

# Caractérisation des sols par SPIR : connaissance des erreurs de mesures et qualité de la prédiction des concentrations en carbone organique et en azote total

Othmane Aoultit<sup>1</sup>, Cyril Girardin<sup>2</sup>, Eric Latrille<sup>3</sup>, Marie-Noël Mistou<sup>4</sup>, Virginie Rossard<sup>3</sup>, Maud Seger<sup>5</sup>

<sup>1</sup> LAS, INRAE

<sup>2</sup> ECOSYS, Paris-Saclay, INRAE

<sup>2</sup> LBE, Université Montpellier, INRAE

<sup>2</sup> Agronomie, Paris-Saclay, INRAE

<sup>2</sup> InfoSol, INRAE

## ➤ Contexte

La caractérisation des sols au niveau agronomique nécessite la détermination des teneurs en carbone et en azote.

La teneur en carbone dans le sol est composée de deux grandeurs : la teneur en carbone organique, utile pour le développement des cultures et la teneur inorganique ou minérale, utile à la structure physique du sol.

Nous nous intéressons à caractériser par spectrométrie proche infrarouge les sols agricoles ayant opérés une transition agroécologique et à prédire les teneurs en carbone organique Corg et l'azote total à l'aide de modèles de calibration.



## ➤ Contexte

Depuis 20 ans, plusieurs systèmes de culture innovants en rupture ont été conçus et testés à l'INRAE dans des « expérimentations systèmes ».

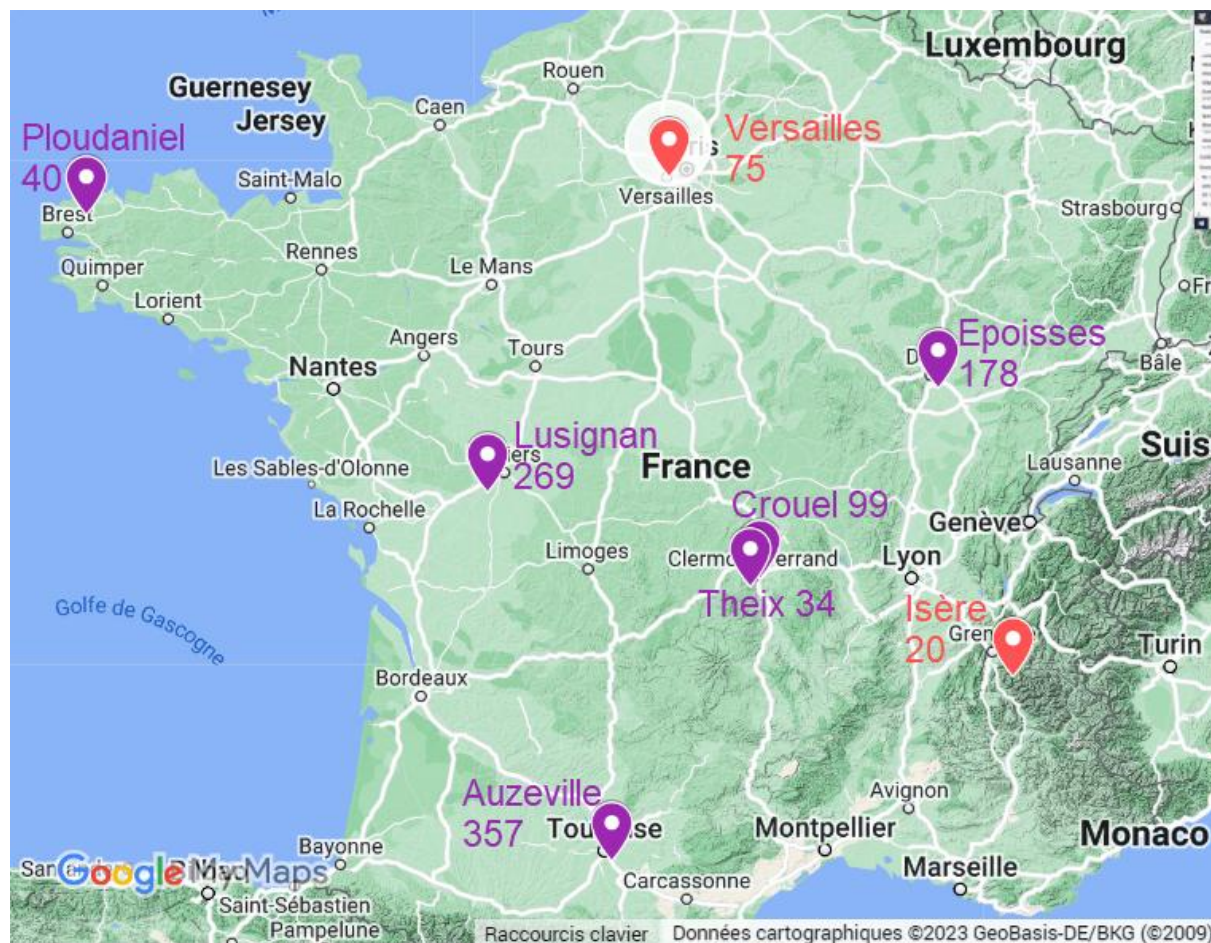
Les échantillons ont été collectés sur 6 sites expérimentaux différents du dispositif CAREX en 2013-2015 et analysés en 2016.

Campagne de caractérisation des parcelles du site d'Epoisses en 2024 après une expérimentation de 10 années en transition agroécologique.



## ➤ Les données : 957 échantillons CAREX en 2013 et 92 d'Epoisses en 2024

| sites      | Nb spectres labo |
|------------|------------------|
| Epoisses   | 178              |
| Auzeville  | 337              |
| Lusignan   | 269              |
| Crouel     | 99               |
| Theix      | 34               |
| Ploudaniel | 40               |



**Campagne 2024  
Epoisses**

92 échantillons avec  
spectres et analyses

# ➤ Les données : protocole de mesures spectrales et analytiques

## ❖ Préparation des échantillons

Selon la norme **NF ISO 11464**

- séchage à l'air (température 35°C)
- émottage et tamisage à 2mm

Avant la prise de spectre

Nouveau séchage à l'air au moins 12h (température 35°C)

- Prise d'essai entre 50 et 100 g d'échantillon dans une clayette puis déposé dans une coupelle

## ❖ Prise de spectre

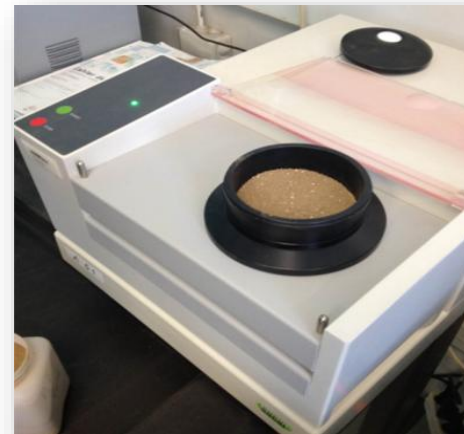
- Büchi NIRFLEX N500. Mode réflectance.
- Acquisition en 4000 et 10000  $\text{cm}^{-1}$  (1000-2500nm)
- Résolution de 8  $\text{cm}^{-1}$  (1501 valeurs)
- Un échantillon de référence est inséré en début et fin de séquence et tous les 10 échantillons « inconnus »



Carbone organique du sol  
(Corg,  $\text{g.kg}^{-1}$ , ISO 14235, NF ISO 10694, 1995).

