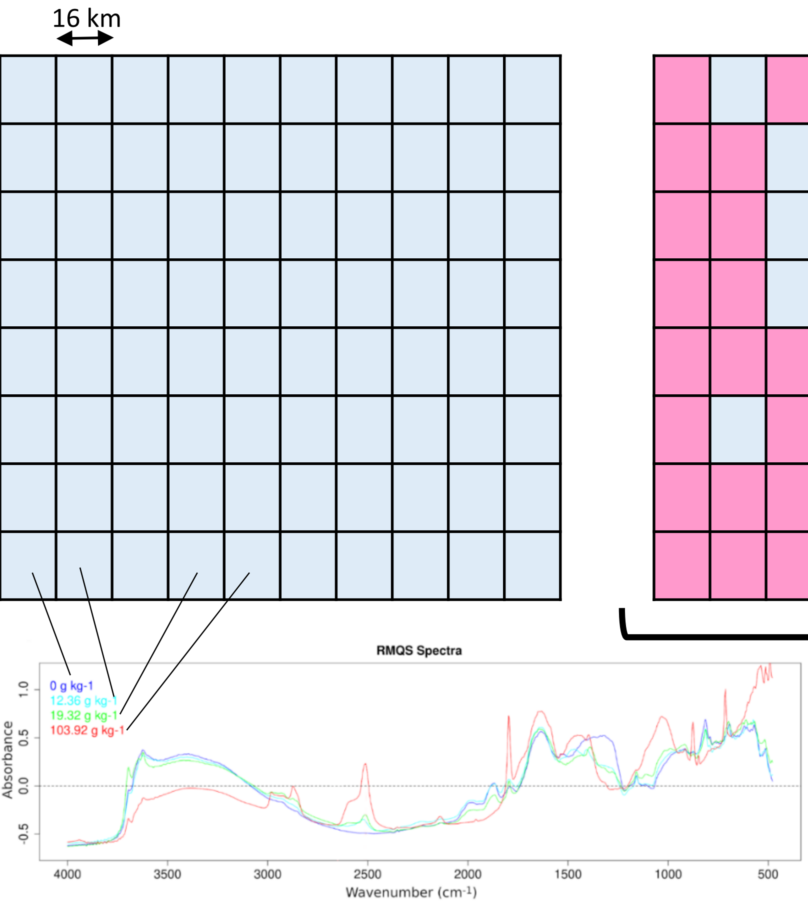
Based on Spectral neighbours

(Shenk et al., 1997)

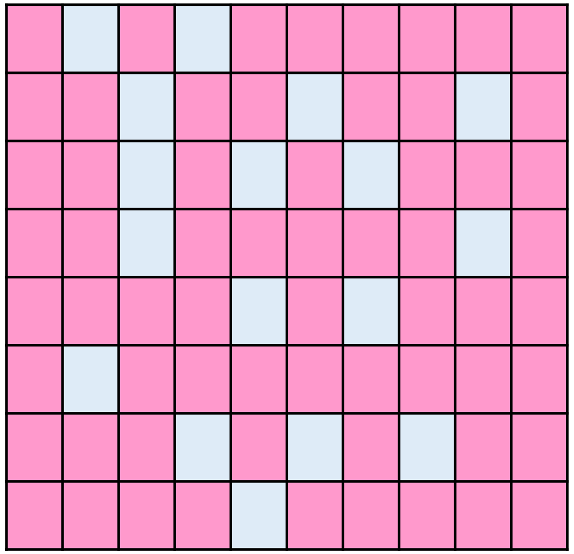


3 ways for the Calibration Data selection

RMQS dataset



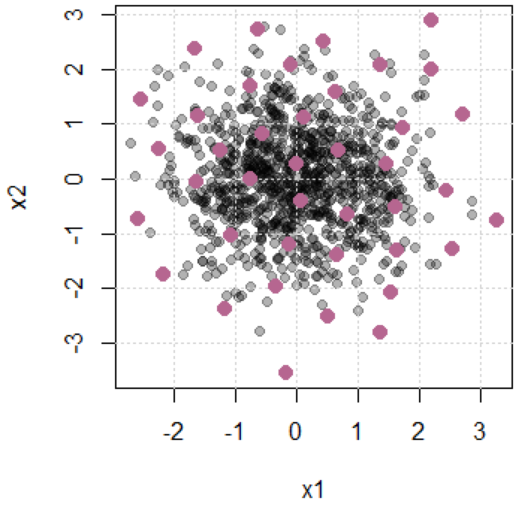
Random selection



Kennard-Stone

Based on spectral representativeness

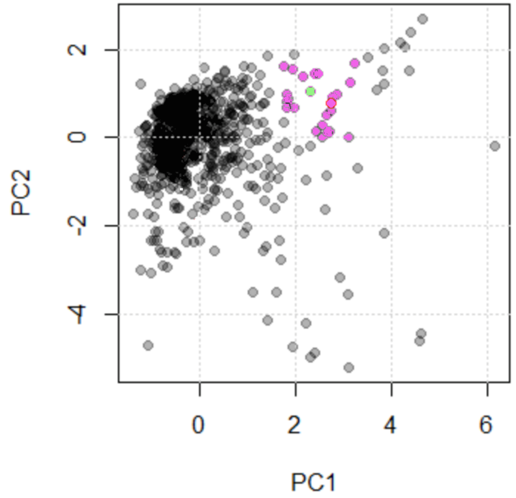
(Kennard and Stone, 1969)



Local calibration samples

Based on Spectral neighbours

(Shenk et al., 1997)



Local PLSR

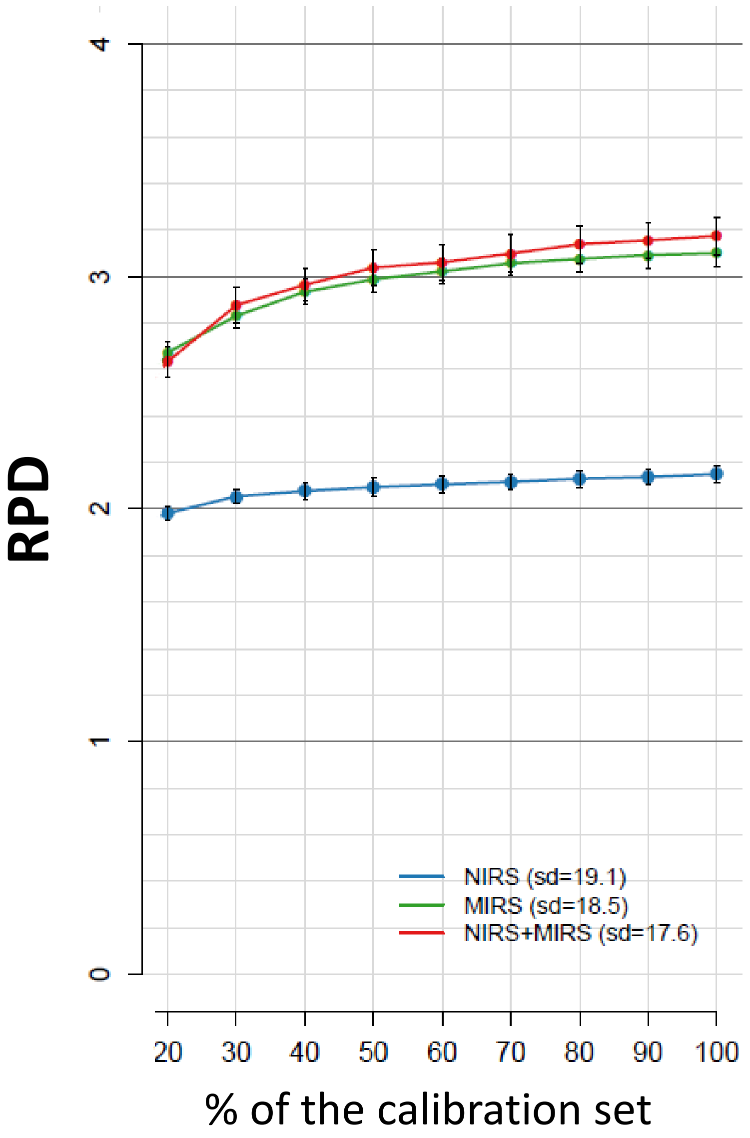
Global PLSR

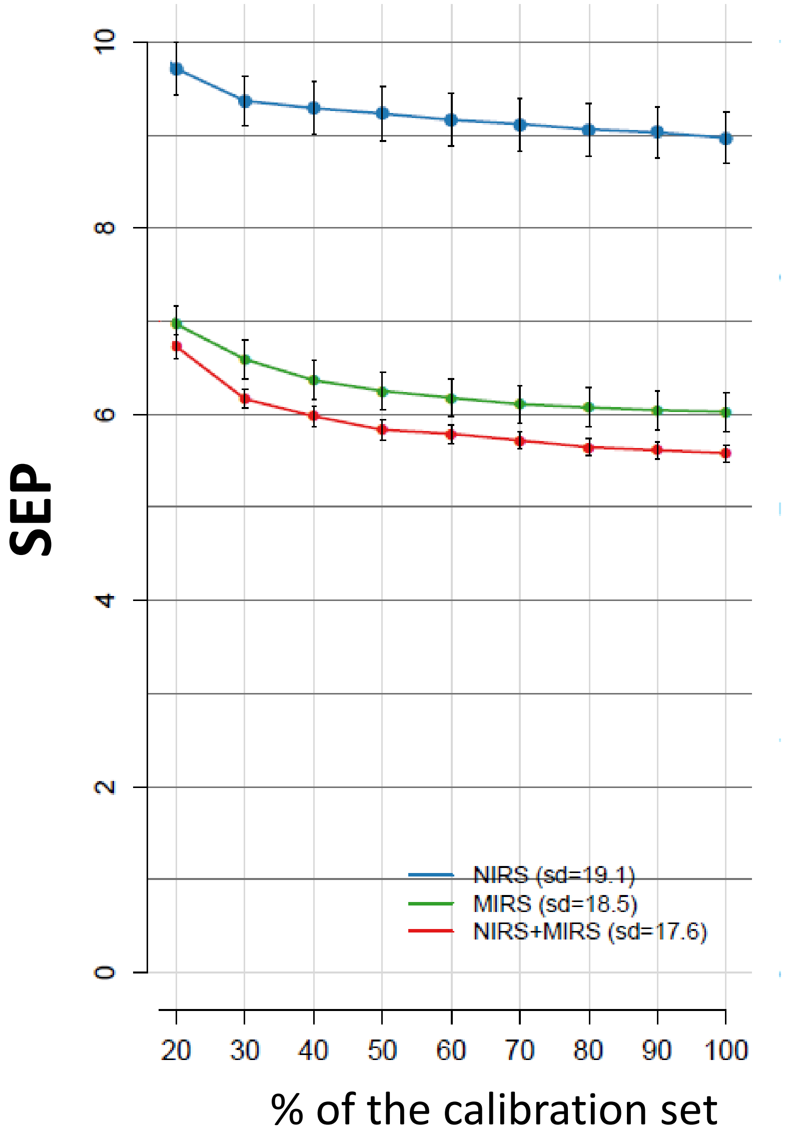
Three strategies for Calibration set Selection

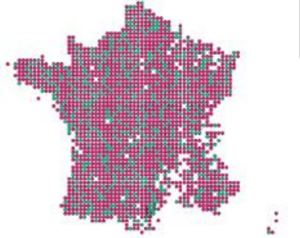
Soil Organic Carbon (g kg^{-1})

___ *Random selection*

VAL.set (380)







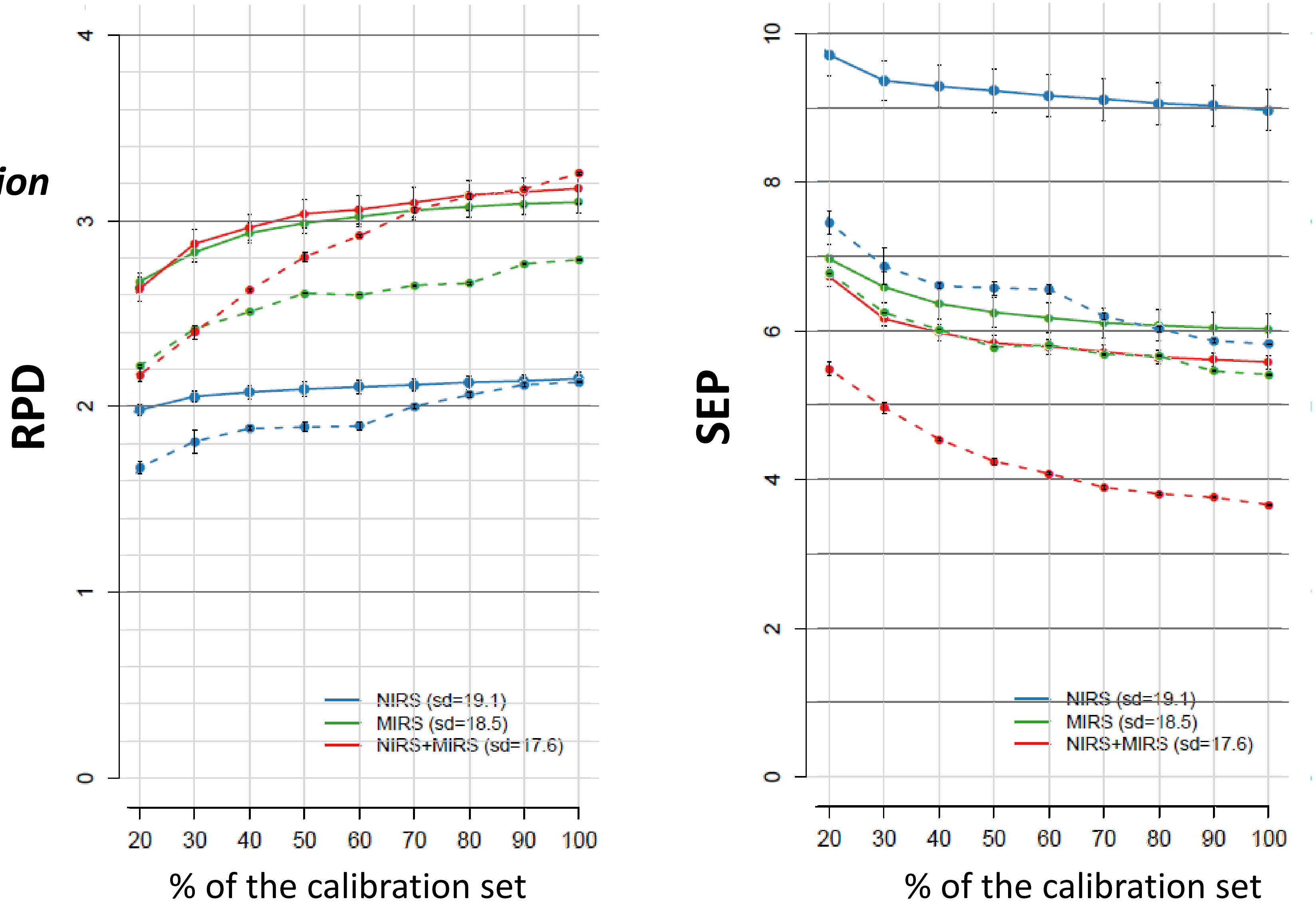
Three strategies for Calibration set Selection

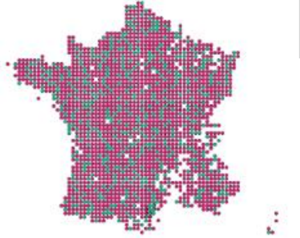
Soil Organic Carbon (g kg^{-1})

____ Random selection

---- Kennard-Stone selection

VAL.set (380)





Three strategies for Calibration set Selection

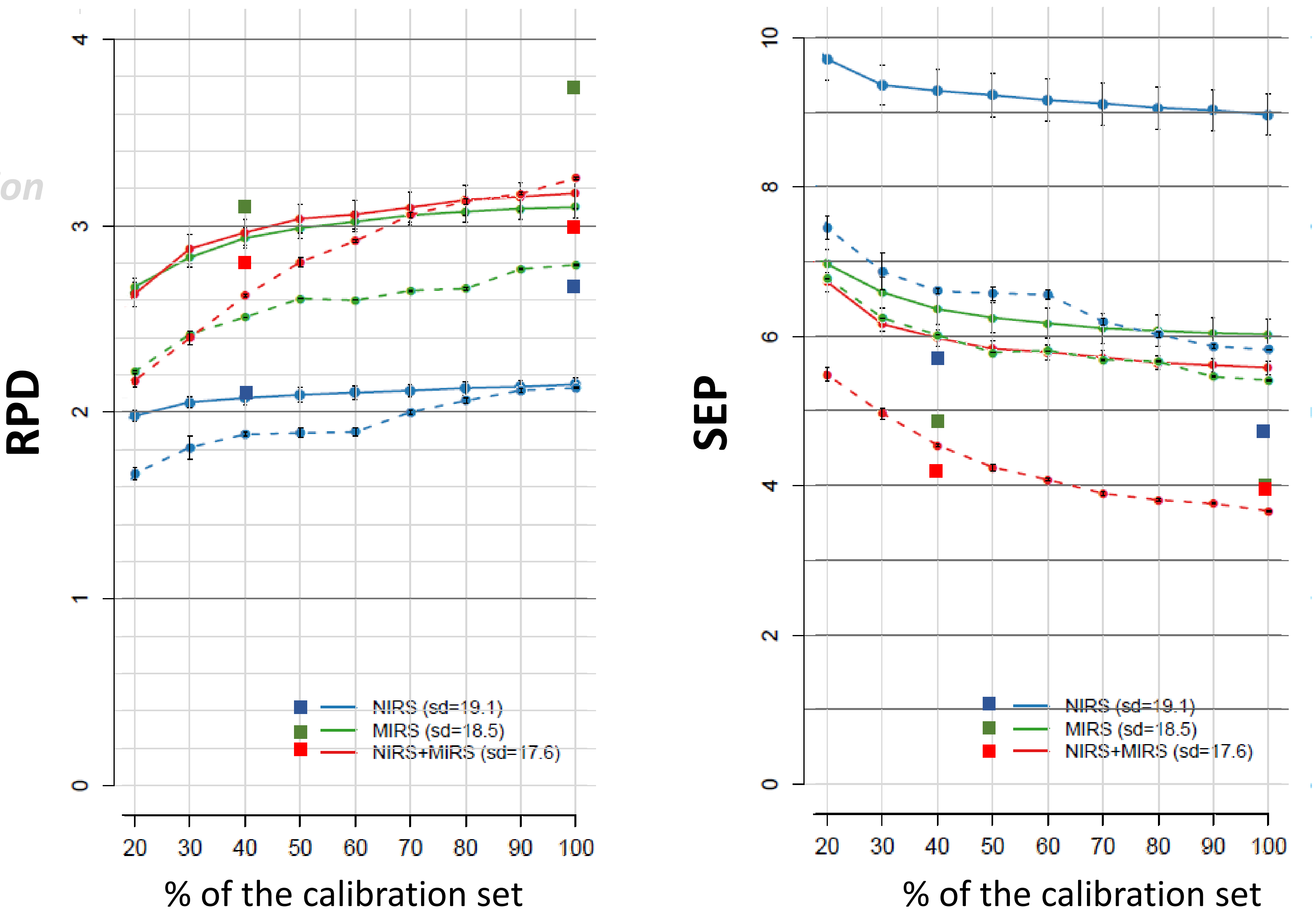
Soil Organic Carbon (g kg^{-1})

