

# Prédiction VNIRS de la teneur en carbone et azote de sols variablement salés du Sénégal

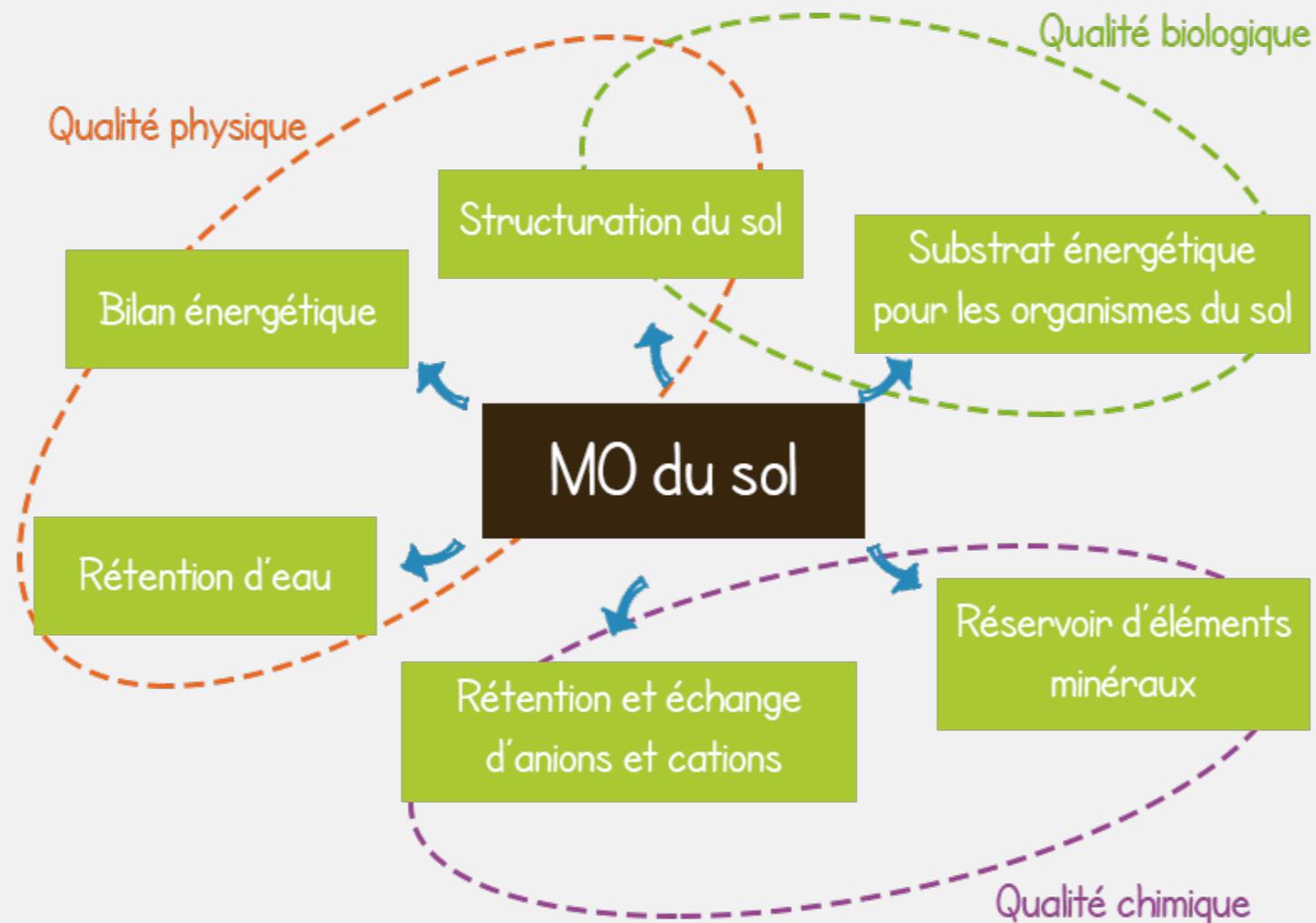
Aurélie CAMBOU, Bernard G. BARTHÈS, Tiphaine CHEVALLIER, Patricia MOULIN, Lydie CHAPUIS-LARDY