Carbonate de calcium,  
carbone inorganique ( $\text{g.kg}^{-1}$ ,  
NF ISO 10693, 1995).

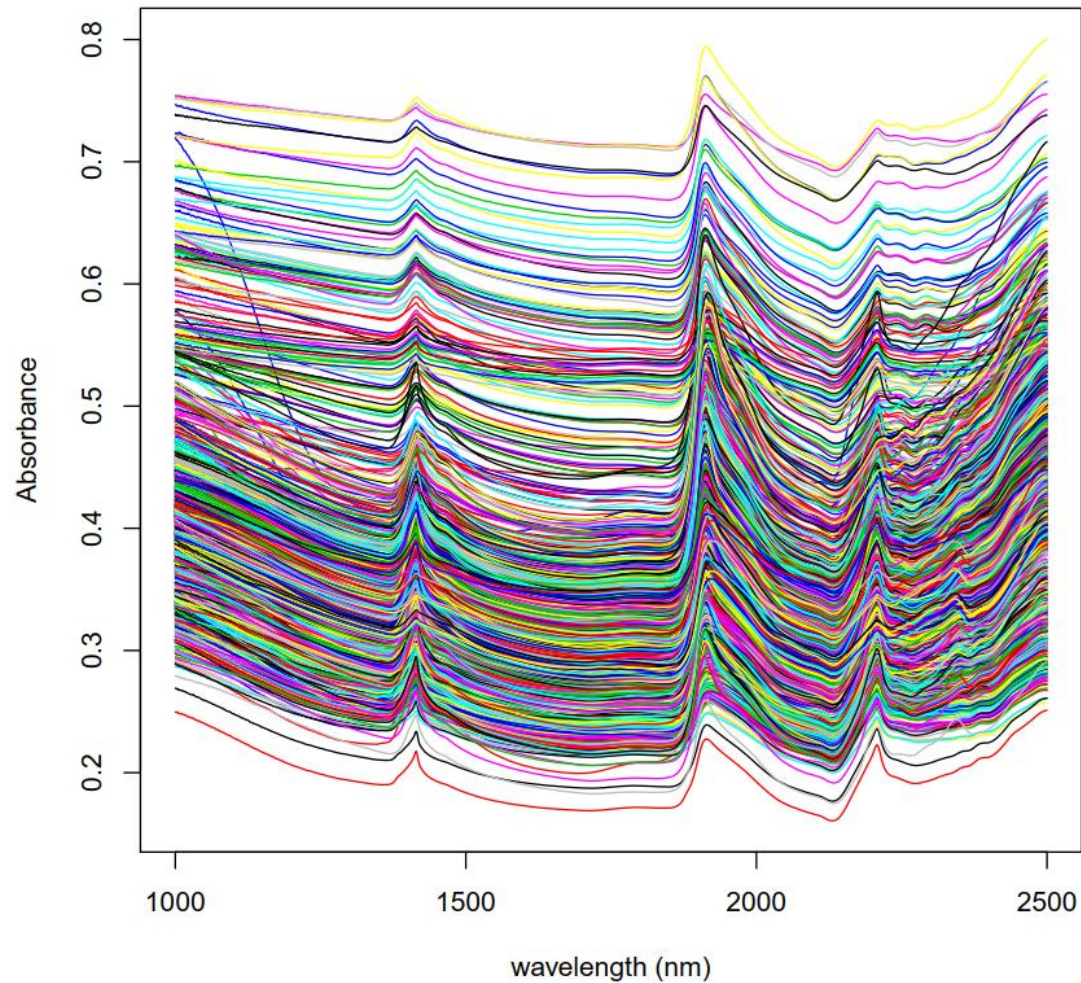
Azote total (Ntot  $\text{g.kg}^{-1}$ , NF ISO 13878, 1998).



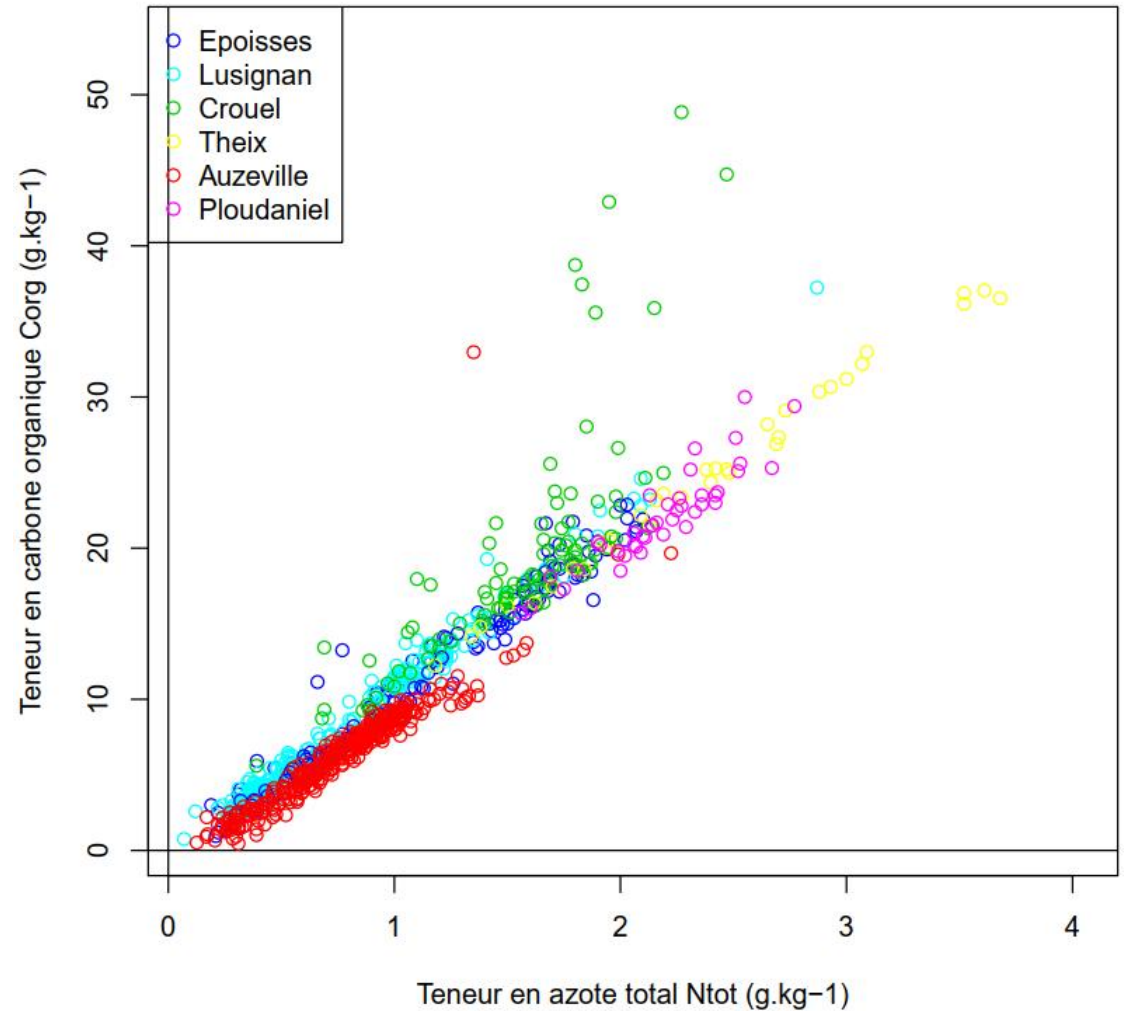
Curation des données :  
observations, analyses  
répétées, calcul de  
Corg/Ntot, suppression de la  
donnée éventuellement.