____ Random selection

---- Kennard-Stone selection

■ ***Local Calibration***

VAL.set (380)

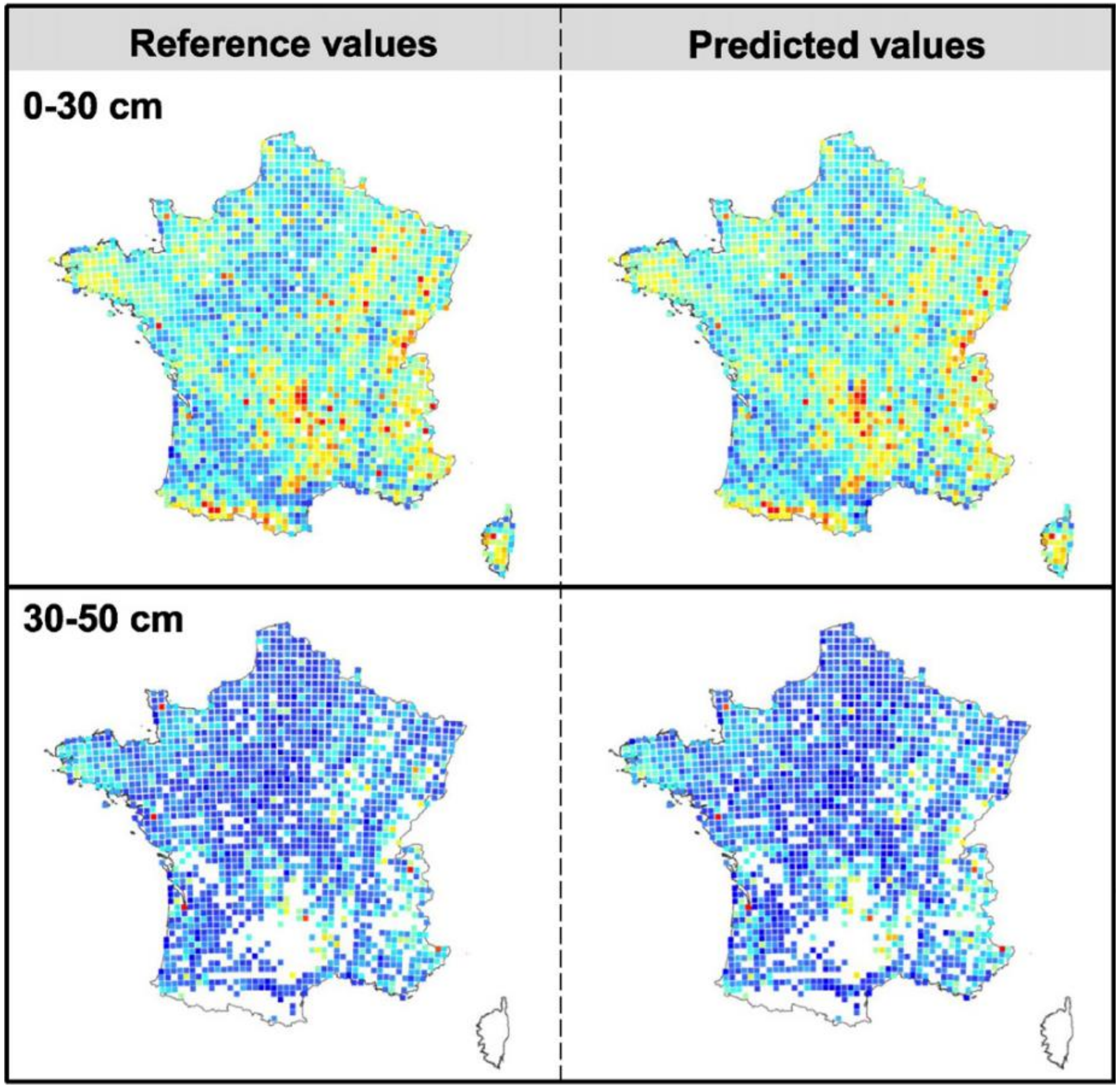


MIRS

Local PLSR

Calibration set = 100 %
= 3076 samples

Validation set = 380 samples



SEP = 4.7 g/kg

Soil Organic Carbon (g kg^{-1})

- 0.0 - 2.5
- 2.6 - 5.0
- 5.1 - 7.5
- 7.6 - 10.0
- 10.1 - 12.5
- 12.6 - 15.0
- 15.1 - 20.0
- 20.1 - 25.0
- 25.1 - 30.0
- 30.1 - 35.0
- 35.1 - 40.0
- 40.1 - 50.0
- 50.1 - 60.0
- 60.1 - 70.0
- 70.1 - 80.0
- 80.1 - 90.0
- 90.1 - 100.0
- 100.1 - 150.0

Clairotte et al., 2016. Geoderma

Questions

- Which wavelength range performs the best? NIR, MIR?
- What is the optimum number of Calibration samples?
- How to select Calibration data?

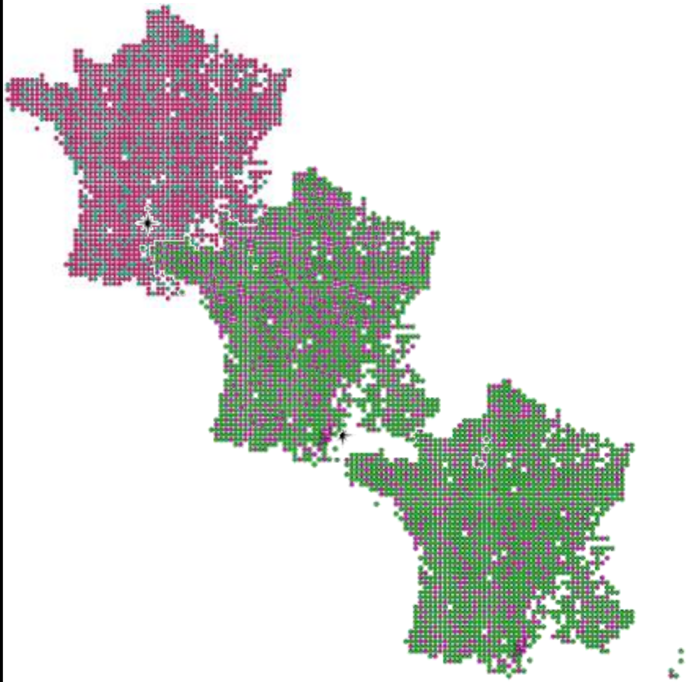
Highlights

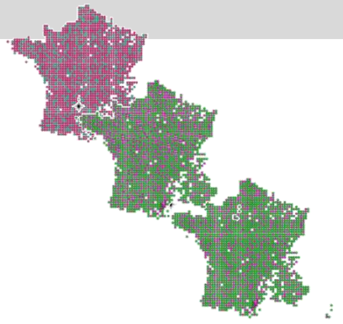
- MIR >> NIR for SOC prediction at our National scale.
- The greater the number of samples, the better the performance.
- Even 20% of the National Dataset provided correct accuracy of SOC estimation.
- Spectral neighbours >> Kennard-Stone >> Random
- Maps of predicted properties appropriate for large-scale soil inventories and mapping studies, but not for accurate carbon monitoring

Can we use the French spectral library for soil properties estimation *at National Scale (in French territory)*?

=> Scenario 2

	1st scenario	2nd scenario
Area coverage	Calib DB = Test DB	Calib DB = Test DB
Database sizes	Calib DB > Test DB	**Nb Calib DB = 100% - Nb Valid DB**

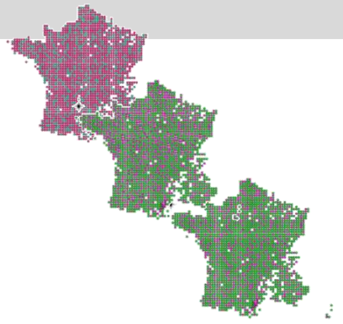




Grinand et al., 2012. EJSS

Question

- Which performances can we expect depending on the number of Calibration samples (from higher to lower than the number of Validation samples)?



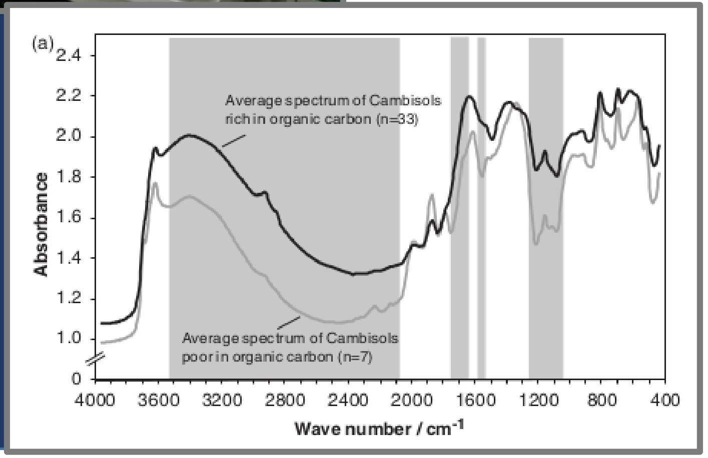
Grinand et al., 2012. EJSS

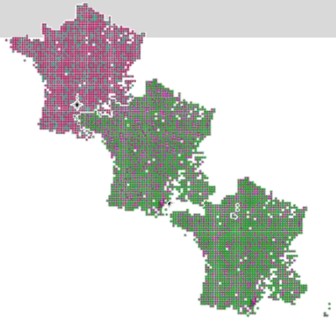
Question

- Which performances can we expect depending on the number of Calibration samples (from higher to lower than the number of Validation samples)?

Data

- ✓ 2086 soil samples of RMQS Collected at 0-30 cm depth
- ✓ MIR spectra
- ✓ Soil Organic Carbon (SOC) and Soil Inorganic Carbon (SIC)





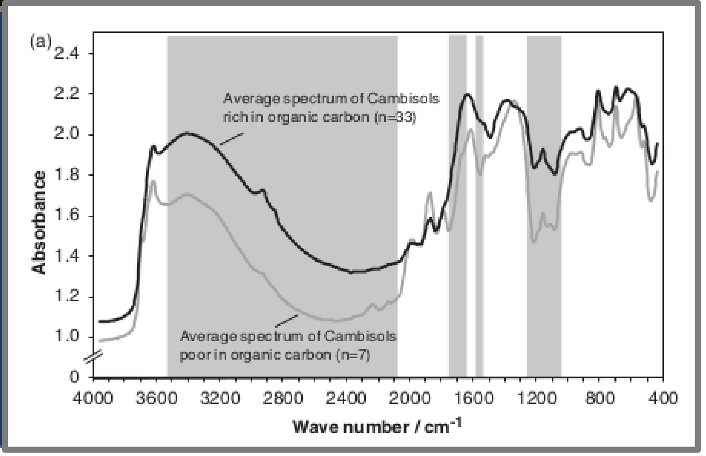
Grinand et al., 2012. EJSS

Question

- Which performances can we expect depending on the number of Calibration samples (from higher to lower than the number of Validation samples)?

Data

- ✓ 2086 soil samples of RMQS Collected at 0-30 cm depth
- ✓ MIR spectra
- ✓ Soil Organic Carbon (SOC) and Soil Inorganic Carbon (SIC)



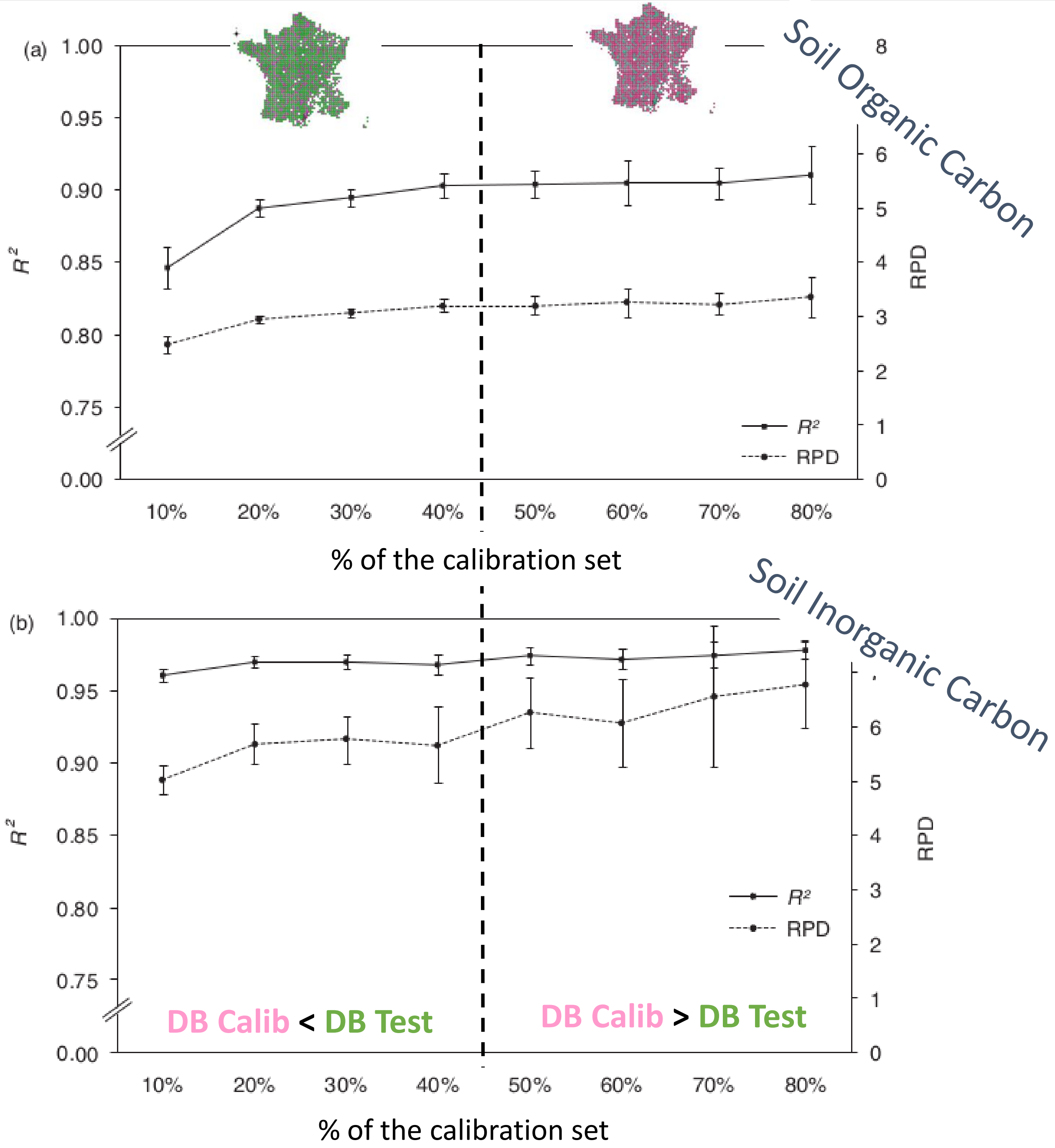
Methods

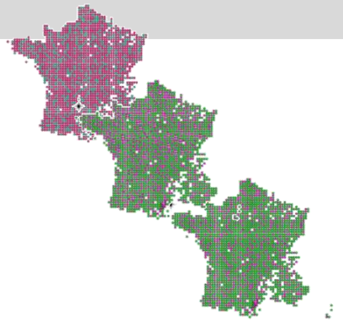
- ✓ PLSR to build regression models
- ✓ Calibration using N_{cal} = *10-80%* of the set
- ✓ Five random selection of samples
- ✓ Validation using remaining samples (100%- N_{cal})

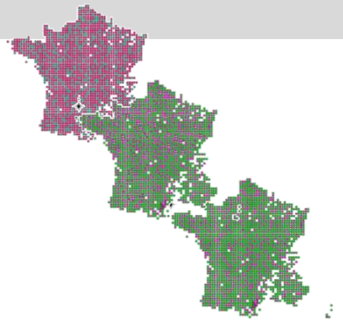
Random selection of Calibration set

CAL.set (X %)

VAL.set (100 -X %)