HelioSPIR– 29/06/2021



## Matières organiques des sols



Les rôles de la matière organique - C.Marsden

- A l'origine de nombreux services écosystémiques rendus par les sols

## Salinisation des sols – Impact à différentes échelles

- **Salinisation des sols** : enrichissement excessif du sol en sels solubles (e.g.,  $\text{Na}^+$ ,  $\text{Cl}^-$ ) – phénomènes naturels ou anthropiques
- **Effet sur les cycles de C et N dans les sols** : e.g., production de biomasse par les plantes limitée (stress osmotique)
- **Au niveau mondial, le phénomène de salinisation des terres** :
  - concerne 7,5 % de la surface terrestre
  - touche principalement les zones arides et semi-arides
- **Au Sénégal, ce phénomène concerne** :
  - **plus de 1,7 million d'hectares**, limitant considérablement les potentialités de production agricole (e.g., région du Sine Saloum)

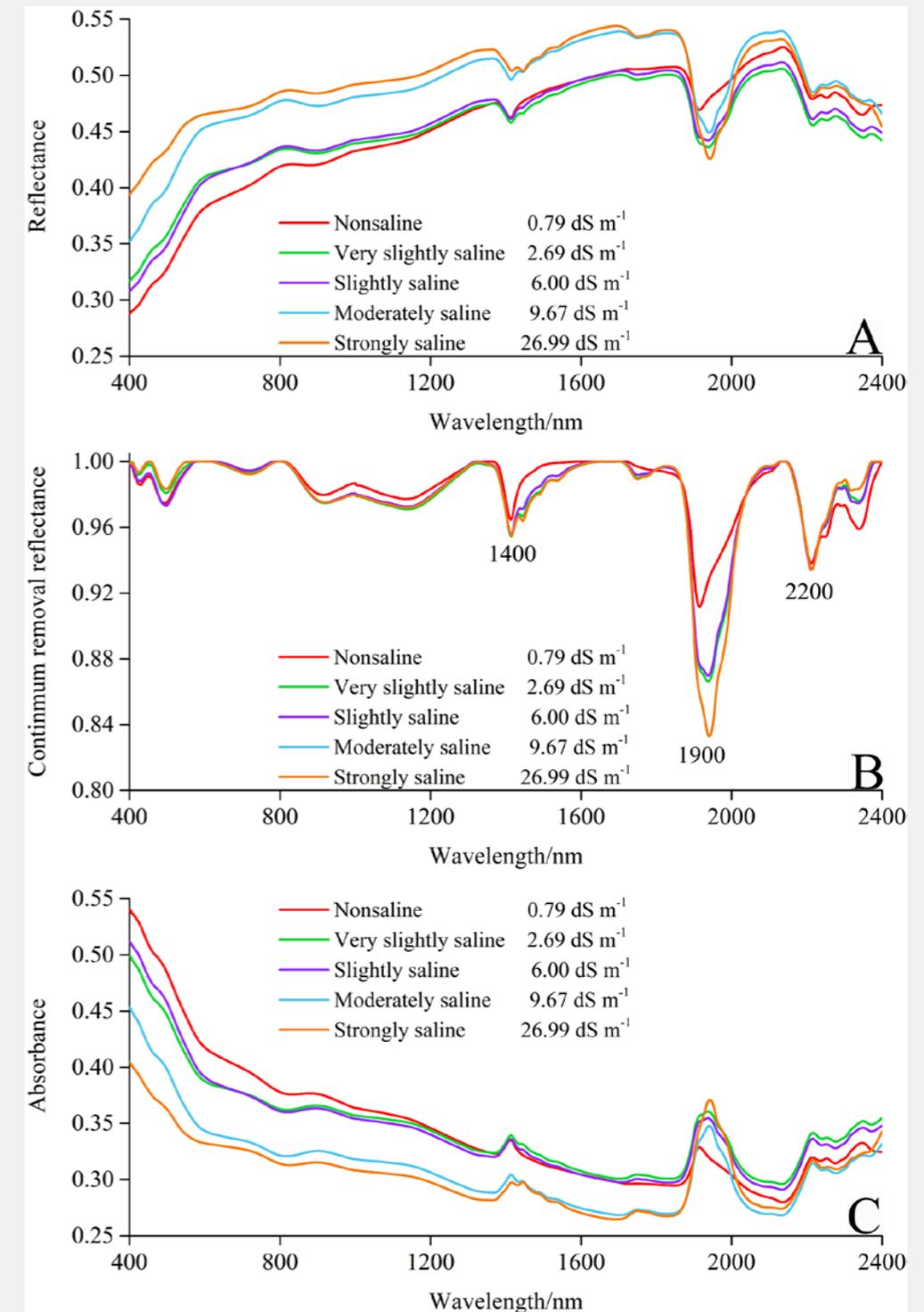


**Nécessité de surveiller la progression de la salinisation des sols et son impact sur les MOS à long terme**

## Salinisation des sols & spectroscopie VNIRS

- Effet de la salinité des sols sur les spectres de réflectance / absorbance :
  - Absorbance dans le visible et proche infrarouge impactée par la présence de sels
  
- Donc, dans les régions affectées par des degrés de salinisation variables :
  - forte hétérogénéité spectrale potentielle

Comment considérer au mieux cette hétérogénéité pour optimiser les modèles de prédictions des teneurs en MOS par VNIRS ?



## Objectifs du travail

**Optimiser les prédictions des teneurs en carbone organique (SOC) et N (= composés principaux de la MOS) par VNIRS dans des sols variablement salés.**

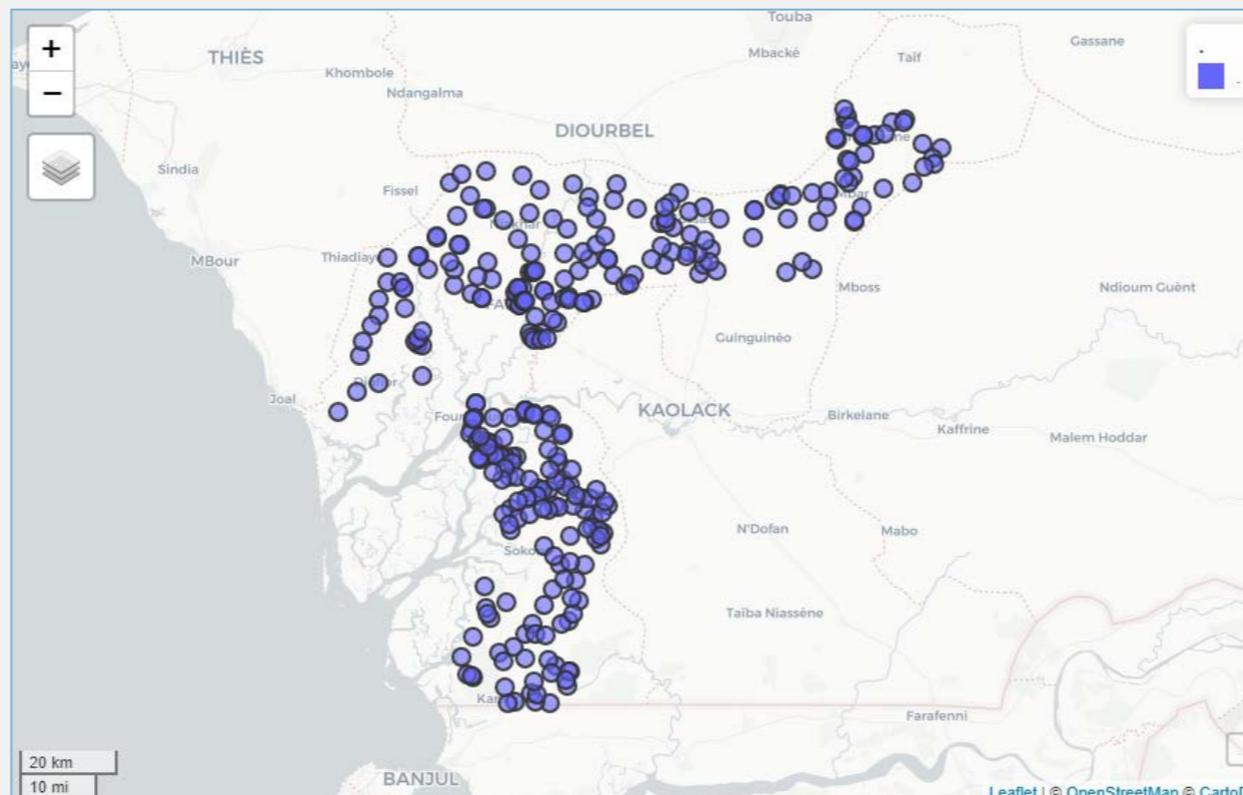
**Pour cela, identifier la meilleure combinaison :**