## ➤ Les données : les mesures spectrales et analytiques

Spectra Plot of CAREX Spectrum (nb 957)



Distributions des teneurs en Corg et Ntot des 957 echantillons



## ➤ Les données : les incertitudes des mesures en laboratoire

**L'incertitude expérimentale de la teneur en carbone organique du sol est variable d'un échantillon à l'autre et n'est pas constante dans le temps.**

Les incertitudes de mesure du Corg sont liées à la teneur en carbonates des échantillons par les formules suivantes :

Avant 2020 :  $\text{Incertitude\_Corg} = 0,50 * \text{Corg} + 0,01800 * \text{CaCO}_3 + 0,1360$

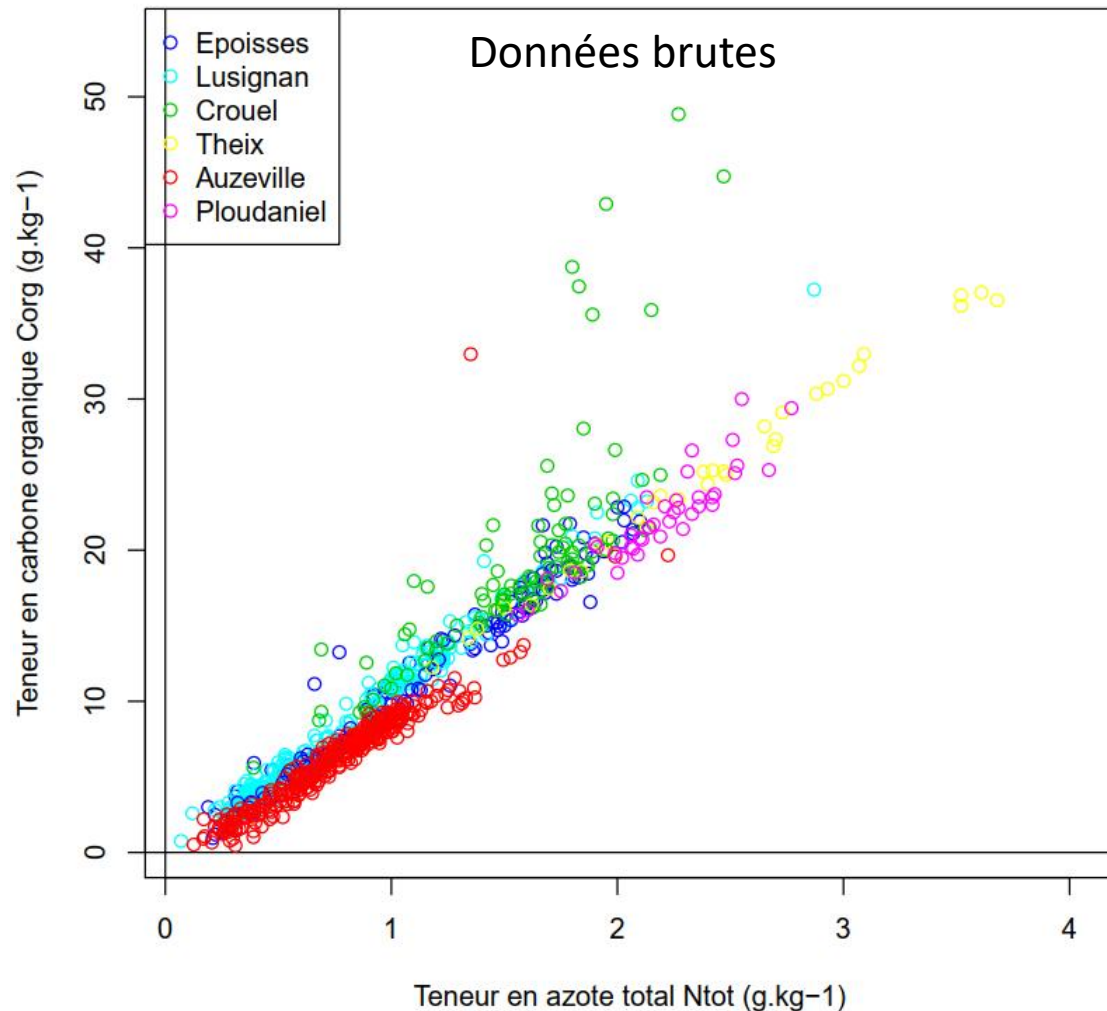
En 2024 :  $\text{Incertitude\_Corg} = 0,028 * \text{Corg} + 0,00590 * \text{CaCO}_3 + 0,2784$

où Corg est la concentration en carbone organique exprimée en g.kg<sup>-1</sup>  
et CaCO<sub>3</sub> est la concentration en carbone inorganique exprimée en g.kg<sup>-1</sup>.