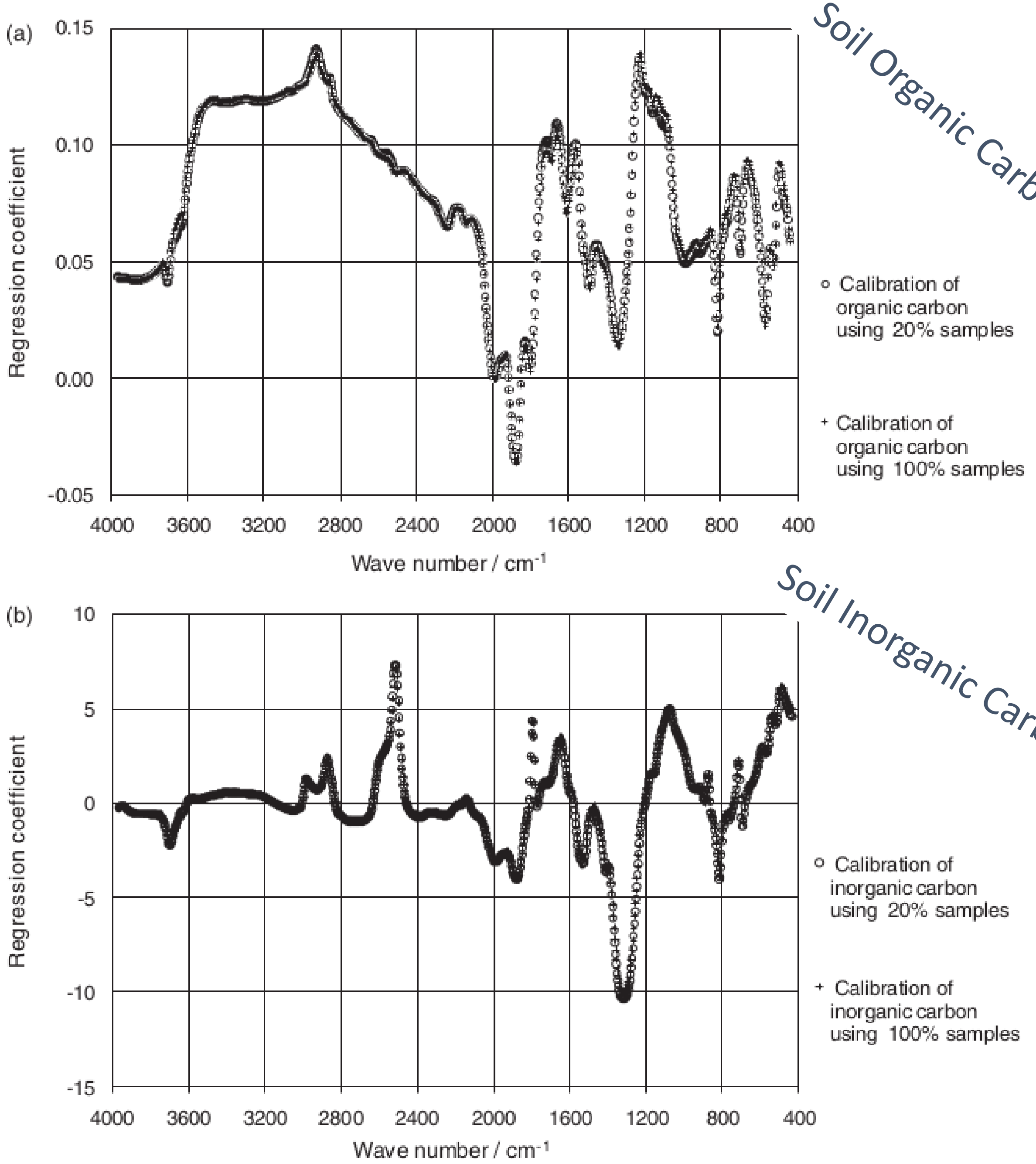


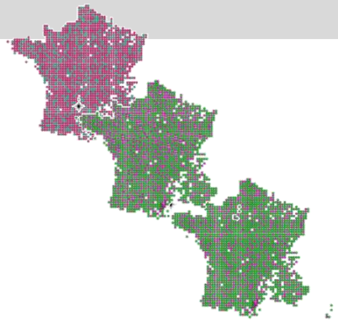




Soil Organic Carbon

Soil Inorganic Carbon





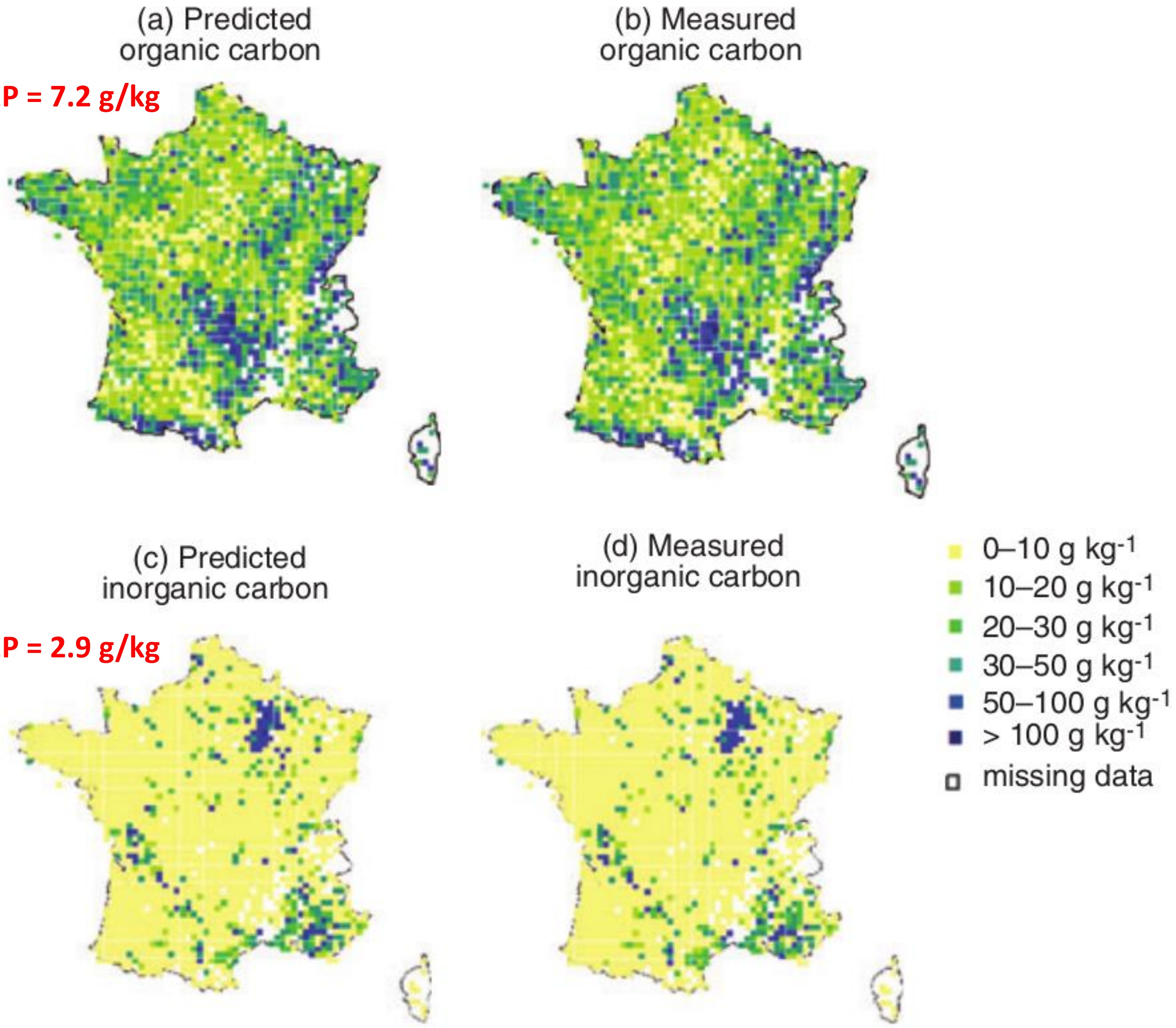
Random selection of Calibration set

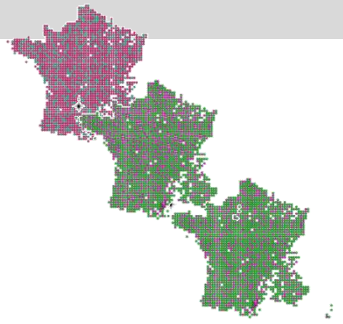
CAL.set (20 %)

VAL.set (80 %)

SEP = 7.2 g/kg

SEP = 2.9 g/kg





Grinand et al., 2012. EJSS

Question

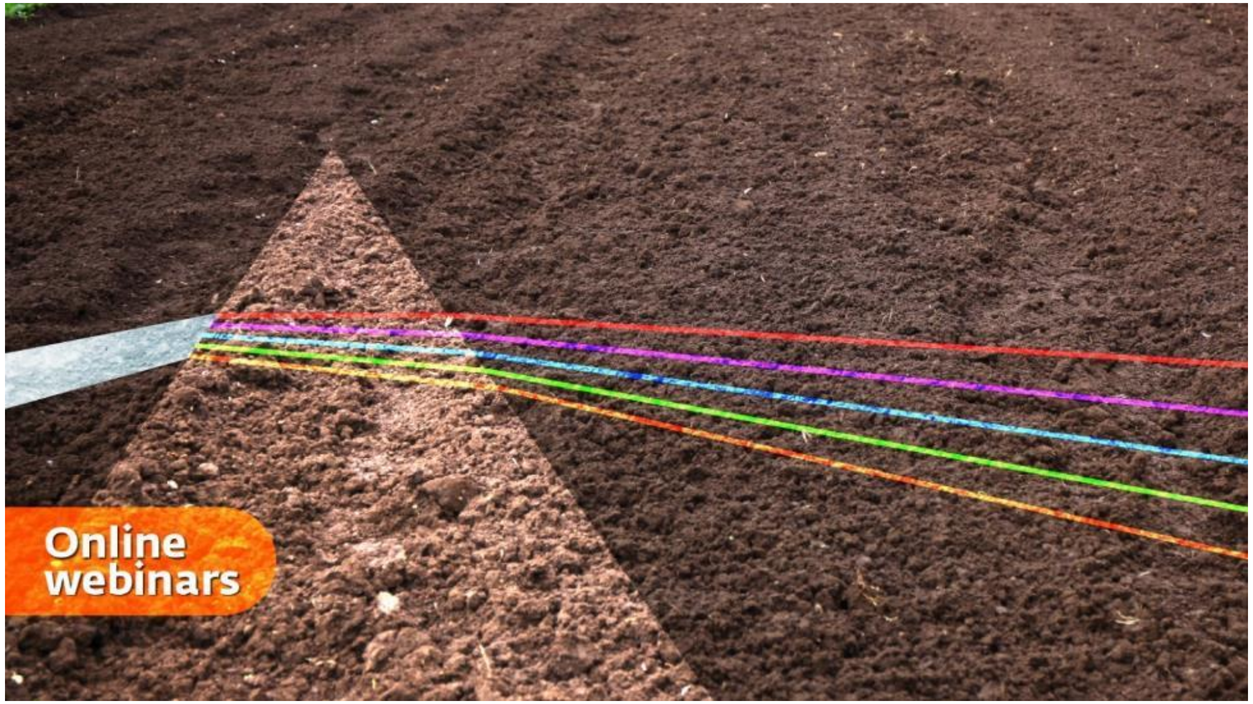
- Which performances can we expect depending on the number of Calibration samples (from higher to lower than the number of Validation samples)?

Highlights

- Optimal calibrations were achieved by using 20% of the RMQS dataset for both SOC and SIC predictions….

  …. With a similarity between regression coefficients of calibration equations that used 20 and 100% of the total set

- Maps of predicted properties still appropriate for large-scale soil inventories and mapping studies, but not for accurate carbon monitoring

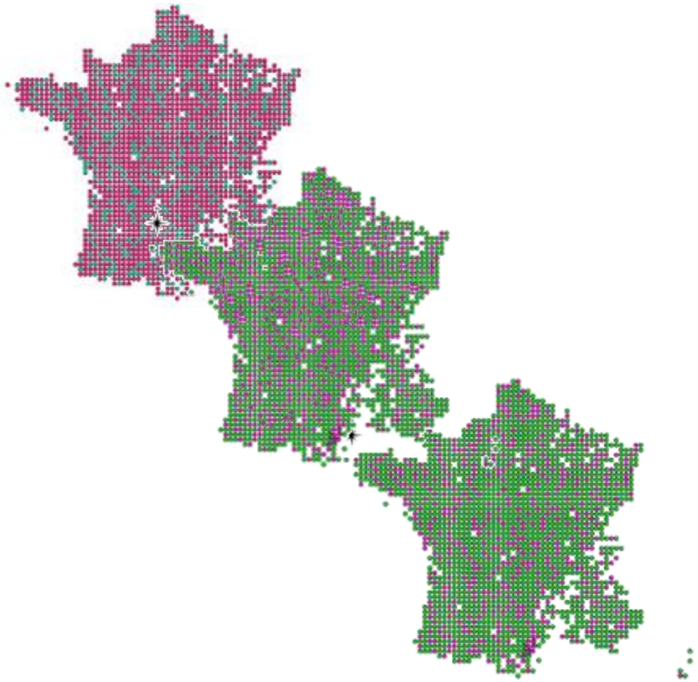


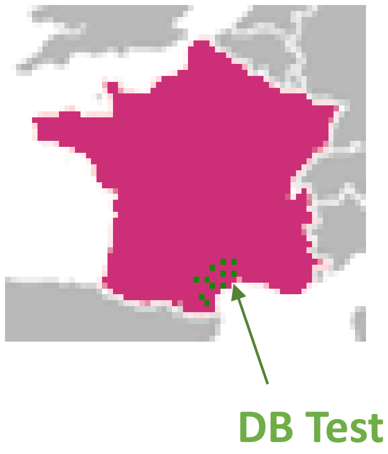
Can we use the French spectral library for soil properties estimation *at Regional Scale (in France)*?

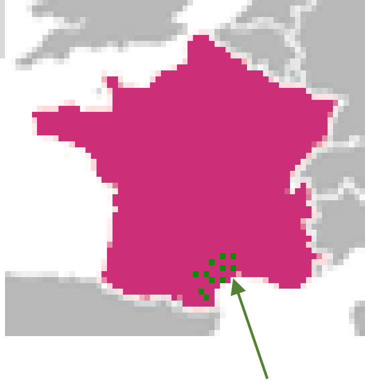
Can we use the French spectral library for soil properties estimation *at Regional Scale (in French territory)*?

=> Scenario 3

	1st scenario	2nd scenario	3rd scenario
Area coverage	Calib DB = Test DB	Calib DB = Test DB	Test DB ϵ Calib DB
Database sizes	Calib DB > Test DB	Nb Calib DB = 100% - Nb Calib DB	Calib DB >> Test DB



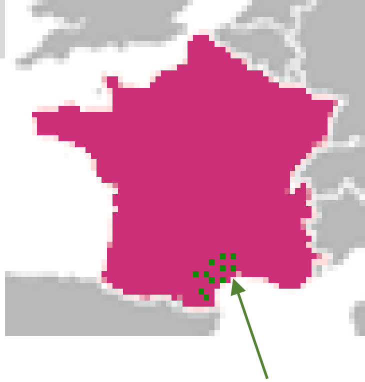




Barthes et al., 2020. Geoderma

Questions

- Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area?
- Which approach is appropriate in this specific situation?



Barthes et al., 2020. Geoderma

Questions

- ➢ Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area?
- ➢ Which approach is appropriate in this specific situation?

Data

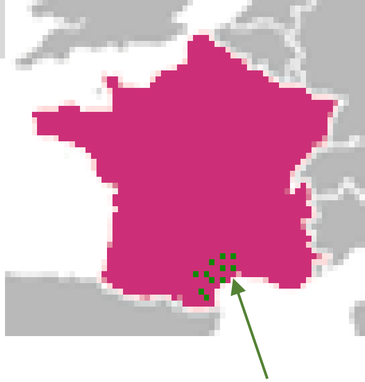
✓ 2178 soil samples of RMQS
Collected at 0-30 cm depth

✓ 164 regional soil samples from
Vineyard plots (South of France)

- A sample per vineyard plot
- Collected at 0-15 cm depth

✓ MIRS spectra

✓ Soil Inorganic Carbon (SIC)



Questions

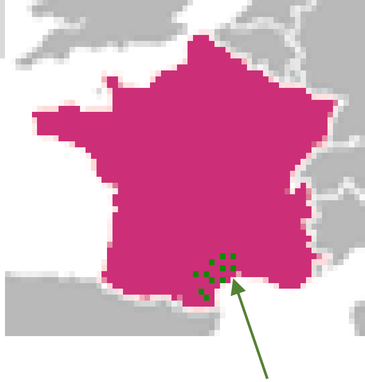
- Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area?
- Which approach is appropriate in this specific situation?

Data

- ✓ 2178 soil samples of RMQS Collected at 0-30 cm depth
- ✓ 164 regional soil samples from Vineyard plots (South of France)
  - A sample per vineyard plot
  - Collected at 0-15 cm depth
- ✓ MIRS spectra
- ✓ Soil Inorganic Carbon (SIC)

Methods

- ✓ PLSR to build regression models
- ✓ Four tested approaches. Calibration from
  1. 2178 soil samples of RMQS
  2. 2178 soil samples of RMQS + spiked Regional samples
  3. spectral neighbours of RMQS
  4. spectral neighbours of RMQS + spiked Regional samples
- ✓ Validation using 134 among the 164 regional samples
- ✓ 30 remaining samples possibly used for spiking



Questions

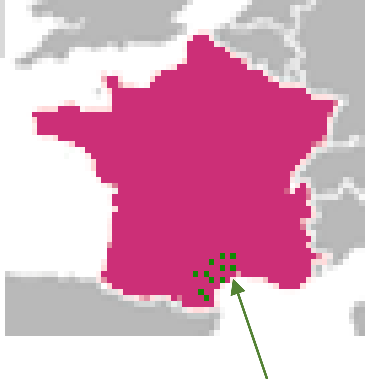
- Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area?
- Which approach is appropriate in this specific situation?

Data

- ✓ 2178 soil samples of RMQS
  Collected at 0-30 cm depth
- ✓ 164 regional soil samples from Vineyard plots (South of France)
  - A sample per vineyard plot
  - Collected at 0-15 cm depth
- ✓ MIRS spectra
- ✓ Soil Inorganic Carbon (SIC)

Methods

- ✓ PLSR to build regression models
- ✓ Four tested approaches. Calibration from
  1. 2178 soil samples of RMQS
  2. 2178 soil samples of RMQS + spiked Regional samples
  3. spectral neighbours of RMQS
  4. spectral neighbours of RMQS + spiked Regional samples
- ✓ Validation using 134 among the 164 regional samples
- ✓ 30 remaining samples possibly used for spiking



Questions

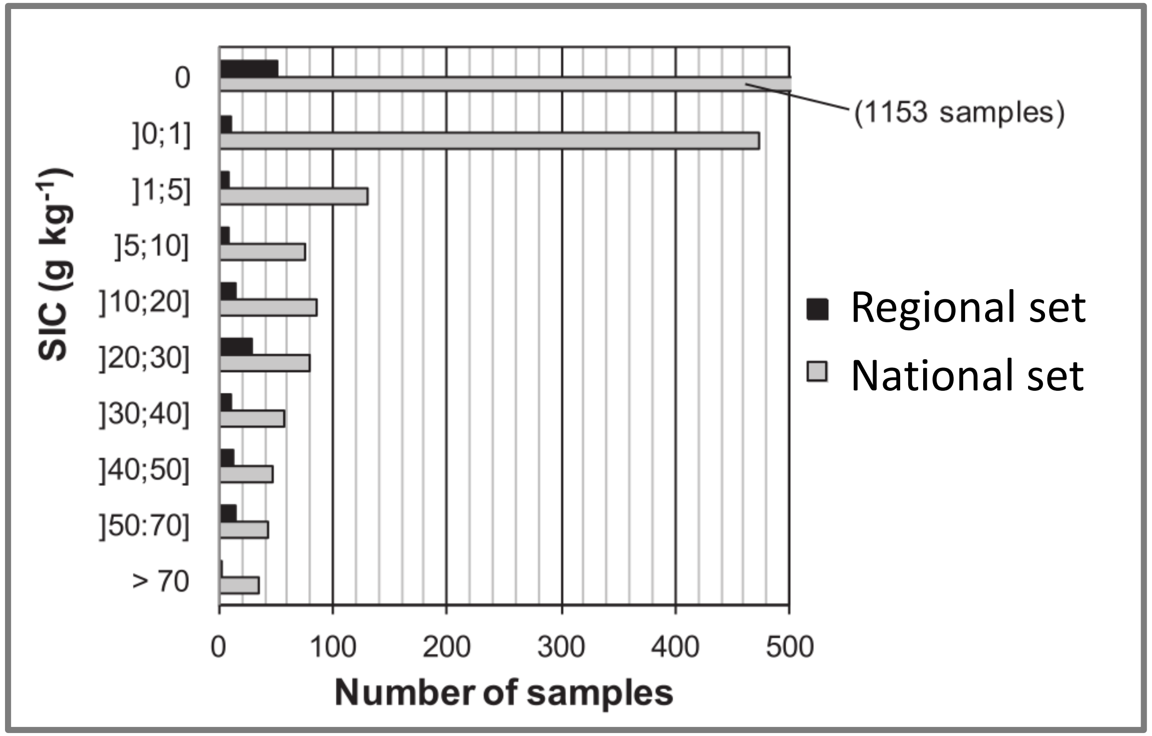
- Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area?
- Which approach is appropriate in this specific situation?