**prétraitement spectral × transformation de la variable × type de modèle de prédiction**

# 6 MATÉRIELS ET MÉTHODES

## Zone d'étude

- Région du Sine Saloum (Sénégal)
  - Climat semi-aride
  - Différents types de sols (e.g., Gleyic Solonchaks, Ferralic/Sideralic Arenosols, Umbric Gleysols)
  - Végétation naturelle principalement constituée de savanes (arborées, arbustives et herbacées)
  - **311 points de prélèvement à 0-25 cm de profondeur (projet de Laure Chauvin, 2012)**

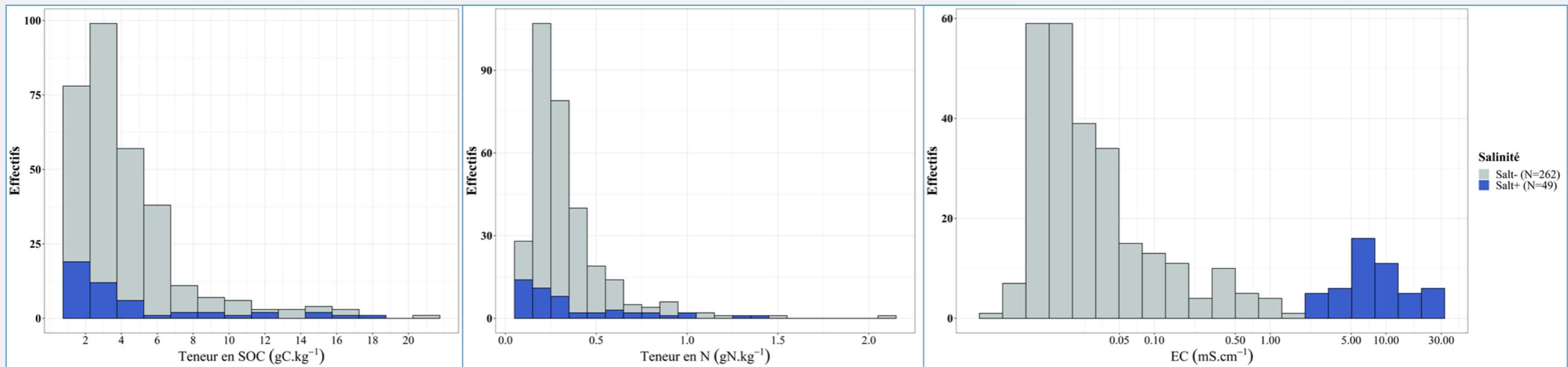


# 7 | MATÉRIELS ET MÉTHODES

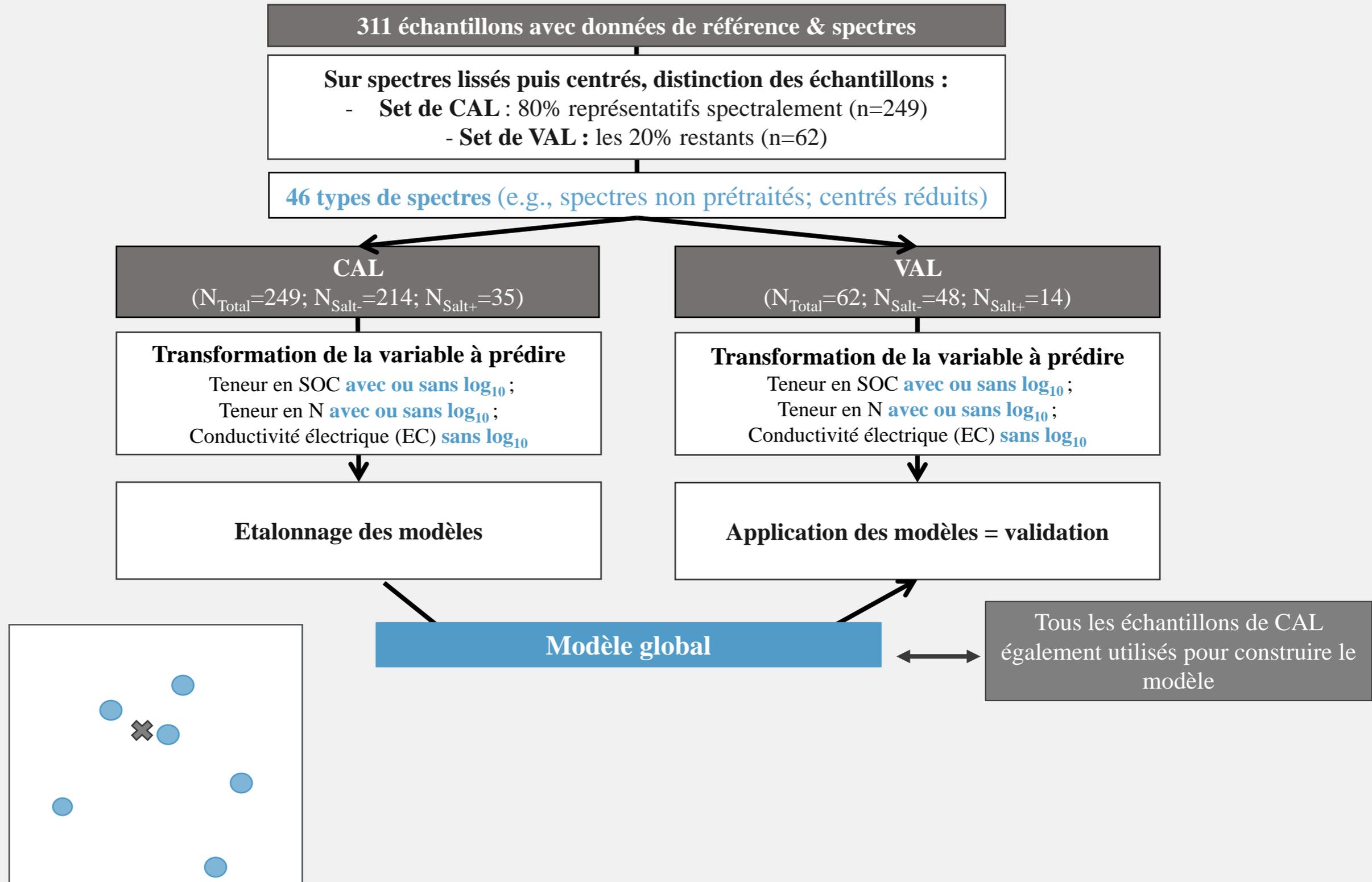
## Variables étudiées et distributions

- Détermination de deux classes de salinité (d'après Sadio 1991)
- Teneurs en SOC & N ; EC : distributions non normales