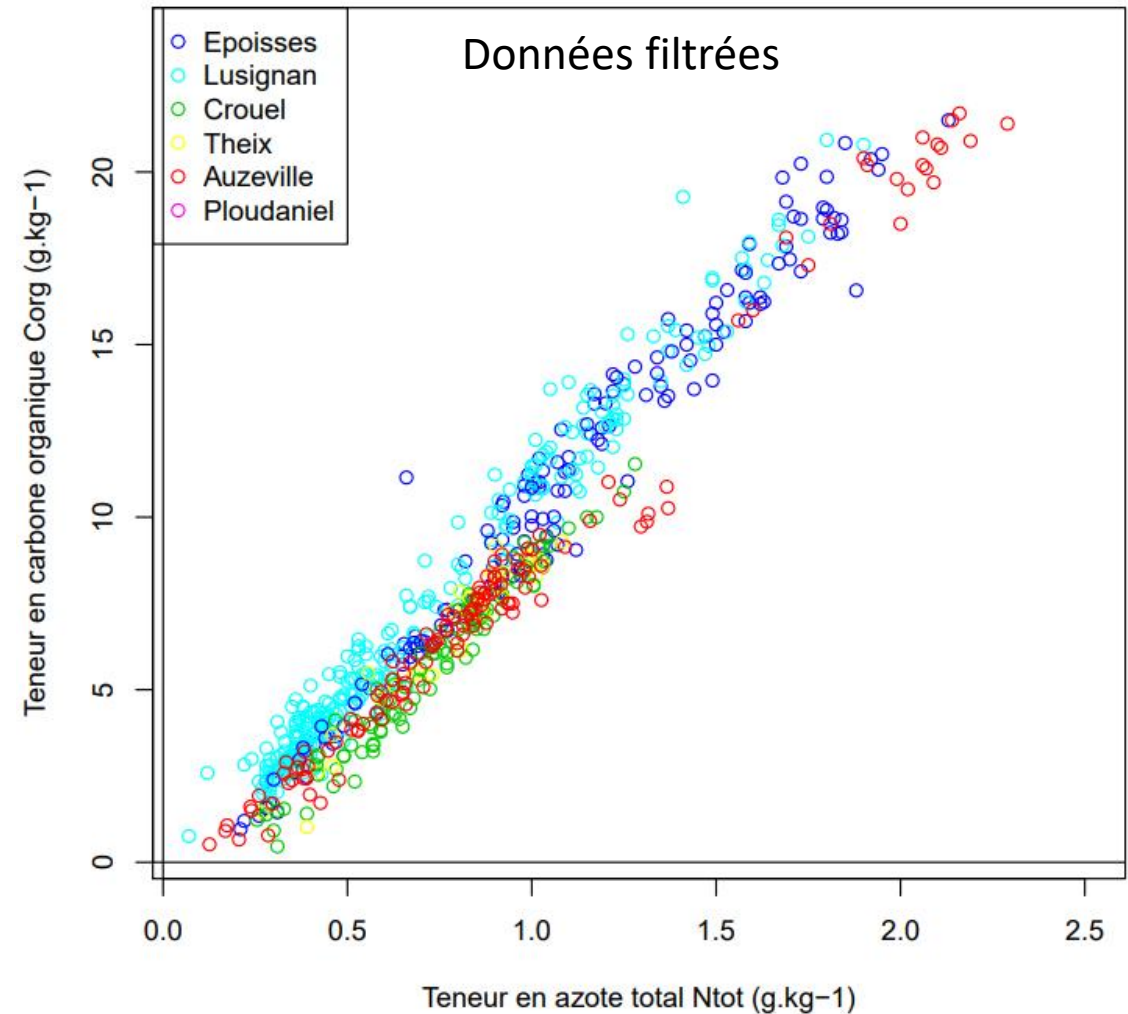
Sur les 957 échantillons de sol retenus et analysés, seuls 705 échantillons ont été utilisés pour la prédiction du Corg car ils offraient une incertitude de la mesure en Corg inférieure ou égale à 1 g.kg<sup>-1</sup>.

## ➤ Les données : les mesures spectrales et analytiques filtrées pour les incertitudes inférieures à 1g.kg-1 en Corg.

Distributions des teneurs en Corg et Ntot des 957 échantillons



Distributions des teneurs en Corg et Ntot des 705 échantillons



## ➤ Les modèles de régressions PLS pour la calibration

Le prétraitement et le traitement des données ont été réalisés avec le logiciel Chemflow. Le prétraitement suffisant a été d'exprimer l'absorbance des spectres.

$X$  = spectres ;  $Y$  = Corg ou Ntot.

La calibration des modèles de prédiction a été réalisée à l'aide de l'algorithme de régression PLS. Une **validation croisée** k-fold ( $k=4$ ) a été utilisée pour sélectionner le nombre optimal de composantes de régression PLS dans les modèles d'étalonnage afin d'éviter un surajustement.

Le nombre de composantes PLS qui a donné le **RMSECV** le plus bas a été identifié et l'information du critère de **Durbin-Watson** a été utilisée pour les situations ambiguës.

L'ensemble des données spectrales et d'analyses physico-chimiques **sont disponibles** sur [recherche.data.gouv.fr](https://recherche.data.gouv.fr)