Data

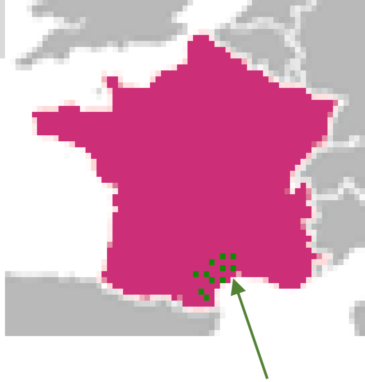
- ✓ 2178 soil samples of RMQS Collected at 0-30 cm depth
- ✓ 164 regional soil samples from Vineyard plots (South of France)
  - A sample per vineyard plot
  - Collected at 0-15 cm depth
- ✓ MIRS spectra
- ✓ Soil Inorganic Carbon (SIC)

Methods

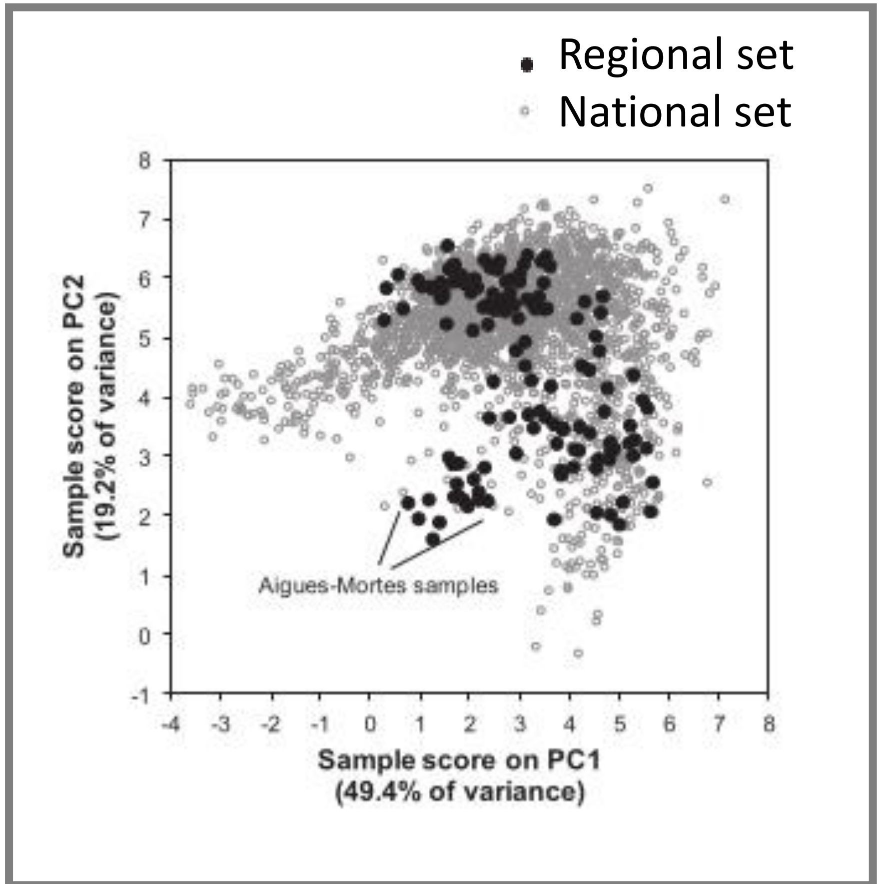
- ✓ PLSR to build regression models
- ✓ Four tested approaches. Calibration from
  1. 2178 soil samples of RMQS
  2. 2178 soil samples of RMQS + spiked Regional samples
  3. spectral neighbours of RMQS
  4. spectral neighbours of RMQS + spiked Regional samples
- ✓ Validation using 134 among the 164 regional samples
- ✓ 30 remaining samples possibly used for spiking



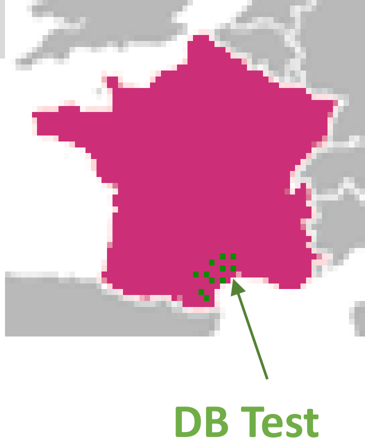
SIC (g kg-1)
0
]0;1]
]1;5]
]5;10]
]10;20]
]20;30]
]30;40]
]40;50]
]50:70]
> 70
(1153 samples)
Regional set
National set
0
100
200
300
400
500
Number of samples



DB Test



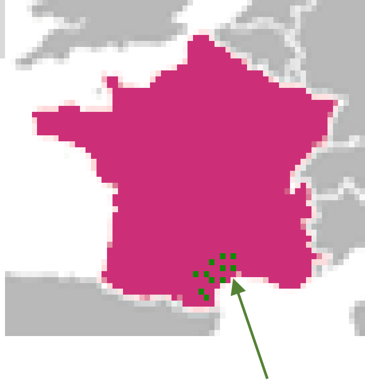
Regional set
National set
Sample score on PC2 (19.2% of variance)
Sample score on PC1 (49.4% of variance)
Aigues-Mortes samples



	N_{cal} from RMQS	**N_{spike}**	**N_{val}**	**SEP_{val}**	**R^2_{val}**	**RPD_{val}**
Global Models for SIC predictions	**2178**	**0**	**134**	**5.2**	**0.96**	**3.7**
	2178	10	134	4.9	0.96	3.9

SEP ↘ & RPD ↗

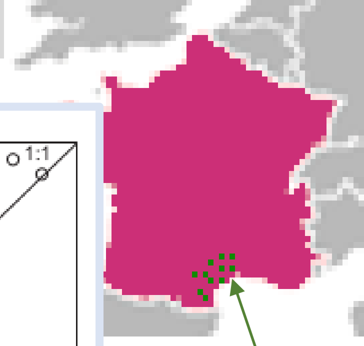
Can we use the French spectral library for soil properties estimation ***at Regional Scale (in French Territory)****? Scenario 3*

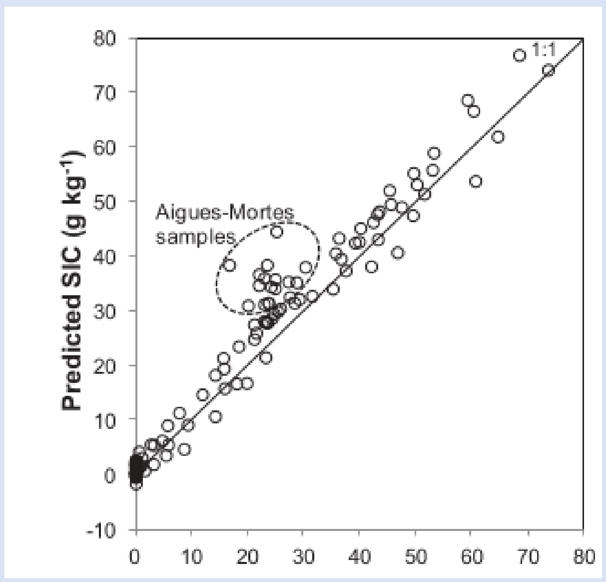


	N_{cal} from RMQS	N_{spike}	N_{val}	SEP_{val}	R^2_{val}	RPD_{val}
Global Models for SIC predictions	**2178**	**0**	**134**	**5.2**	**0.96**	**3.7**
	2178	10	134	4.9	0.96	3.9
Local Models for SIC predictions	From 4 to 2178	0	133	4.8	0.95	4.1
	From 4 to 2178	**10**	**134**	**2.7**	**0.98**	**7.3**
	From 50 to 2178	10	115	2.3	0.99	8.9

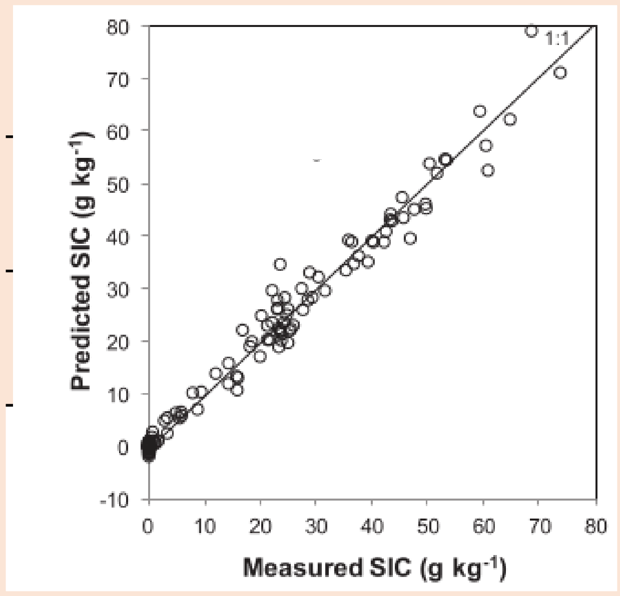
SEP ↘ & RPD ↗

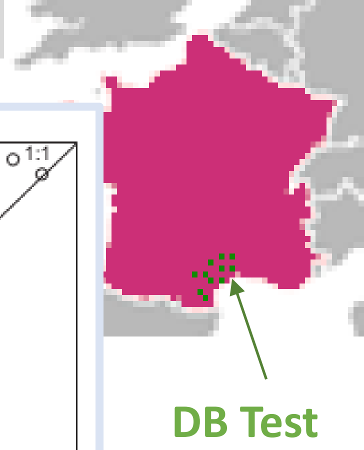
SEP ↘ & RPD ↗



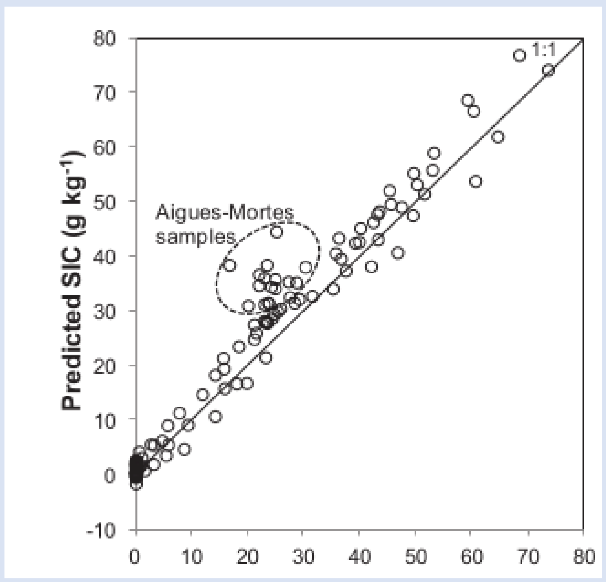


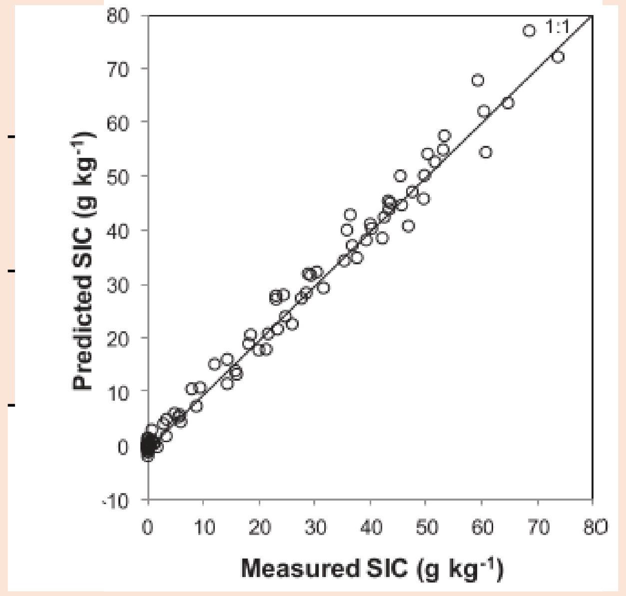
	N_{cal} from RMQS	N_{spike}	N_{val}	SEP_{val}	R^2_{val}	RPD_{val}
Global Models for SIC predictions	**2178**	**0**	**134**	**5.2**	**0.96**	**3.7**
	2178	10	134	4.9	0.96	3.9
Local Models for SIC predictions	From 4 to 2178	0	133	4.8	0.95	4.1
	From 4 to 2178	**10**	**134**	**2.7**	**0.98**	**7.3**
	From 50 to 2178	10	115	2.3	0.99	8.9

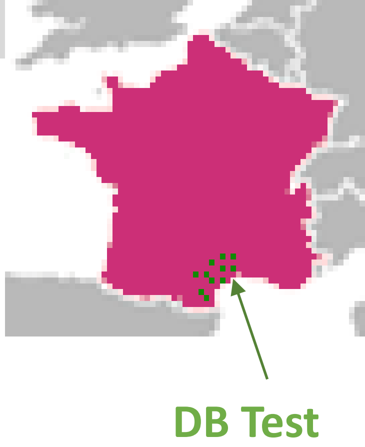




	N_{cal} from RMQS	N_{spike}	N_{val}	SEP_{val}	R^2_{val}	RPD_{val}
Global Models for SIC predictions	**2178**	**0**	**134**	**5.2**	**0.96**	**3.7**
	2178	10	134	4.9	0.96	3.9
Local Models for SIC predictions	From 4 to 2178	0	133	4.8	0.95	4.1
	From 4 to 2178	10	134	2.7	0.98	7.3
	From 50 to 2178	**10**	**115**	**2.3**	**0.99**	**8.9**







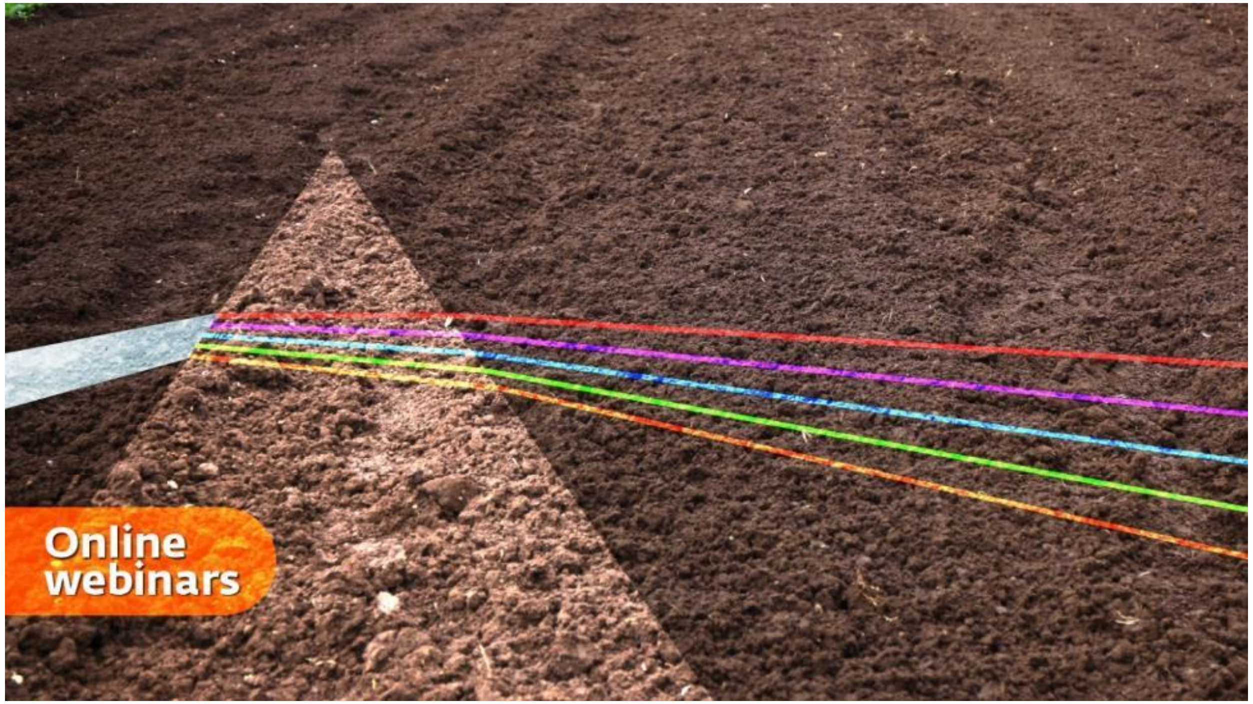
Barthes et al., 2020. Geoderma

Questions

- Which performances can we expect when the validation samples come from a region compared to Calibration samples coming from National area ?
- Which approach is appropriate in this specific situation?

Highlights

- The French dataset may provide accurate SIC prediction at regional scale from Global Model (« classical approach »)
- SIC predictions may be improved by using spiked samples and Local calibration (=Local-PLSR).
- Using the Local-PLSR approach with sufficient number of neighbours may allow to avoid bad predictions on some specific samples poorly represented in the National dataset.