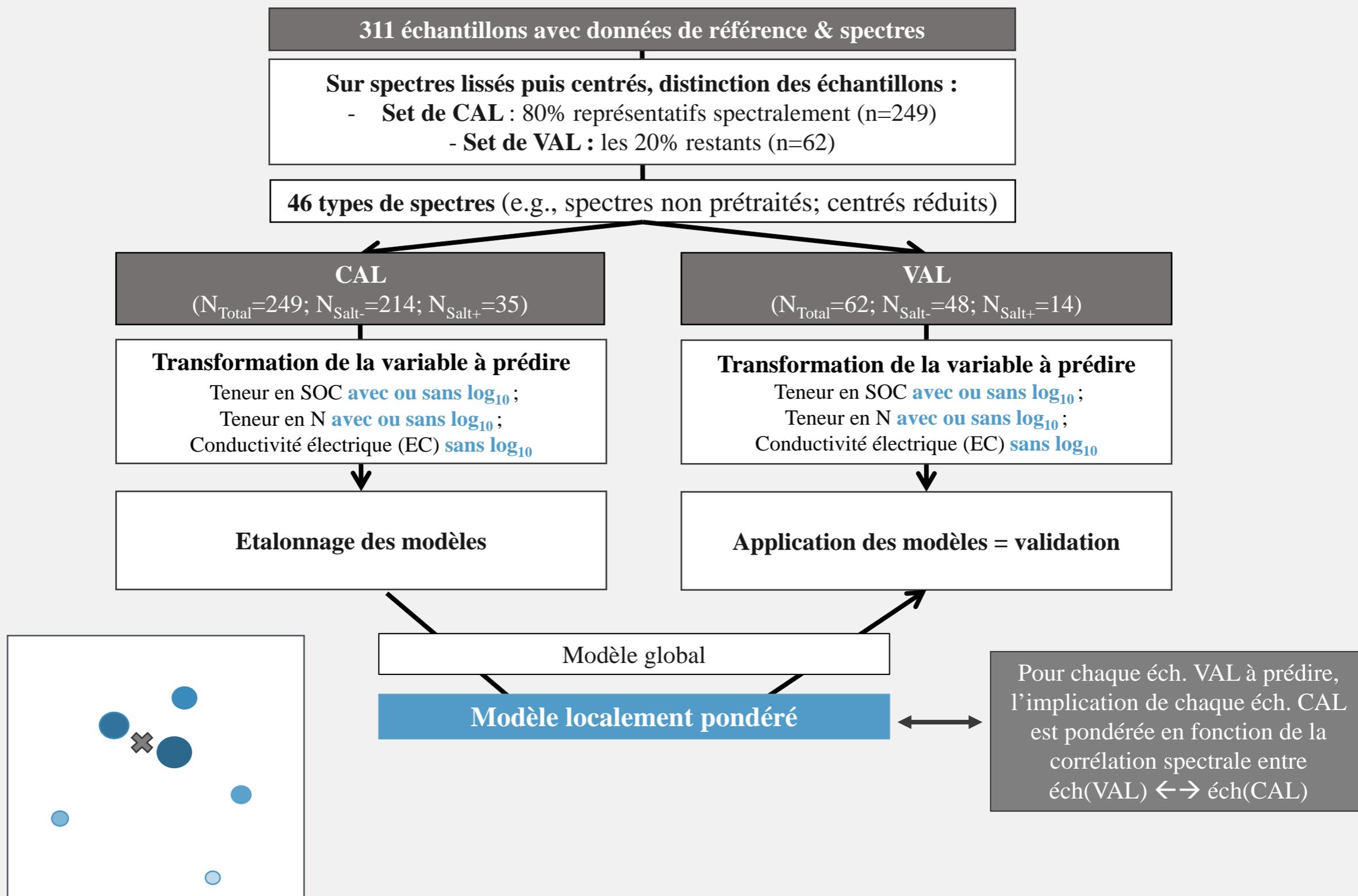
Sols non ou peu salés (Salt-)	Sols salés (Salt+)
EC 0 à 2.0 mS.cm <sup>-1</sup>	EC > 2.0 mS.cm <sup>-1</sup>
N = 262	N = 49