## ➤ Objectif

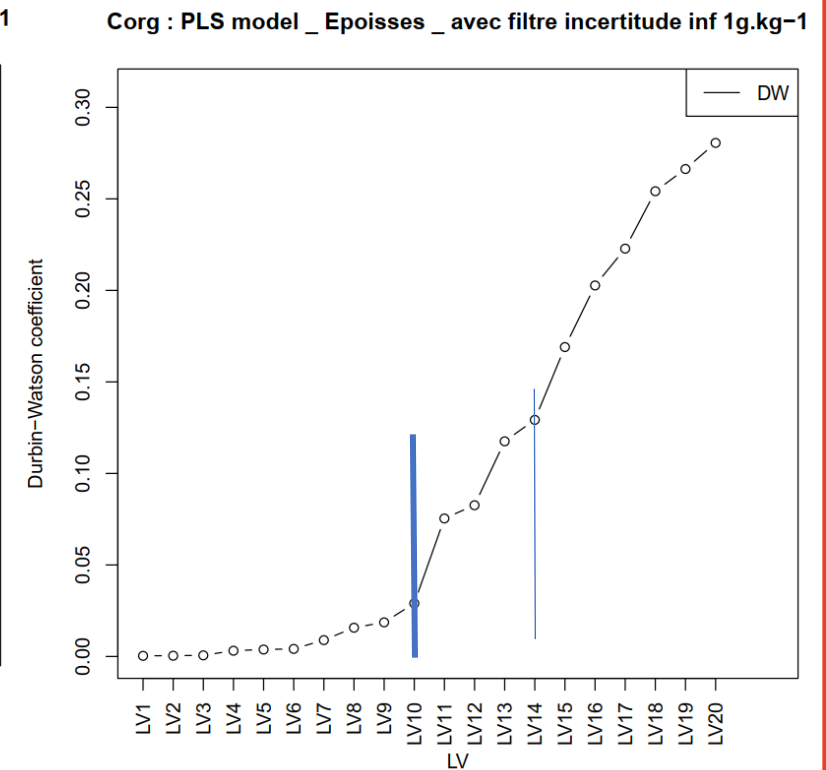
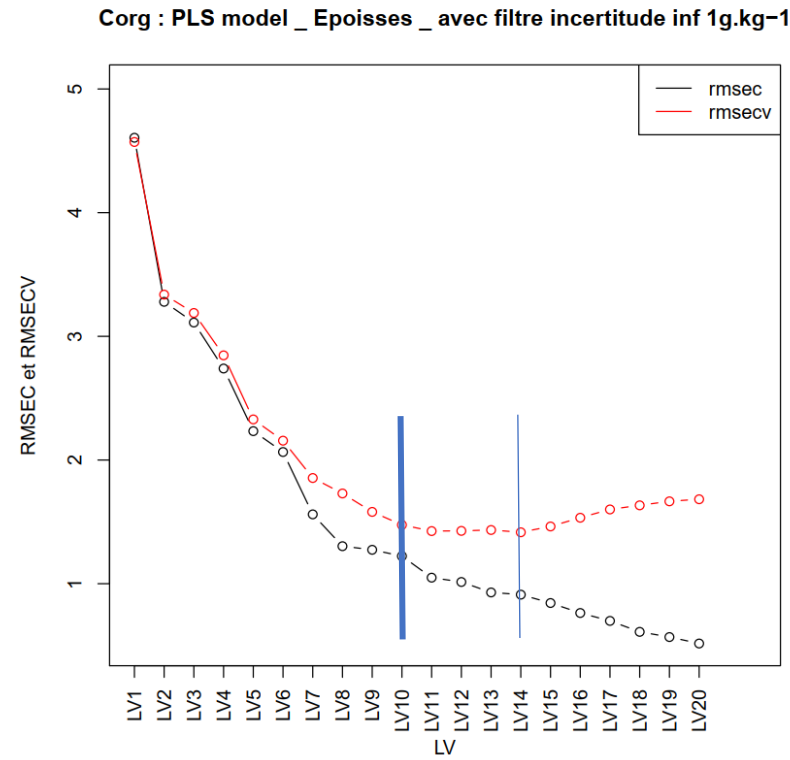
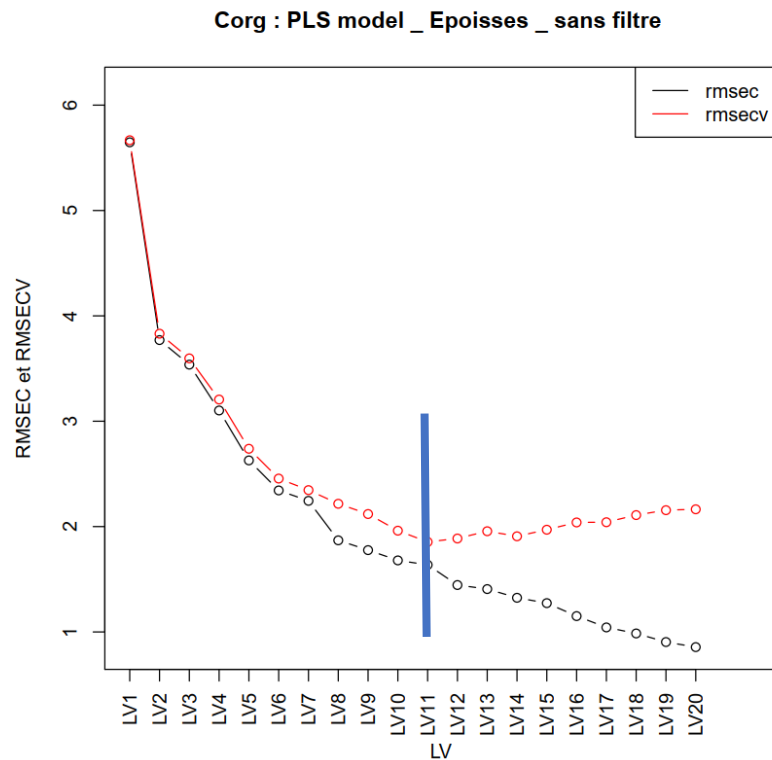
L'objectif est de prédire les concentrations en carbone organique des sols avec des précisions inférieures à 2 g.kg<sup>-1</sup> sur des données de l'unité expérimentale analysées en 2024 à partir d'un **modèle global** obtenu sur les 6 sites expérimentaux analysés en 2013 ou à partir d'un **modèle local**, c'est-à-dire à partir des données obtenues sur ce même site mais 10 ans plus tôt.



## ➤ Modèles de calibration locale utilisant les 178 ou 140 spectres du site d'Epoisses sur des échantillons de 2013.

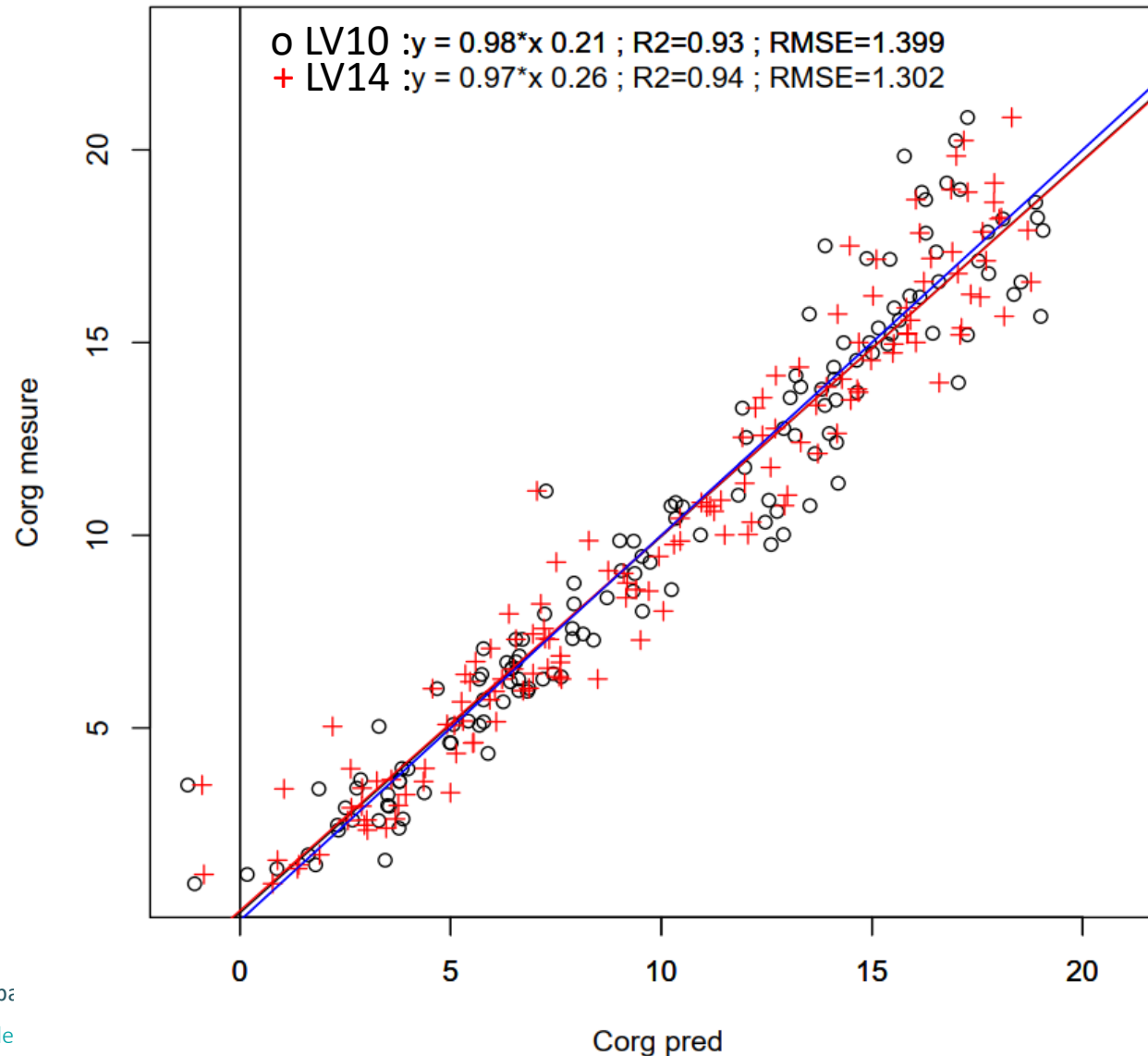
**RMSECV=1,85 ; RMSEC=1,64 ; DW=0,02 ;**  
**RPD=3.26. 178 données.**  
**LV=11**