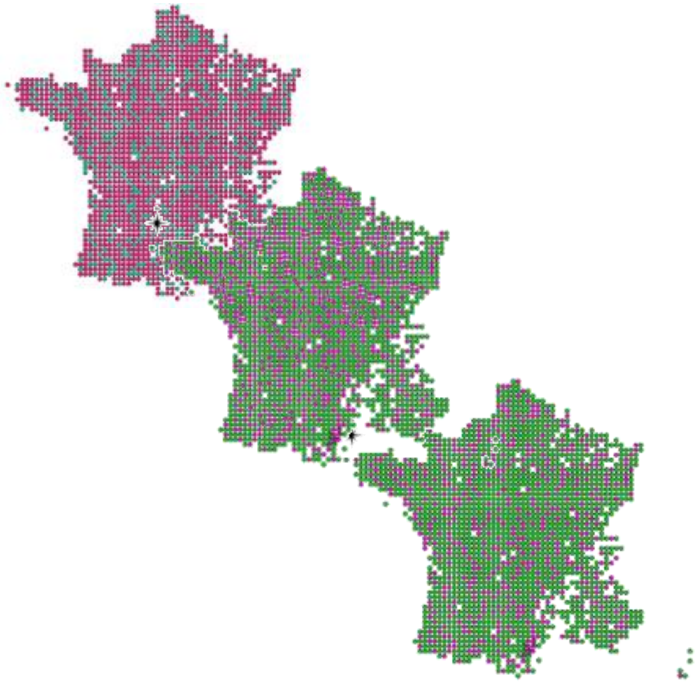


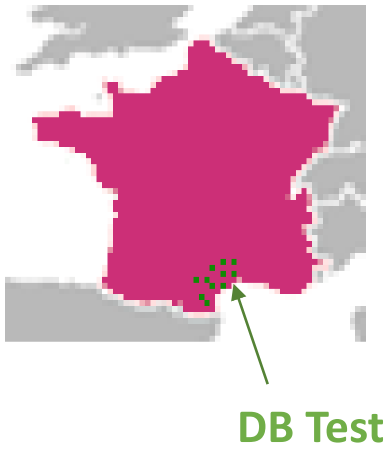
Can we use the French spectral library for soil properties estimation *at Regional Scale (Outside of France)*?

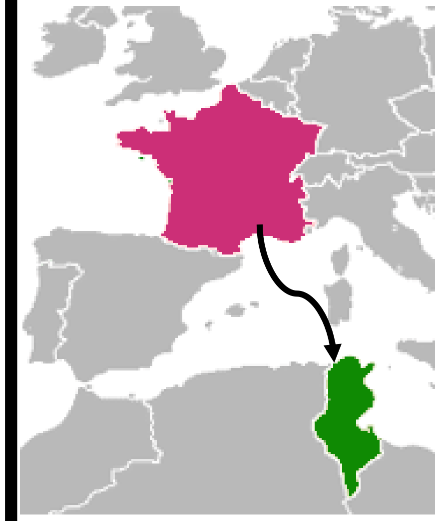
Can we use the French spectral library for soil properties estimation *at Regional Scale (Outside of France)*?

=> Scenario 4

	1st scenario	2nd scenario	3rd scenario	4th scenario
Area coverage	Calib DB = Test DB	Calib DB = Test DB	Test DB ϵ Calib DB	Calib DB ≠ Test DB
Database sizes	Calib DB > Test DB	Nb Calib DB = 100% - Nb Calib DB	Calib DB >> Test DB	DB Calib >> DB Test

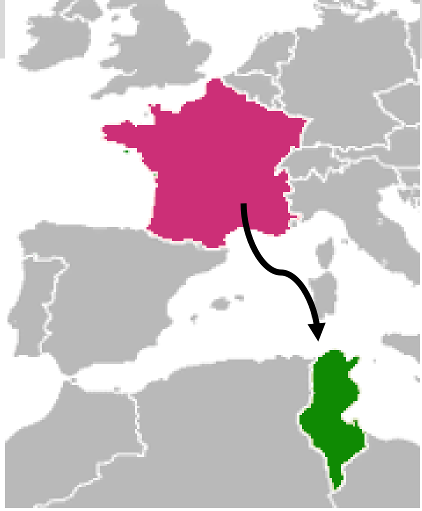




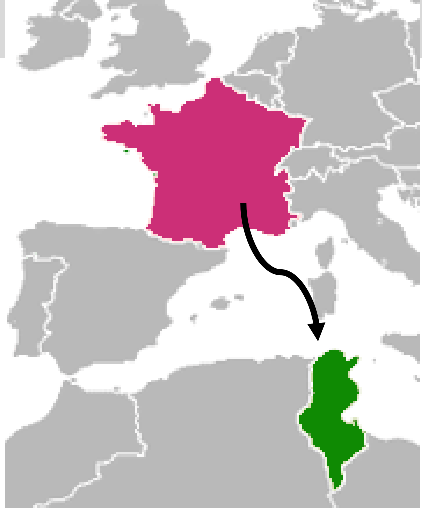


Questions

- Which performances can we expect when the validation samples come from a region OUTSIDE of the territory covered by the Calibration samples?
- Which approach is appropriate in this specific situation?



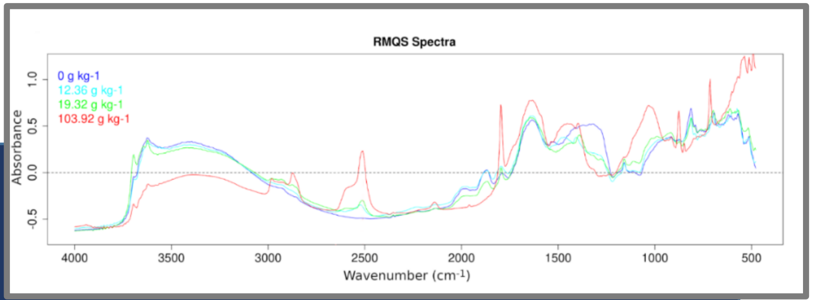
Gomez et al., 2020. Geoderma



Gomez et al., 2020. Geoderma

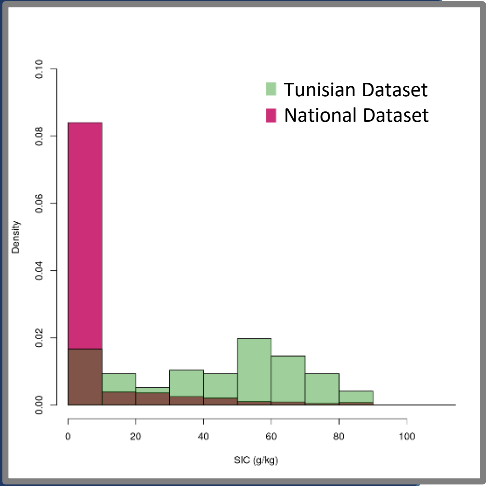
Questions

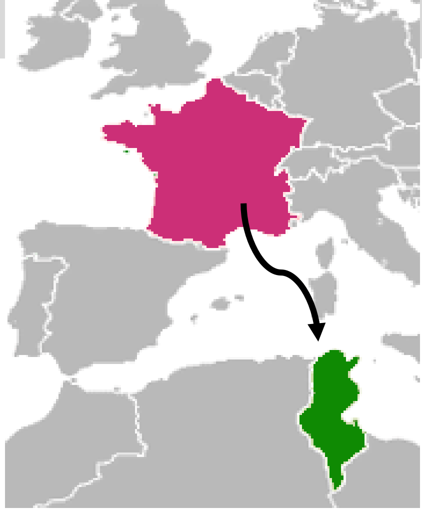
- Which performances can we expect when the validation samples come from a region OUTSIDE of the territory covered by the Calibration samples?
- Which approach is appropriate in this specific situation?



Data

- ✓ 2178 soil samples of RMQS Collected at 0-30 cm depth
- ✓ 96 soil samples from Tunisia Collected at 0-10 cm depth
- ✓ MIRS spectra
- ✓ Soil Organic Carbon (SOC) and Soil Inorganic Carbon (SIC)



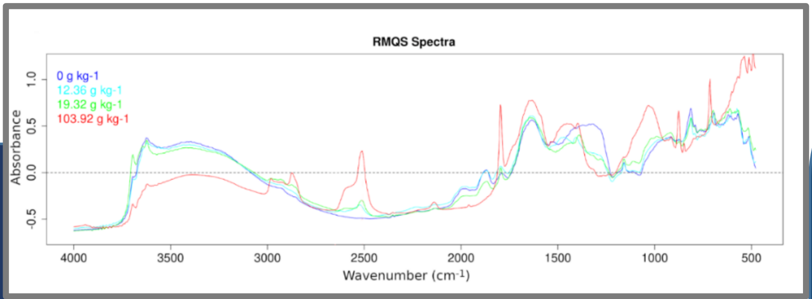


Gomez et al., 2020. Geoderma

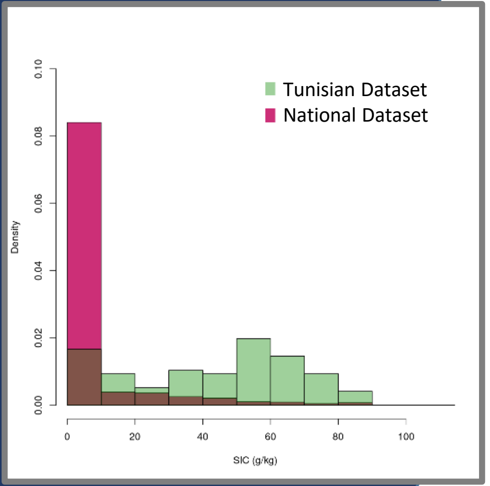
Questions

- ➢ Which performances can we expect when the validation samples come from a region OUTSIDE of the territory covered by the Calibration samples?
- ➢ Which approach is appropriate in this specific situation?

Data

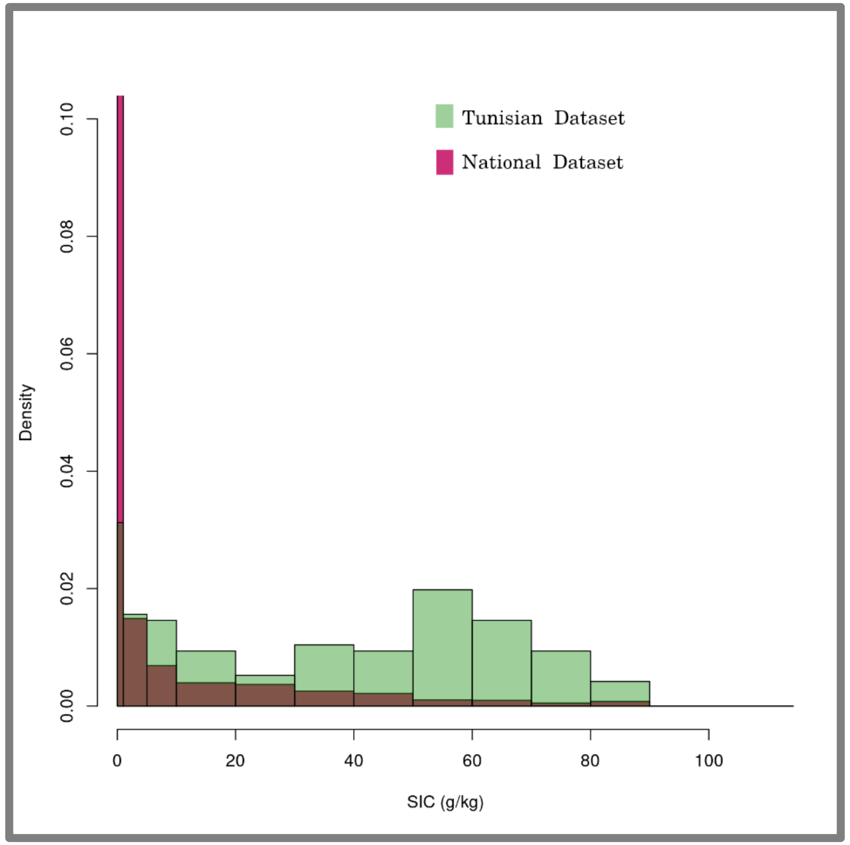


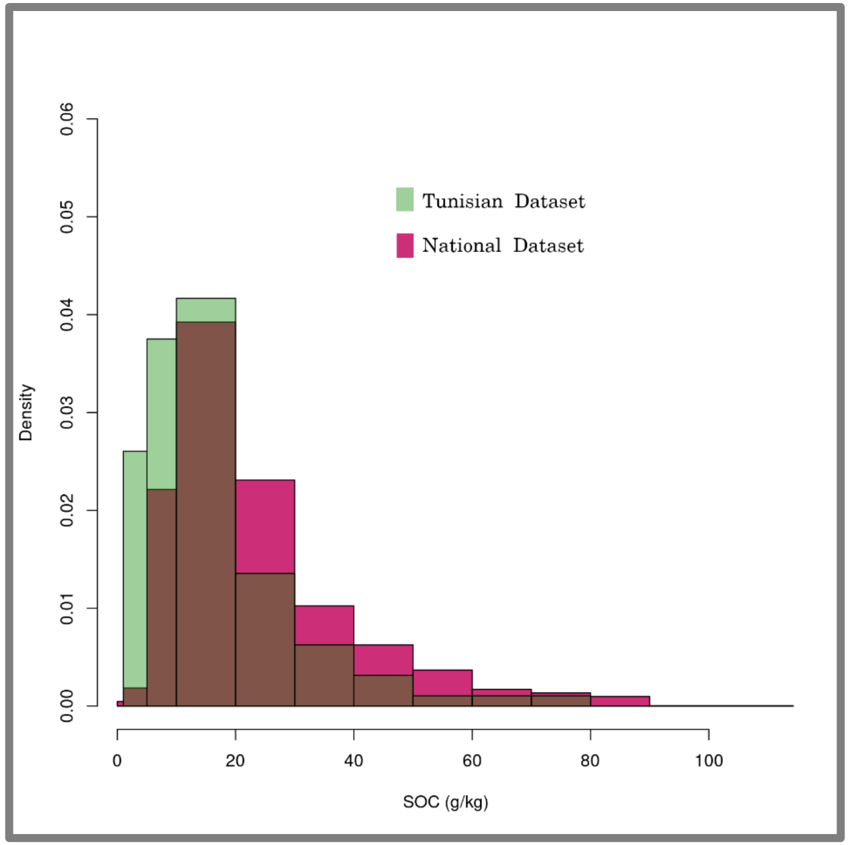
- ✓ 2178 soil samples of RMQS Collected at 0-30 cm depth
- ✓ 96 soil samples from Tunisia Collected at 0-10 cm depth
- ✓ MIRS spectra
- ✓ Soil Organic Carbon (SOC) and Soil Inorganic Carbon (SIC)

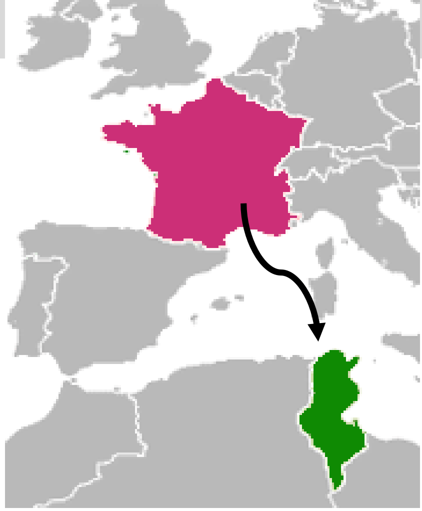


Methods

- ✓ PLSR to build regression models
- ✓ RMQS dataset divided in 75% for Calibration and 25 % for Validation
  (same SOC and SIC distributions in both)
- ✓ 2 Tested Approaches. Calibration from
  - ❖ PLSR using all calibration soil samples of RMQS
  - ❖ PLSR using spectral neighbours of calibration dataset
- ✓ Test using the 96 Tunisian soil samples



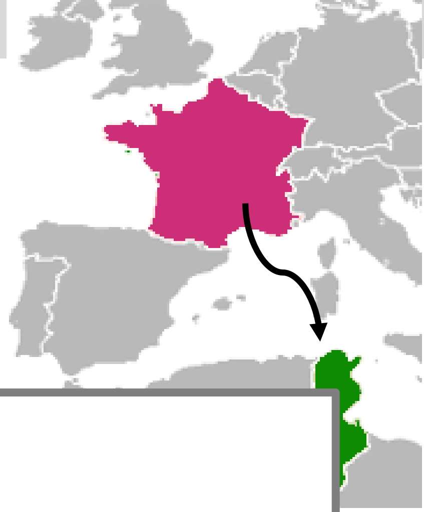


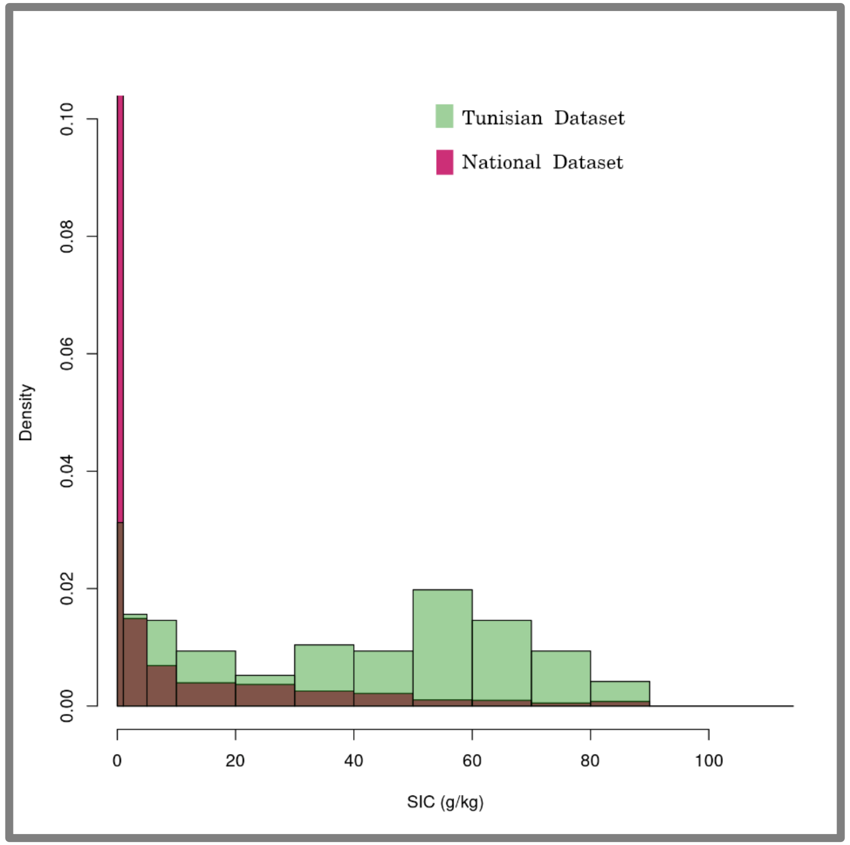


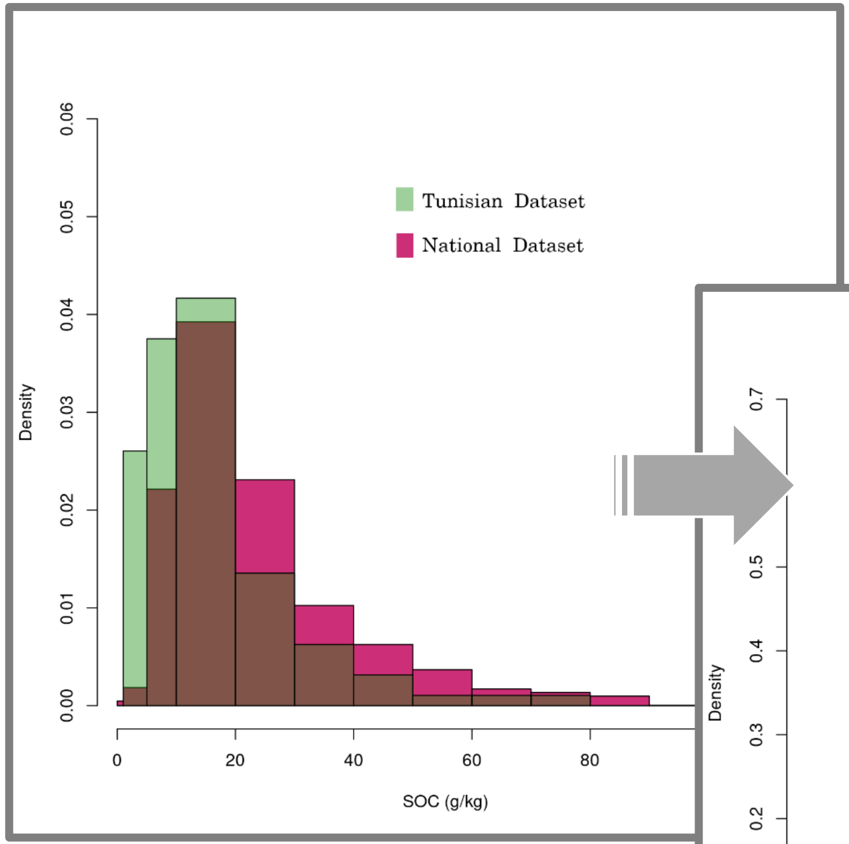
Similar SIC ranges but different SIC distributions