**RMSECV=1,47 ; RMSEC=1,22 ; DW=0,029 ;**  
**RPD=3.65. 140 données, incertitude Corg**  
 **$\leq 1 \text{ g.kg}^{-1}$ ; LV=10**



## ➤ Modèles de calibration locale utilisant les 140 spectres du site d'Epoisses sur des échantillons de 2013.

Corg pred versus Corg mesure : LV10 et LV14

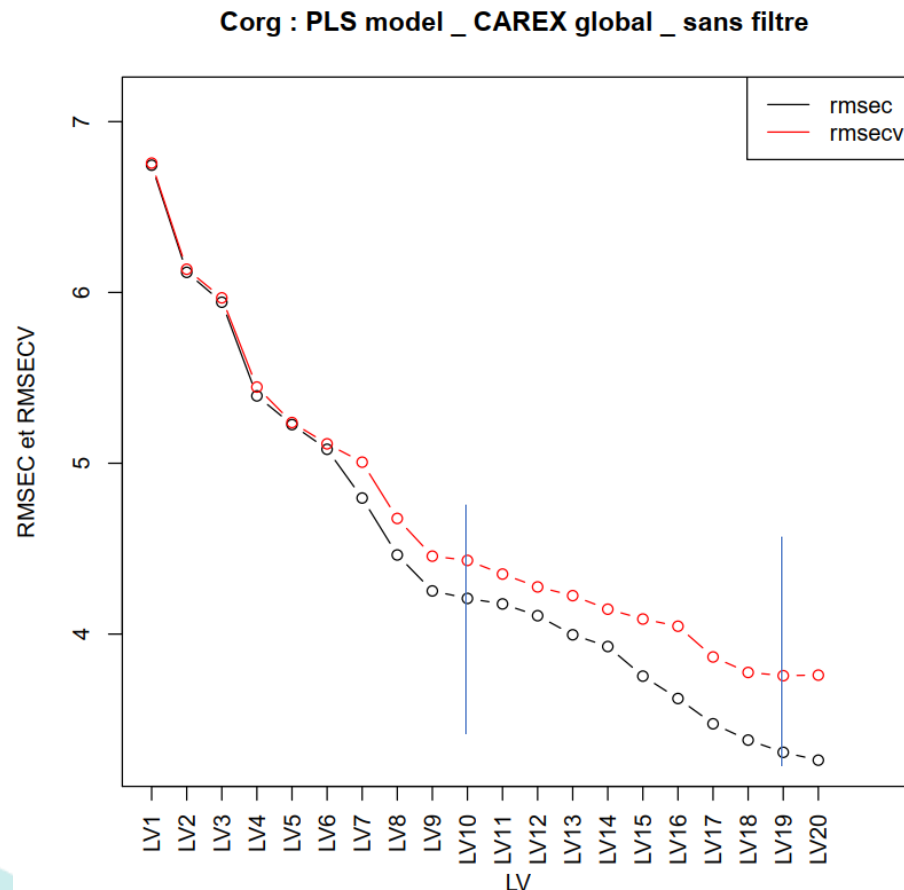


## ➤ Modèles de calibration globale CAREX utilisant les 957 spectres des 6 sites expérimentaux avec des échantillons de 2013.

**RMSECV=3,8 ; RMSEC=3,3 ; DW=0,12.**

957 données.

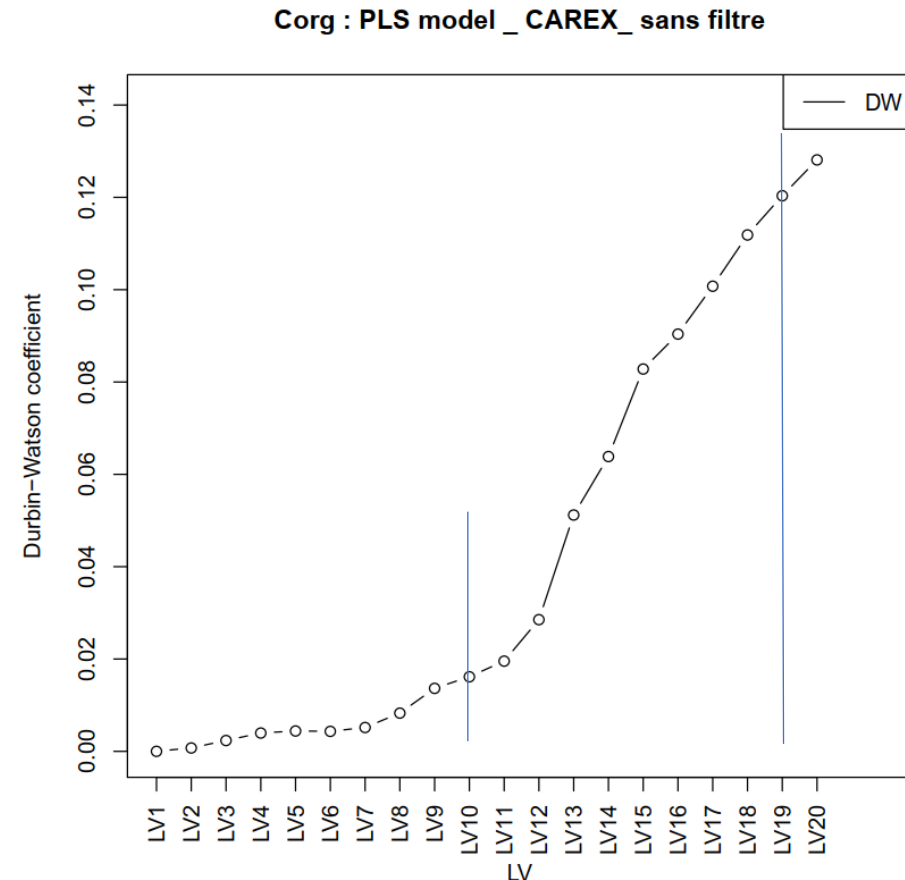
LV=19



**RMSECV=4,4 ; RMSEC=4,2 ; DW=0,016.**

957 données.

LV=10



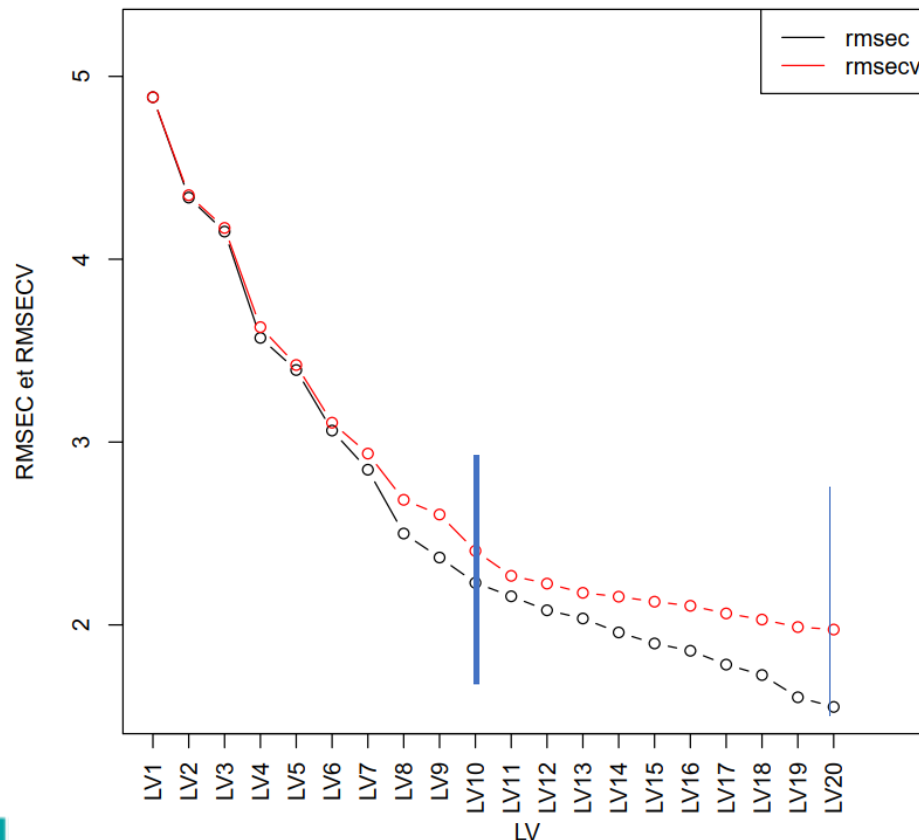
## ➤ Modèles de calibration globale CAREX utilisant les 705 spectres des 6 sites expérimentaux dont l'incertitude sur Corg est $\leq 1 \text{ g.kg}^{-1}$

**RMSECV=1,97 ; RMSEC=1,55 ; DW=0,13.**

705 données, incertitude Corg  $\leq 1 \text{ g.kg}^{-1}$

LV=20

Corg : PLS model \_ CAREX global \_ avec filtre incertitude inf 1g.kg-1

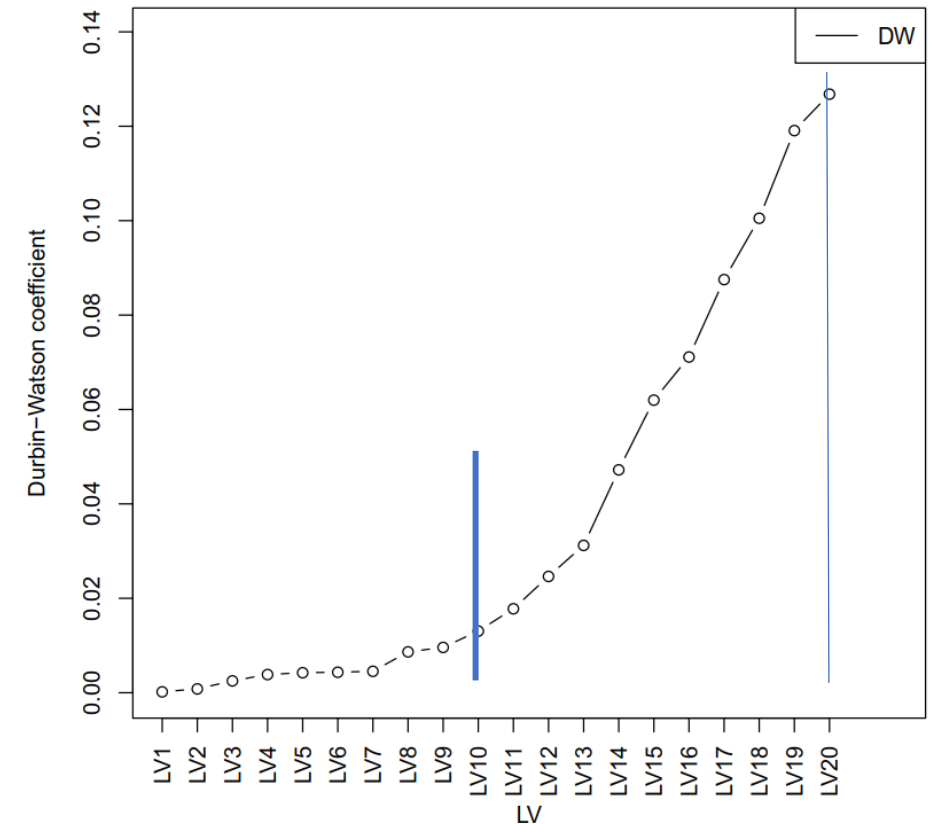


**RMSECV=2,40 ; RMSEC=2,23 ; DW=0,013.**

705 données, incertitude Corg  $\leq 1 \text{ g.kg}^{-1}$

LV=10

Corg : PLS model \_ CAREX \_ avec filtre incertitude inf 1g.kg-1



## ➤ Résumé des modèles de calibration et prise en compte des incertitudes des mesures de référence Corg ≤ 1 g.kg<sup>-1</sup>

| Modèles                                 | nb data    | SD data     | LV        | DW           | RMSECV/RMSEP | RMSEC       | R <sup>2</sup> | RPD         |
|---|------------|-------------|-----------|--------------|--------------|-------------|----------------|-------------|
| Global CAREX brut non filtré            | 957        | 7,4         | 10        | 0,016        | <b>4,43</b>  | 4,21        | 0,68           | 1,67        |
| Global CAREX brut non filtré            | 957        | 7,4         | 19        | 0,12         | 3,76         | 3,3         | 0,8            | 1,97        |
| Global CAREX filtré incertitude inf1g   | 705        | 4,94        | 10        | 0,013        | <b>2,4</b>   | 2,23        | 0,8            | 2,06        |
| Global CAREX filtré incertitude inf1g   | 705        | 4,94        | 20        | 0,13         | 1,97         | 1,55        | 0,9            | 2,51        |
| Local EPOISSES brut non filtré          | 178        | 6,04        | 11        | 0,02         | <b>1,85</b>  | 1,64        | 0,93           | 3,26        |
| Local EPOISSES filtré incertitude inf1g | <b>140</b> | <b>5,36</b> | <b>10</b> | <b>0,029</b> | <b>1,47</b>  | <b>1,22</b> | <b>0,95</b>    | <b>3,65</b> |
| Local EPOISSES filtré incertitude inf1g | 140        | 5,36        | 14        | 0,13         | 1,41         | 0,91        | 0,97           | 3,80        |