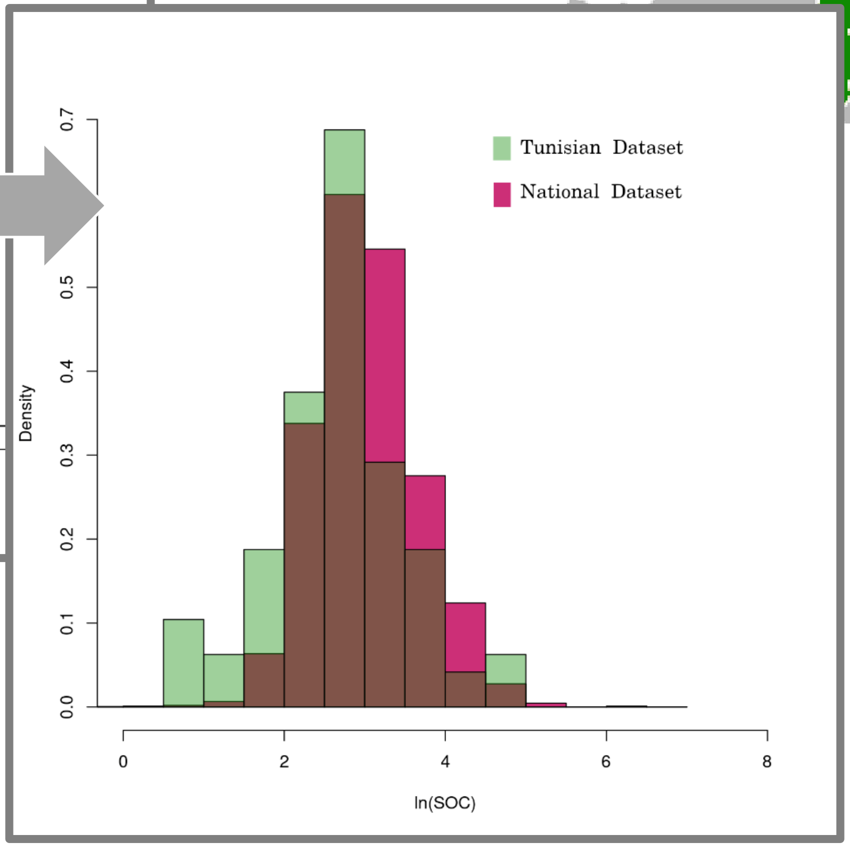
Similar SOC ranges and distributions

=> non-normal distributions of SOC and SIC contents for RMQS samples







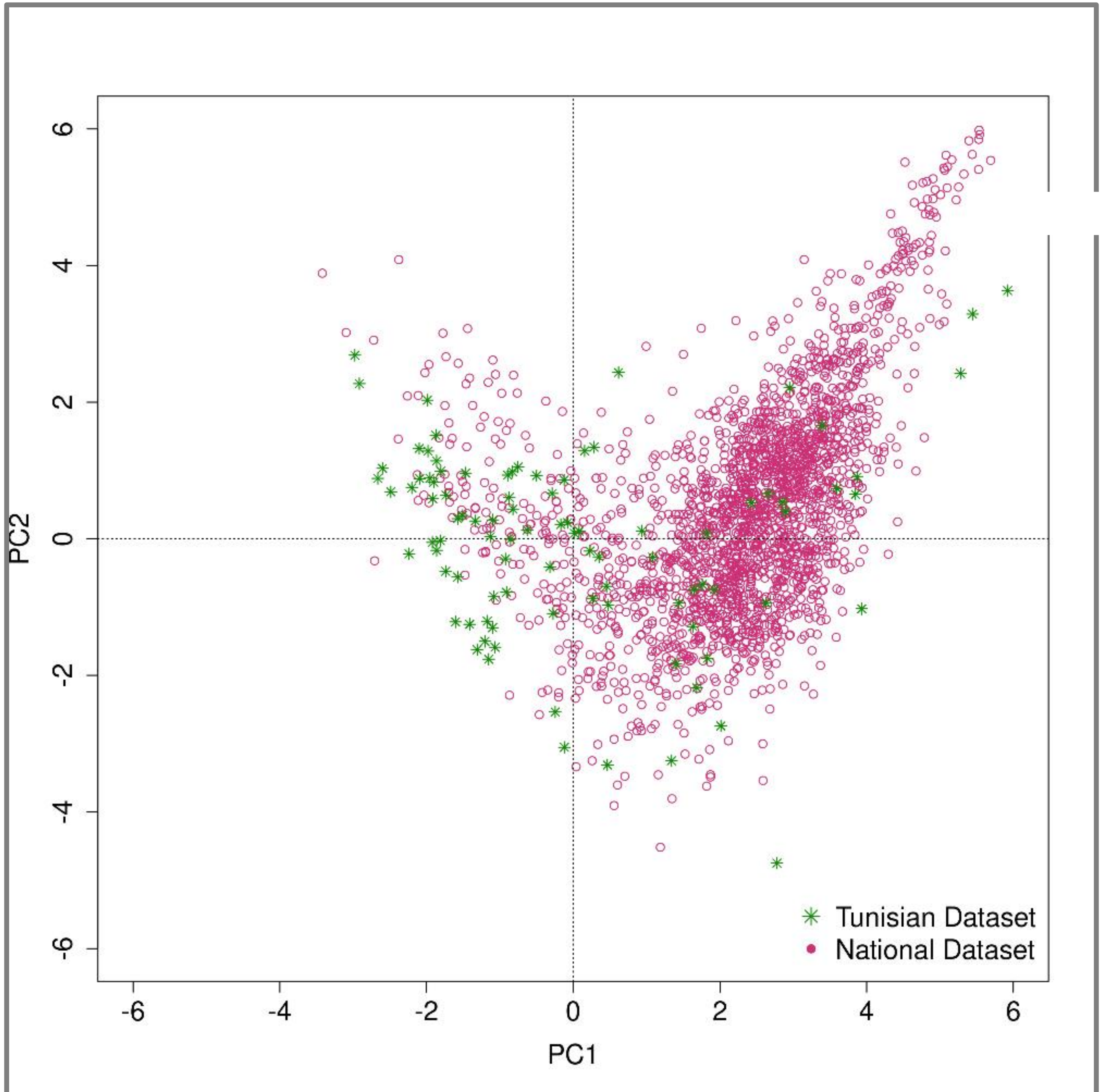


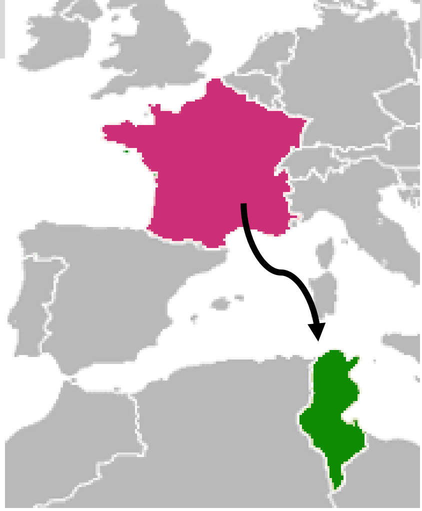
Similar SIC ranges but different SIC distributions

Similar SOC ranges and distributions

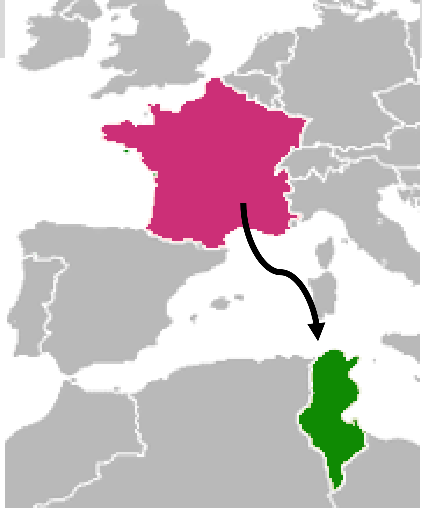
⇒ Transformation to normal distribution of SOC content of RMQS samples

(*Jaconi et al., 2019*)

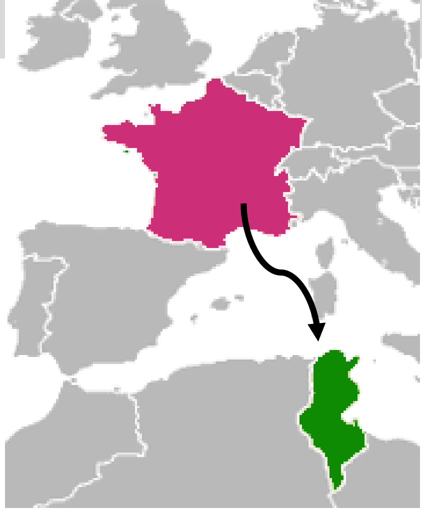

PC2
PC1
6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6
Tunisian Dataset
National Dataset



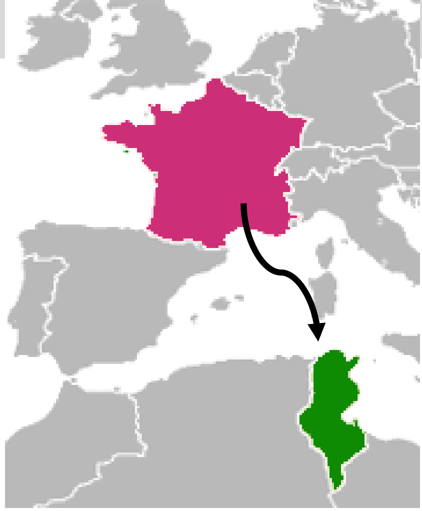
Can we use the French spectral library for soil properties estimation at Regional Scale (**Outside of France**)? Scenario 4



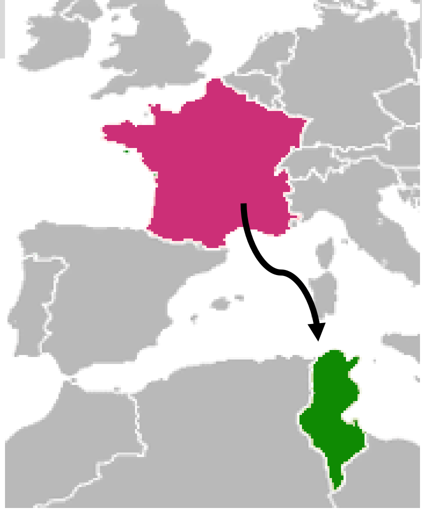
	R^2_{val}	$RMSE_{val}$ g kg[-1]	RPD_{val}	R^2_{test}	$RMSE_{test}$ g kg[-1]	RPD_{test}
Global-Model for SIC predictions	0.98	2.1	7.6			
Global-Model for SOC predictions	0.88	7.2	2.7	0.64	16.0	1.3
Global-Model for ln(SOC) predictions*	0.90	6.6	2.9			



	R^2_{val}	$RMSE_{val}$ g kg-1	RPD_{val}	R^2_{test}	$RMSE_{test}$ g kg-1	RPD_{test}
Global-Model for SIC predictions	0.98	2.1	7.6	0.96	5.2	4.9
Global-Model for SOC predictions	0.88	7.2	2.7	0.64	16.0	1.3
Global-Model for ln(SOC) predictions*	0.90	6.6	2.9	0.97	4.2	4.9



	R^2_{val}	$RMSE_{val}$ g kg[-1]	RPD_{val}	R^2_{test}	$RMSE_{test}$ g kg[-1]	RPD_{test}
Global-Model for SIC predictions	0.98	2.1	7.6	0.96	5.2	4.9
Local-Model for SIC predictions	0.99	1.8	8.8	0.96	5.6	4.6
Global-Model for SOC predictions	0.88	7.2	2.7	0.64	16.0	1.3
Local-Model for SOC predictions	0.93	5.4	3.6	0.89	6.9	3.0
Global-Model for ln(SOC) predictions*	0.90	6.6	2.9	**0.97**	**4.2**	**4.9**
Local-Model for ln(SOC) predictions*	0.92	5.7	3.4	0.93	5.8	3.6



Gomez et al., 2020. Geoderma

Questions

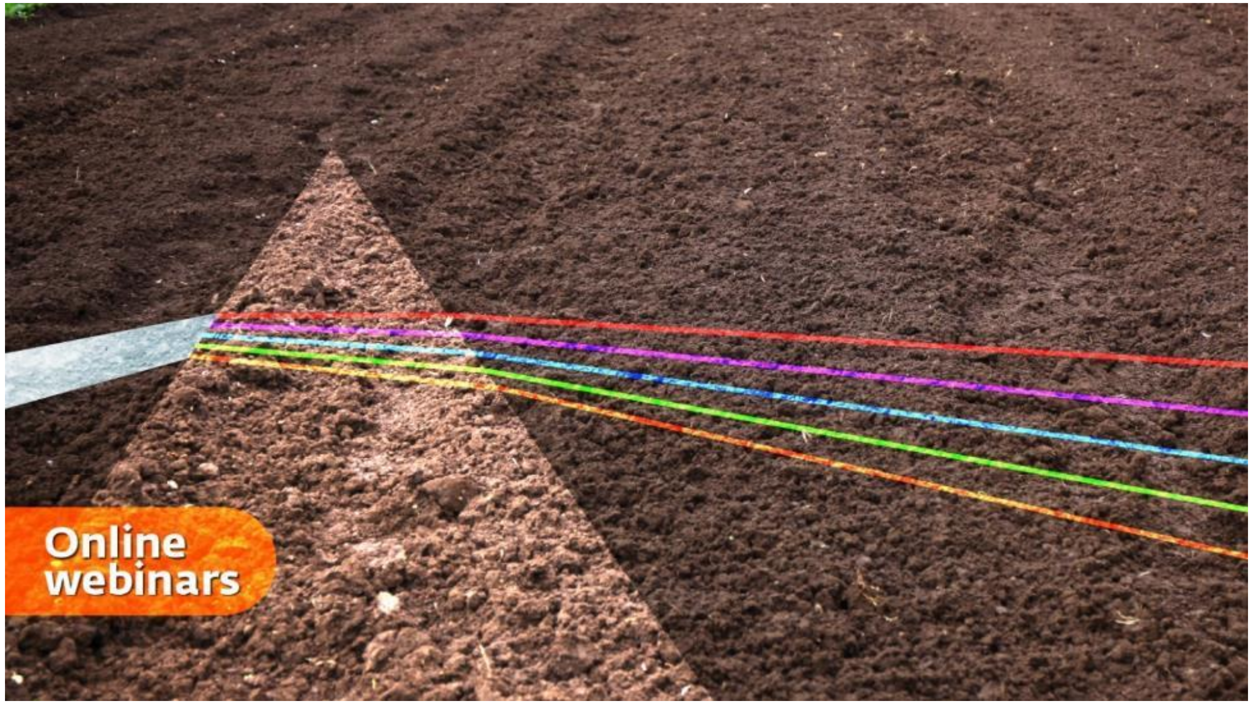
- Which performances can we expect when the validation samples come from a region OUTSIDE of the territory covered by the Calibration samples?
- Which approach is appropriate in this specific situation?

Highlights

- MIRS is a promising tool for SIC determination, even when the calibration and test samples originate from different contexts….

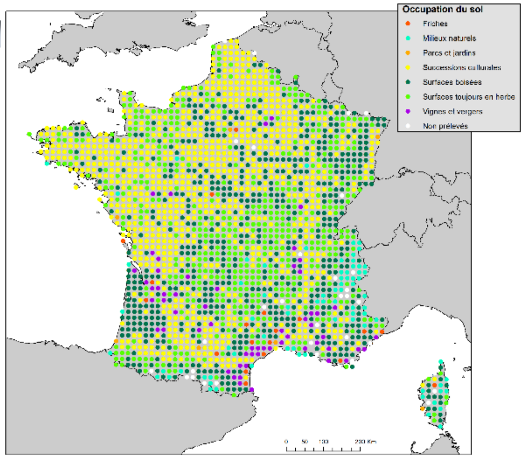
  … while the SOC prediction performance decreases.

- Local calibration significantly improved SOC prediction of Tunisian samples.
- Log-transformation of SOC were more efficient than the local-model approach.



Main Highlights / Recommendations

Main Highlights



Development of a Danish national vis-NIR soil spectral library for soil organic carbon determination

M. Knadel, F. Deng, A. Thomsen, M.H

Dept. of Agroecology, Faculty of Agricultural Sciences, Aa

Blichers Alle 20, PO box 50, DK-8830 Tjele

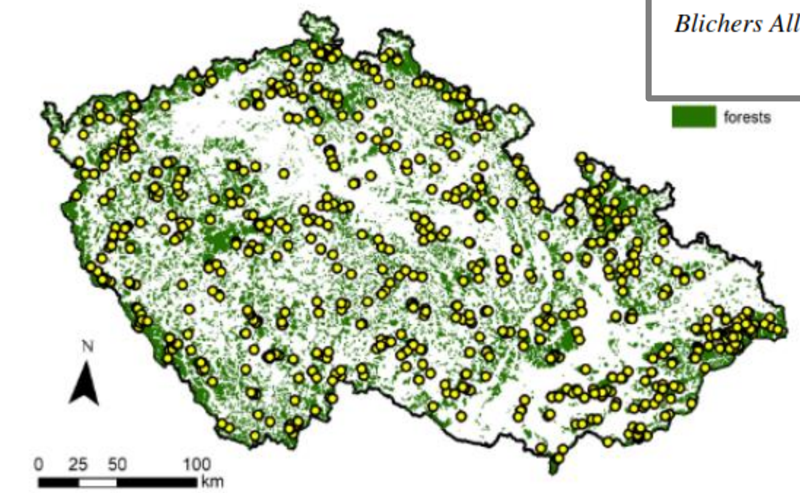


Fig. 1. Location of Czech forests and soil sampling sites.