*Clairotte et al., Geoderma 276 (2016) sols RMQS N= 3000,  
Corg moyen=18,9 g.kg<sup>-1</sup>, sd: 18,3 g.kg<sup>-1</sup>, **SEP=4,4 g.kg<sup>-1</sup>, RPD = 2,3***

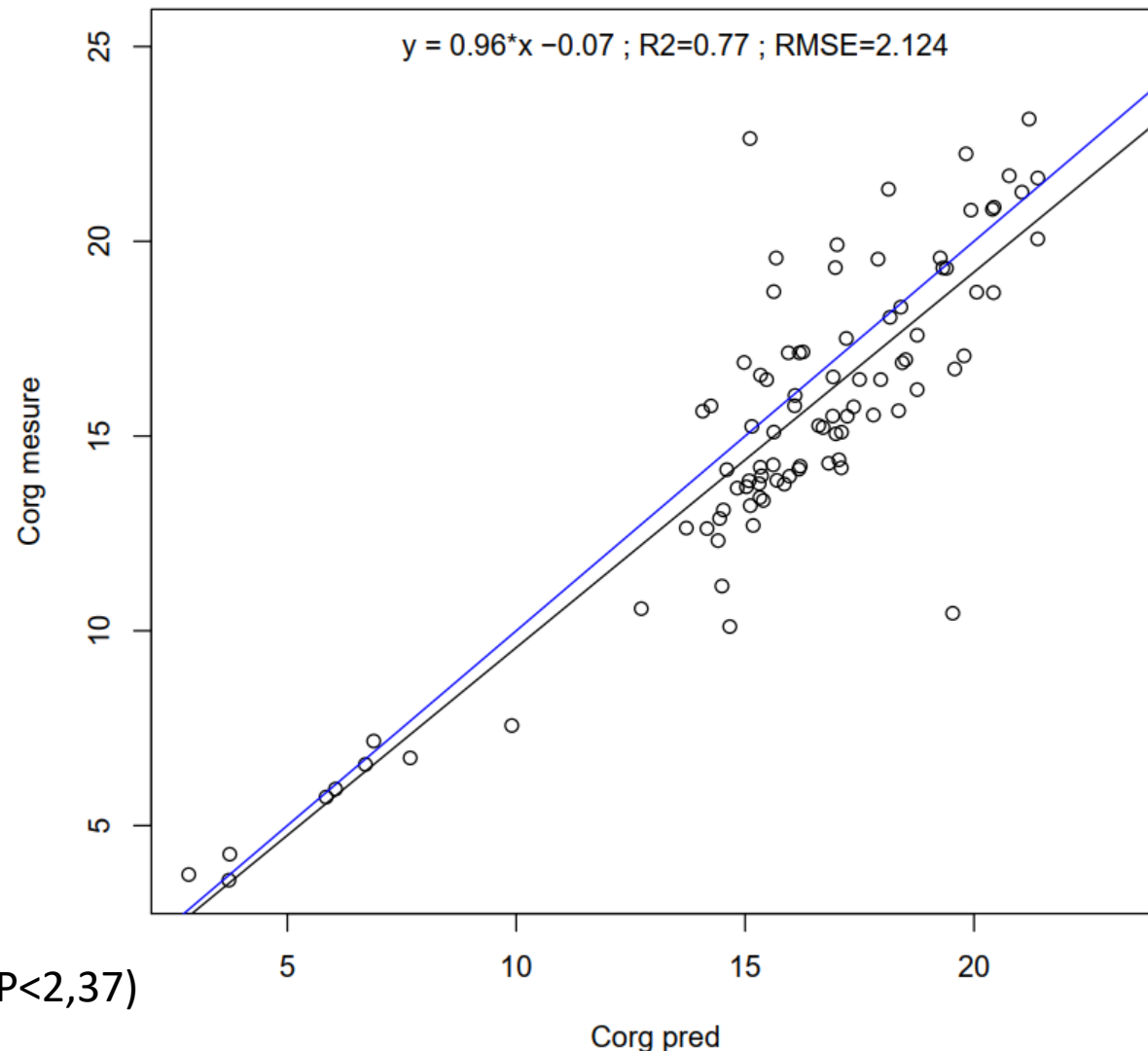
## ➤ Performances de prédiction sur les données du site d'Epoisses en 2024

Prédiction de la teneur en Corg à partir des spectres obtenus sur le site d'Epoisses en 2024, en utilisant un modèle établi sur des données d'Epoisses obtenues en 2013 et avec 10 variables latentes (LV10).

**RMSEP = 2,12 g.kg<sup>-1</sup>**

**RPD = 2,02**

(Tous les autres modèles ont  $2,32 < \text{RMSEP} < 2,37$ )



# ➤ Perspectives et remerciements

## ➤ Perspectives

- Application du modèle local Epoisses sur les échantillons sur lesquels les spectres ont été acquis mais pas les mesures analytiques. Prédiction de la teneur en Corg pour 60 nouveaux points et élaboration d'une **cartographie des teneurs en Corg** sur le domaine d'Epoisses.

## ➤ Merci aux participants

- *LAS : Amandine Etayo, Nicolas Proix, Etienne Lamy*
- *UMR SAS, INRAE, Institut Agro - Agrocampus Ouest : Youssef Fouad, Pascal Pichelin, Fabien Goge*
- *UMR ECOSYS : Jeanne-Chantal Thoisy, Cyril Girardin*
- *UMR EMMAH : Claude Doussan*
- *UE DIASCOPE : Brigitte Montegano, Pascal Sartre*
- *UE Auzeville : André Gavaland*
- *UE Ferlus : Sandra Novak, Stéphanie Mahieu*
- *UR LBE : Virginie Rossard, Julie Jimenez, Eric Latrille*
- *Agronomie, Grignon pour le réseau CAREX : Marie-Noël Mistou*
- *Info&Sols : Maud Seger*
- *L'UE d'Epoisses*

## ➤ Et aux experts externes

- *Bernard Barthès (Eco&Sols, IRD), Mathieu Lesnoff (SELMET, Cirad), Jean-Michel Roger (UMR ITPA, Inrae), Emmanuelle Vaudour (Ecosys, INRAE)*