ELSEVIER
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

National-scale spectroscopic assessment of soil organic carbon in forests of the Czech Republic

Asa Gholizadeh[a,*], Raphael A. Viscarra Rossel[b], Mohammadmehdi Saberioon[c], Luboš Borůvka[a], Josef Kratina[a], Lenka Pavlů[a]

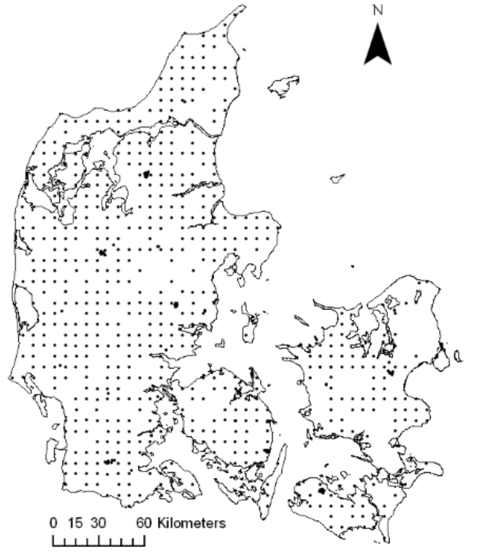


Figure 1. Distribution of soil profiles across Denmark.

SOIL, 7, 52
https://doi.org/10.5194/soil-7-525-2021

Figure 1. Map of Switzerland with sampling locations of the mid-IR spectral library, including the sites of the Biodiversity Monitoring program (BDM; 6 × 4 km; $n = 1094$) and the National Soil Monitoring Network (NABO; $n = 71$). In total, 71 NABO sites (10 m × 10 m) were sampled with a grid-based stratified design, and 1094 BDM samples were obtained from single sampling events. The NABO sites have been continuously sampled and measured in 5-year intervals since 1985.

Developing the Swiss mid-infrared soil spectral library for local estimation and monitoring

Philipp Baumann[1,2], Anatol Helfenstein[1,3], Andreas Gubler[4], Armin Keller[2], Reto Giulio Meuli[4], Daniel Wächter[4], Juhwan Lee[5], Raphael Viscarra Rossel[6], and Johan Six[1]

Main Highlights

- ➢ Current French National Dataset : ~3800 soil samples (collected every 16 km at two depth layers when feasible : 0-30 and 30-50 cm) associated to Vis-NIR, NIR, MIR spectra, soil attributes, soil biodiversity indicators…

- ➢ Future French National Dataset (collected every 16 km at four depth layers when feasible) :

> 7000 soil samples in 2030, > 11 000 in 2045, and > 15 000 in 2060

⇒ *still associated to soil attributes*

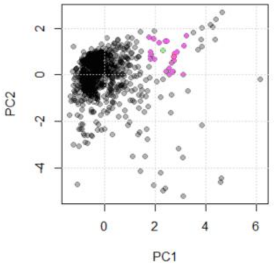
⇒ *NIR and/or MIR spectra would be also acquired*

- ➢ French National Dataset may provide soil properties predictions at both national and local scale for large-scale soil inventories and mapping studies

- ➢ Models using spectral neighbors (Local-PLSR) seem to be more appropriate than Global Models when using the French National Dataset

Main Highlights / Recommendations

- Have a look on test data (Spectra). Overlap Calibration data?



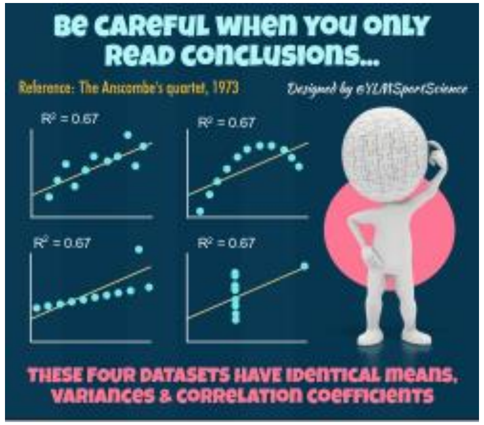
- Knowledge on pedological context of the test area. Same context than Calibration dataset?

- Define the aim of the soil properties predictions. Pre-classification of soils? Spatial pattern analysis? Monitoring?

  $\Rightarrow$ so define the acceptable SEP for your study

- As explained by Budiman in a previous webinar, be careful to the R^2 !

@Webinar Budiman Minasny

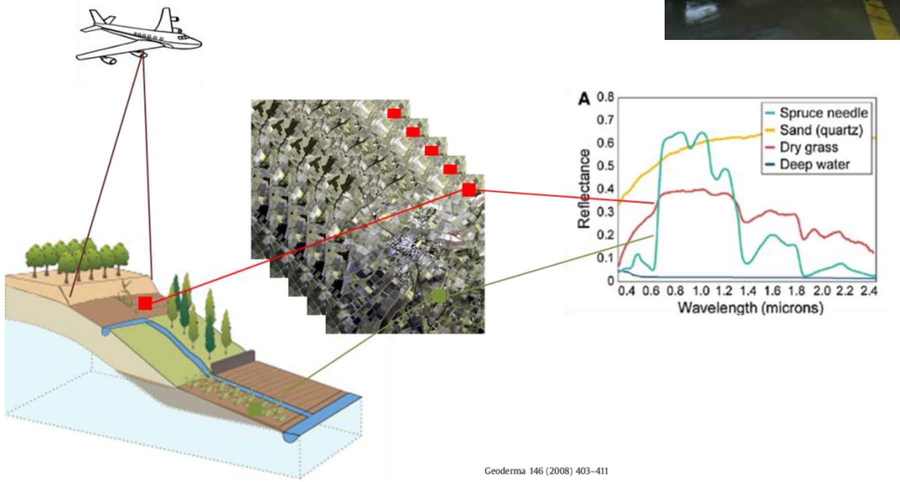


Main Highlights / Recommendations

- Remote sensing sensors acquire data in Vis-NIR range…

  … and provide a synoptic view of the topsoil and a possible repetitive coverage

- Needs to go further in Vis-NIR developments to get benefit from the high quality of scattered Lab Vis-NIR spectra and the high quantity of remote sensing spectra



Geoderma 146 (2008) 403–411
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma
ELSEVIER

Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study

Cécile Gomez [a,b,*], Raphael A. Viscarra Rossel [b], Alex B. McBratney [b]
[a] IRD, Laboratoire d'étude des Interactions Sols-Agrosystèmes-Hydrosystèmes (LISAH), Campus AGRO-Bat.24, 34060 Montpellier, France
[b] Australian Centre for Precision Agriculture, Food and Natural Resources, McMillan Building, A05, The University of Sydney, NSW 2006, Australia

Geoderma xxx (2010) xxx–xxx
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma
ELSEVIER

Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy

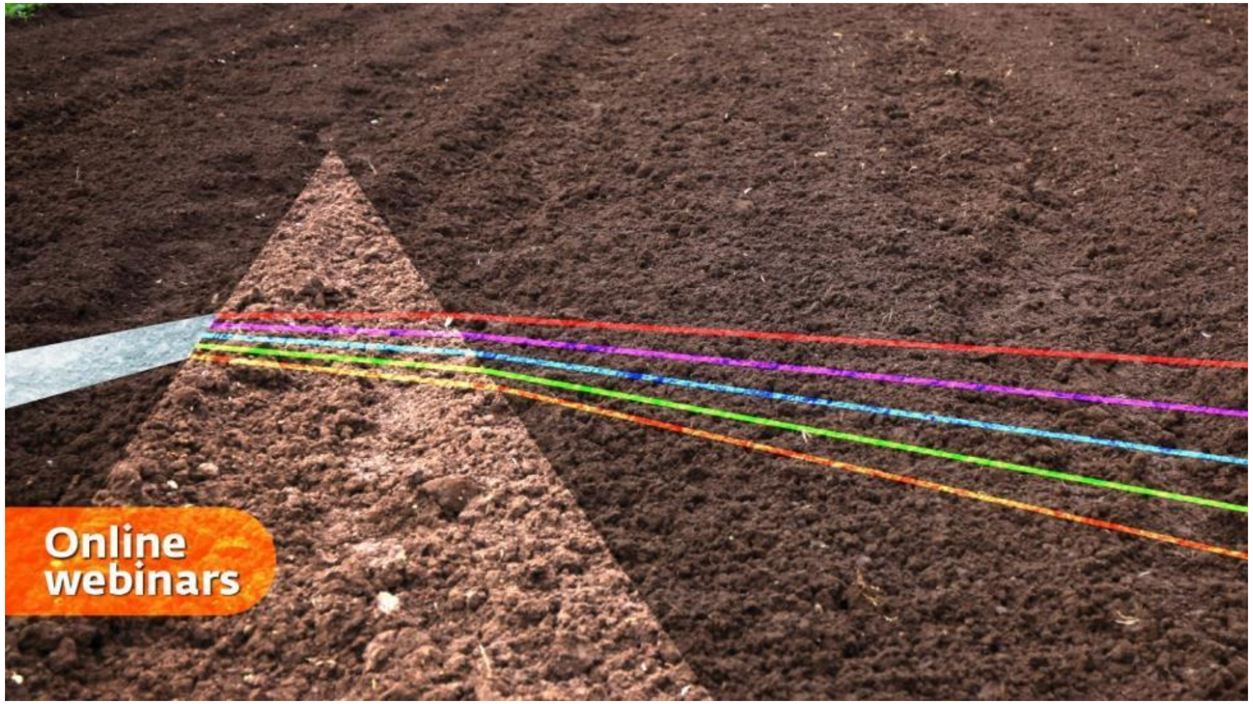
Antoine Stevens [a,*], Thomas Udelhoven [b], Antoine Denis [c], Bernard Tychon [c], Rocco Lioy [d], Lucien Hoffmann [b], Bas van Wesemael [a]

ELSEVIER
ScienceDirect
Geoderma 136 (2006) 235 – 244
GEODERMA
www.elsevier.com/locate/geoderma

High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures

Thomas Selige [a,*], Jürgen Böhner [b], Urs Schmidhalter [a]
[a] Chair of Plant Nutrition, Department of Plant Sciences, Technical University of Munich, Am Hochanger 2, D-85350 Freising, Germany
[b] Department of Physical Geography, University Göttingen, Goldschmidtstrasse 5, D-37077 Göttingen, Germany

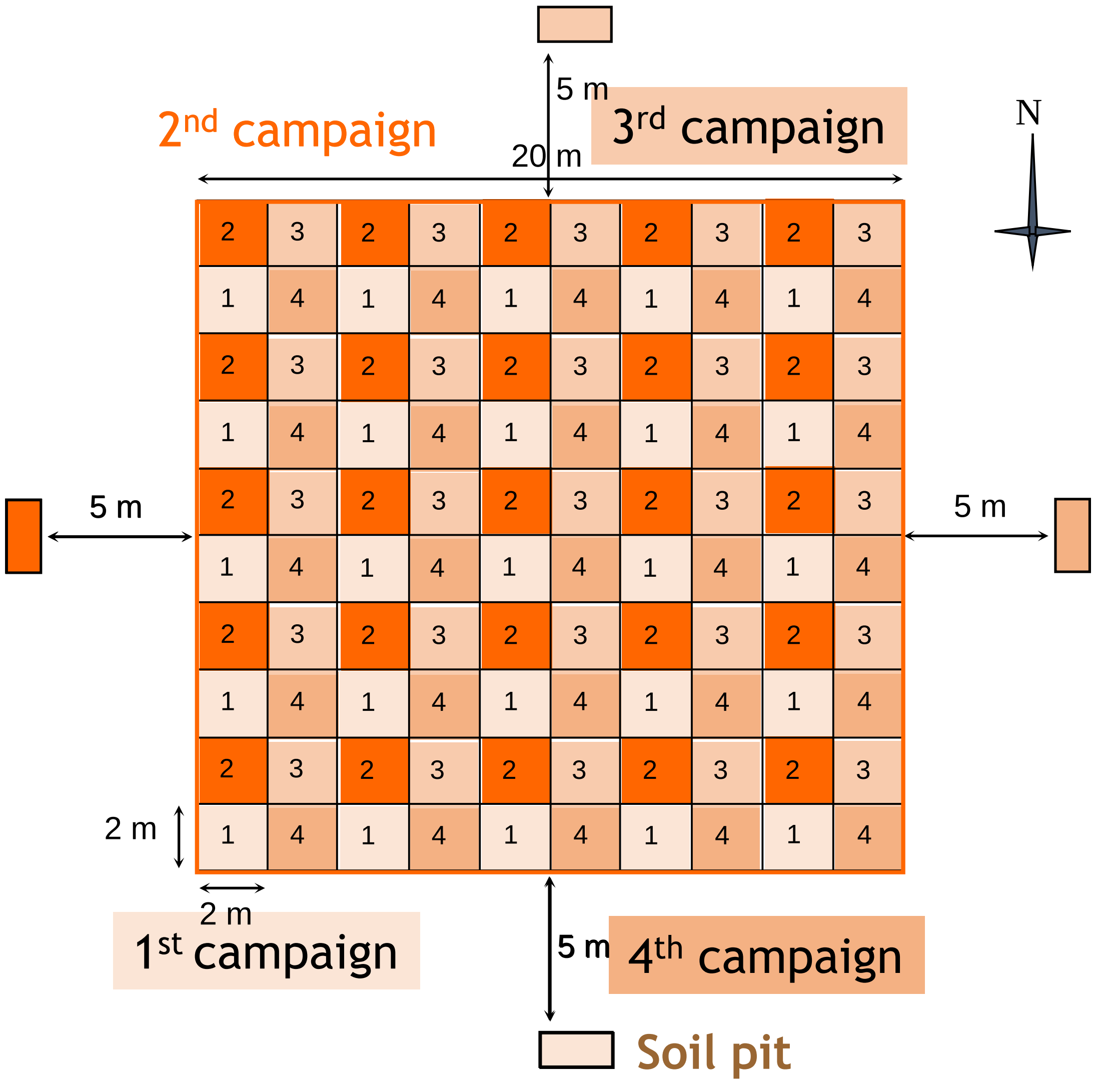
High potential for soil properties mapping (% Clay, SOC, SIC, Iron, …)



Next challenges?

Next challenge?

Could we use these NIR or MIR data to estimate soil properties on soil samples collected during next campaigns?

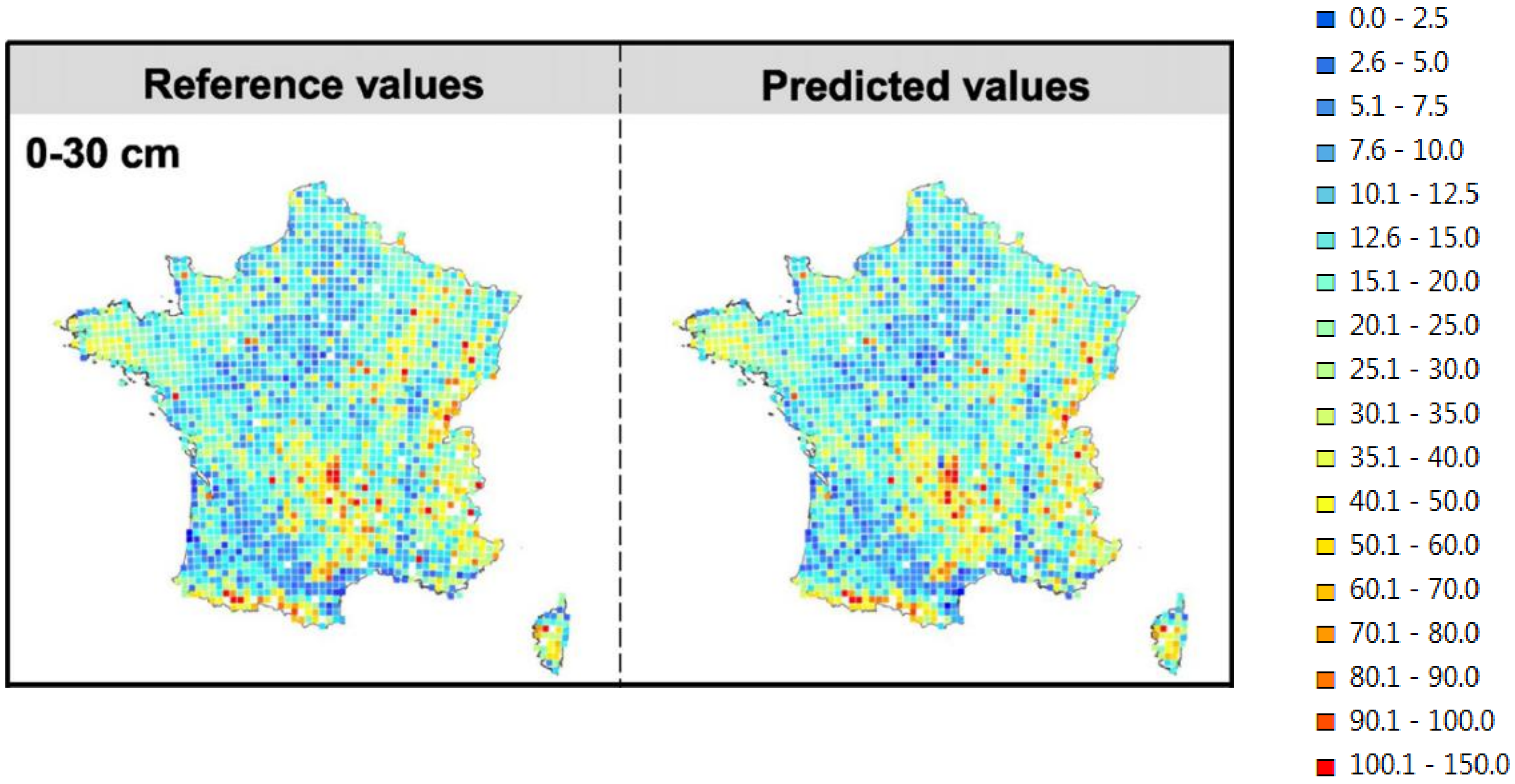


Next challenge?

Regarding the prediction errors, which temporal change of soil properties could we detect?

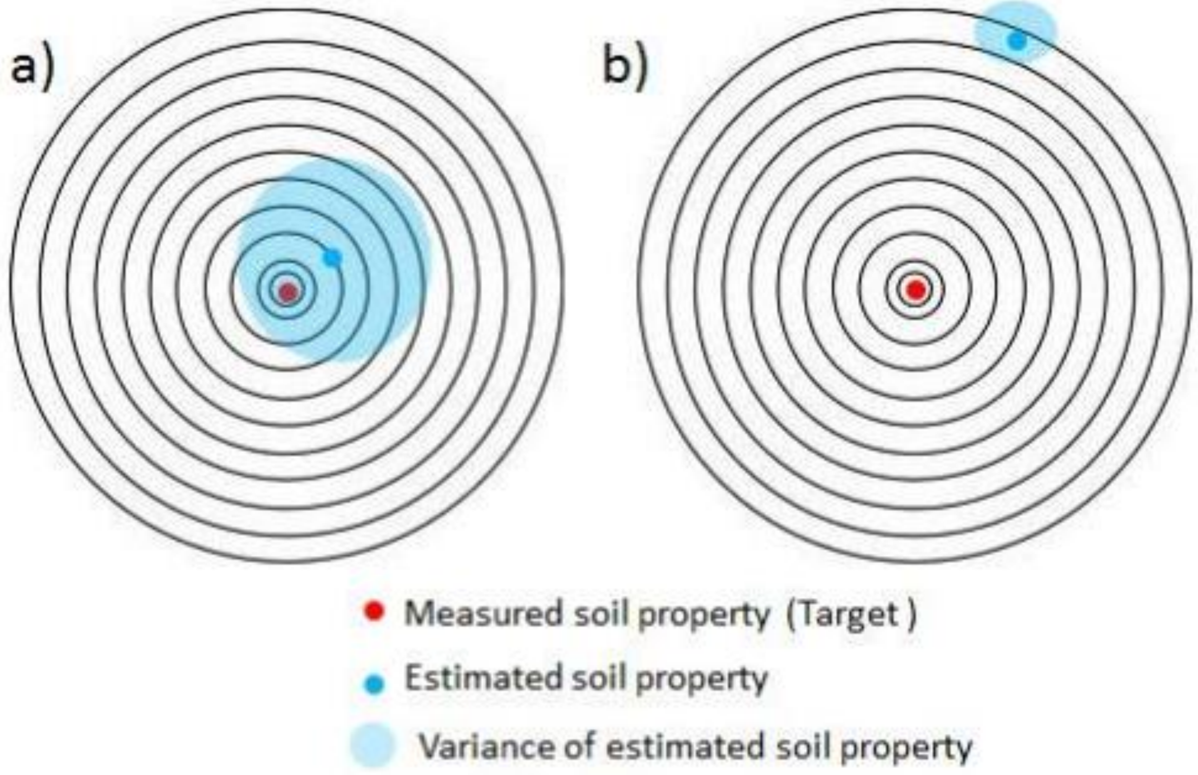
Soil Organic Carbon (g kg^{-1})

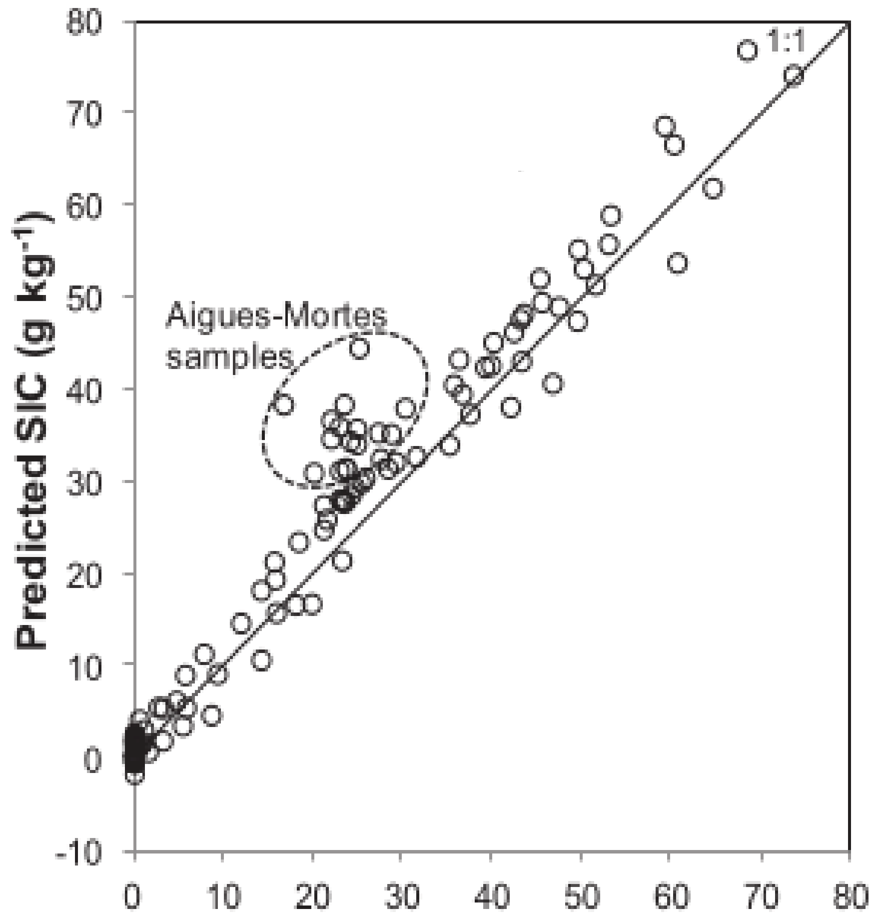
SEP = 4.7 g/kg



Next challenge?

Do we need to estimate an uncertainty and error to each new prediction?





Next challenge?

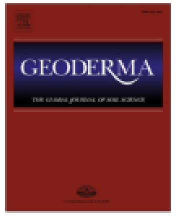
Could we stop testing regression methods, pre-treatments, calibration data selection and fix a protocol to use this kind of dataset?

Geoderma 406 (2022) 115501

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma



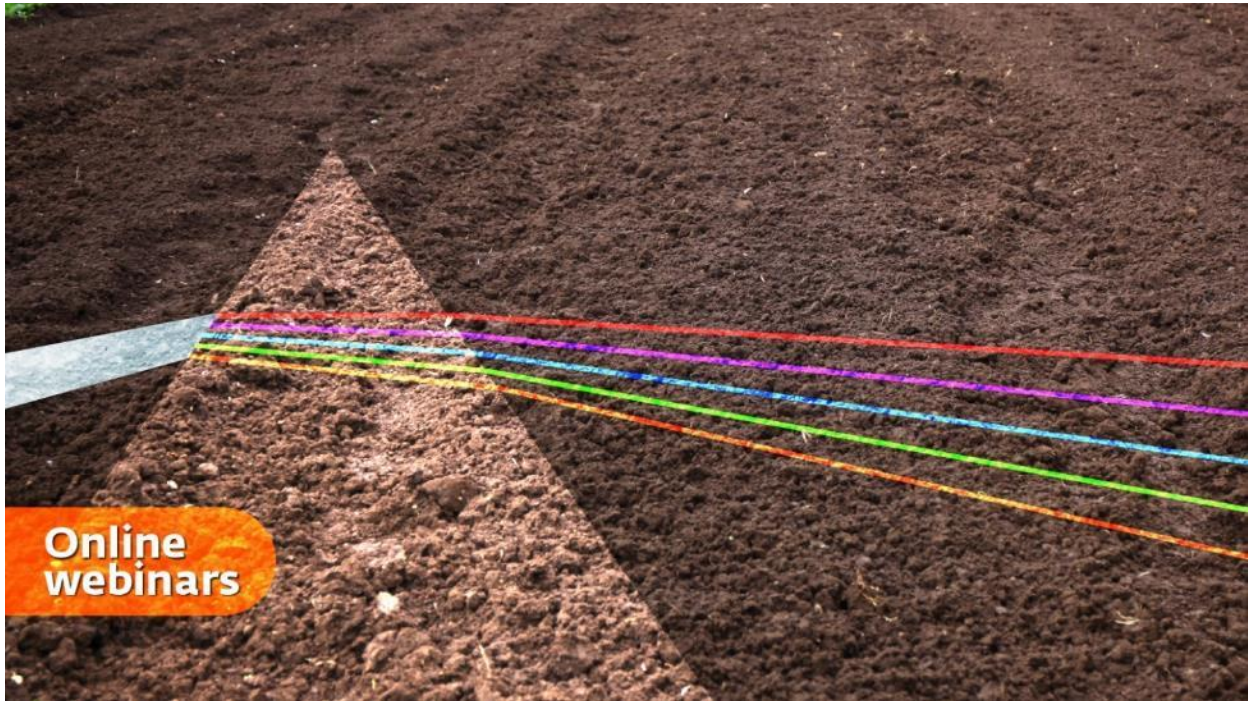
ELSEVIER

Check for updates

To spike or to localize? Strategies to improve the prediction of local soil properties using regional spectral library

Wartini Ng *, Budiman Minasny, Edward Jones, Alex McBratney

School of Life and Environmental Sciences & Sydney Institute of Agriculture, The University of Sydney, NSW, Australia



Acknowledgement

INRAE

IRD

IGN

brgm
Géosciences pour une Terre durable

MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE
Liberté
Égalité
Fraternité

MINISTÈRE
DE L'AGRICULTURE
ET DE L'ALIMENTATION
Liberté
Égalité
Fraternité

ADEME
AGENCE DE LA
TRANSITION
ÉCOLOGIQUE

OFB
OFFICE FRANÇAIS
DE LA BIODIVERSITÉ

Pedosphere 23(4): 422–436, 2013
ISSN 1002-0160/CN 32-1315/P
© 2013 Soil Science Society of China
Published by Elsevier B.V. and Science Press
PEDOSPHERE
www.elsevier.com/locate/pedosphere

Estimation of Soil Carbon Input in France: An Inverse Modelling Approach*[1]

J. MEERSMANS[1,*2], M. P. MARTIN[1], E. LACARCE[1], T. G. ORTON[1], S. DE BAETS[2,4], M. GOURRAT[1], N. P. A. SABY[1], J. WETTERLIND[1], A. BISPO[3], T. A. QUINE[4] and D. ARROUAYS[1]

Geoderma Regional 23 (2020) e00337
Contents lists available at ScienceDirect
Geoderma Regional
journal homepage: www.elsevier.com/locate/geodrs

Impacts of national scale digital soil mapping programs in France

Dominique Arrouays [a,*], Anne C. Richer-de-Forges [a], Floren
Songchao Chen [a,c], Manuel P. Martin [a], Mercedes Román D
Bertrand Laroche [a], Thomas Loiseau [a], Isabelle Cousin [e], M
Christine Le Bas [a], Thomas Eglin [f], Marion Bardy [g], Véroniq
Céline Ratié [a], Antonio Bispo [a]

Geoderma 213 (2014) 133–143
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Statistical sampling design impact on predictive quality of harmonization functions between soil monitoring networks

B.P. Louis *, N.P.A. Saby, T.G. Orton, E. Lacarce, L. Boulonne, C. Jolivet, C. Ratié, D. Arrouays

Geoderma 336 (2019) 81–95
https://doi.org/10.1016/j.geoderma.2018.08.022
Contents lists available
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty

Mercedes Román Dobarco[a,*], Isabelle Cousin[b], Christine Le Bas[a], Manuel P. Martin[a]

Geoderma 381 (2021) 114645
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Developing pedotransfer functions to harmonize extractable soil phosphorus content measured with different methods: A case study across the mainland of France

Bifeng Hu[a,b], Hocine Bourennane[b], Dominique Arrouays[a], Pascal Denoroy[c], Blandine Lemercier[d],
P.A. Saby[a,*]

Science of the Total Environment 441 (2012) 239–247
Contents lists available at SciVerse ScienceDirect
Science of the Total Environment
journal homepage: www.elsevier.com/locate/scitotenv

Assessment of diffuse contamination of agricultural soil by copper in Aquitaine region by using French national databases

Hind El Hadri [a,b,*], Philippe Chéry [b], Stéphanie Jalabert [b], Alexandre Lee [b],
Martine Potin-Gautier [a], Gaëtane Lespes [a]

Geoderma Regional 7 (2016) 137–152
Contents lists available at ScienceDirect
Geoderma Regional
journal homepage: www.elsevier.com/locat

Evaluating large-extent spatial modeling approaches: A case study for soil depth for France

M. Lacoste [a,*], V.L. Mulder [b], A.C. Richer-de-Forges [b], M.P. Martin [b], D. Arrouays [b]

Chemosphere 181 (2017) 635–644
Contents lists available at ScienceDirect
Chemosphere
journal homepage: www.elsevier.com/locate/chemosphere

A survey of topsoil arsenic and mercury concentrations across France

B.P. Marchant [a,*], N.P.A. Saby [b], D. Arrouays [b]
[a] British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK
[b] INRA, US1106 Unité Infosol, Orléans, France

Geoderma 331 (2018) 70–80
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Prediction of total silicon concentrations in French soils using pedotransfer functions from mid-infrared spectrum and pedological attributes

A. Landré[a], N.P.A. Saby[a,*], B.G. Barthès[b], C. Ratié[a], A. Guerin[c], A. Etayo[c], B. Minasny[d], M. Bardy[a], J.-D. Meunier[e], S. Cornu[e]

Pedosphere 23(4): 422–436, 2013
ISSN 1002-0160/CN 32-1315/P
© 2013 Soil Science Society of China
Published by Elsevier B.V. and Science Press
PEDOSPHERE
www.elsevier.com/locate/pedosphere

n Inverse

S. DE BAETS[2,4], M. GOURRAT[1],
ARROUAYS[1]

Geoderma 213 (2014) 133–143
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

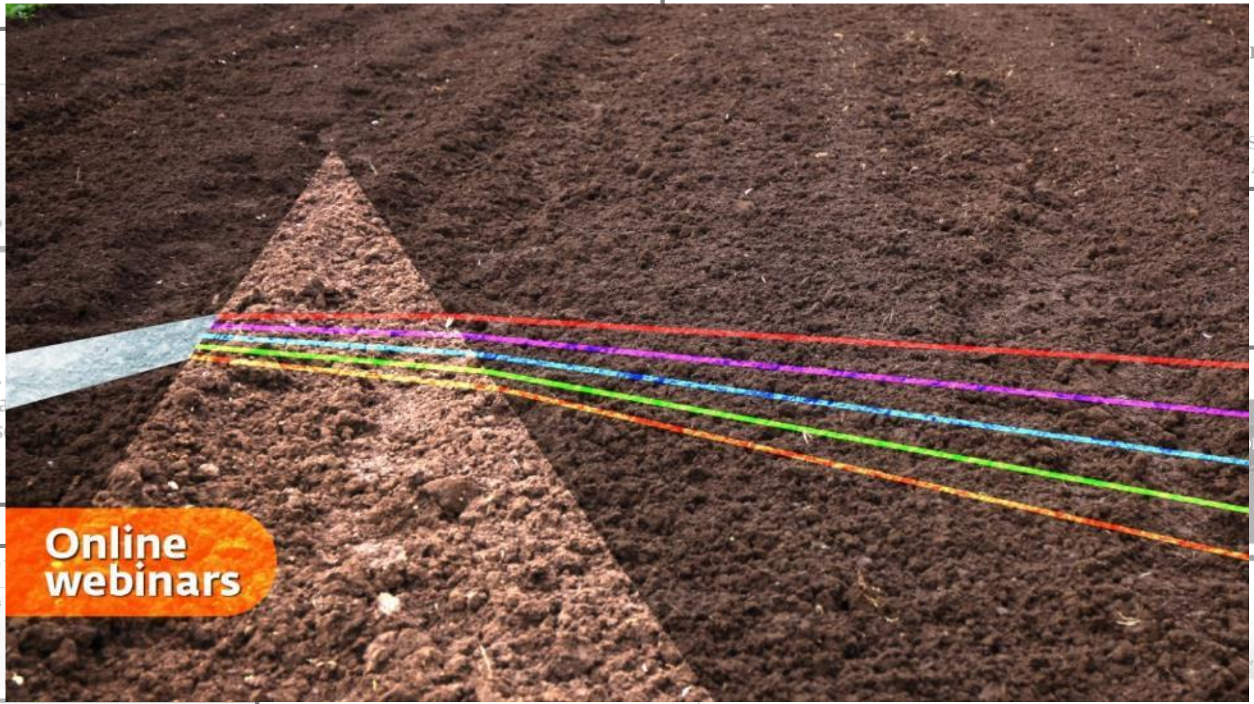
Statistical sampling design impact on predictive quality of harmonization functions between soil monitoring networks

B.P. Louis *, N.P.A. Saby, T.G. Orton, E. Lacarce, L. Boulonne, C. Jolivet, C. Ratié, D. Arrouays

ELSEVIER
jo

Impacts of national scale

Dominique Arrouays [a,*], Anne C.
Songchao Chen [a,c], Manuel P. Ma
Bertrand Laroche [a], Thomas Lois
Christine Le Bas [a], Thomas Eglin
Céline Ratié [a], Antonio Bispo [a]



Mercedes Román Dobarco[a,*], Isabelle Cousin[b], Christine Le Bas[a], Manuel P. Martin[a]

Geoderma 381 (2021) 114645
Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Developing pedotransfer functions to harmonize extractable soil phosphorus content measured with different methods: A case study across the mainland

Bifeng Hu[a,b], Hocine Bourennane[b], Dominique Arrouays[a], Pascal Denoroy[c], Blandine Lemercier[d],

Science of the
Contents lists
Science of
journal homepage:

Assessment of diffuse contamination
by using French national databases

Hind El Hadri [a,b,*], Philippe Chéry [b], Stéphanie
Martine Potin-Gautier [a], Gaëtane Lespes [a]

THANKS FOR YOUR ATTENTION

Contents lists available at ScienceDirect
Geoderma Regional
journal homepage: www.elsevier.com/locat

Evaluating large-extent spatial modeling approaches: A case study for soil depth for France

M. Lacoste [a,*], V.L. Mulder [b], A.C. Richer-de-Forges [b], M.P. Martin [b], D. Arrouays [b]

Chemosphere
journal homepage: www.elsevier.com/locate/chemosphere

A survey of topsoil arsenic and mercury concentrations across France

B.P. Marchant [a,*], N.P.A. Saby [b], D. Arrouays [b]
[a] British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK
[b] INRA, US1106 Unité Infosol, Orléans, France

Contents lists available at ScienceDirect
Geoderma
journal homepage: www.elsevier.com/locate/geoderma

Prediction of total silicon concentrations in French soils using pedotransfer functions from mid-infrared spectrum and pedological attributes

A. Landré[a], N.P.A. Saby[a,*], B.G. Barthès[b], C. Ratié[a], A. Guerin[c], A. Etayo[c], B. Minasny[d], M. Bardy[a],
J.-D. Meunier[e], S. Cornu[e]