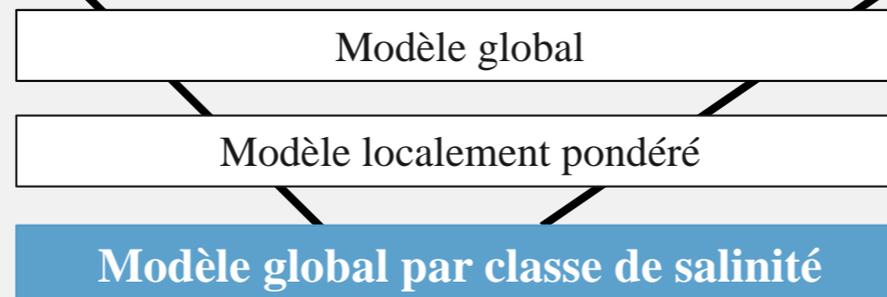
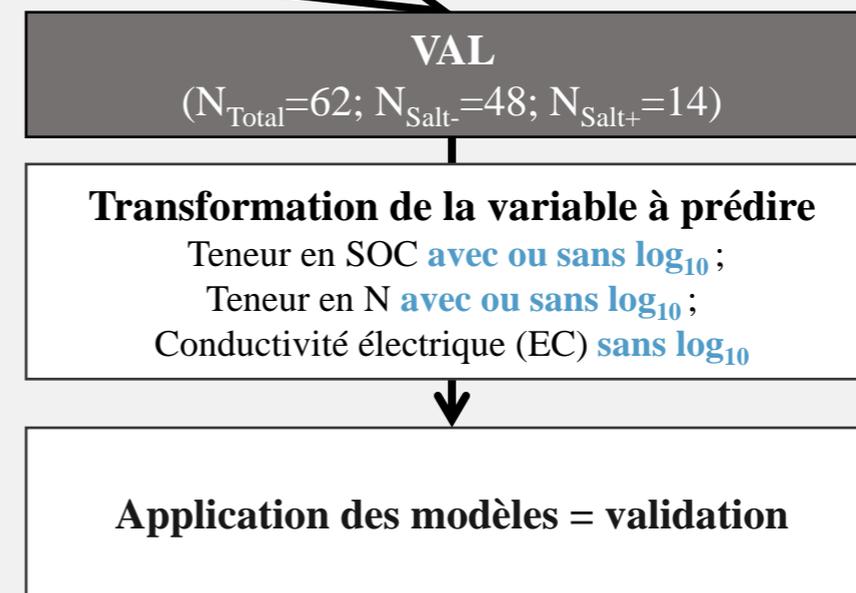
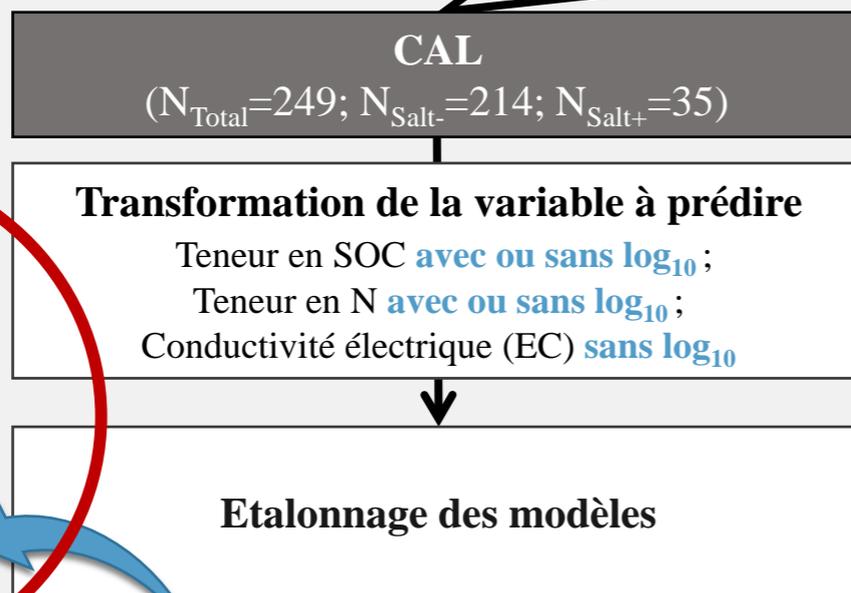
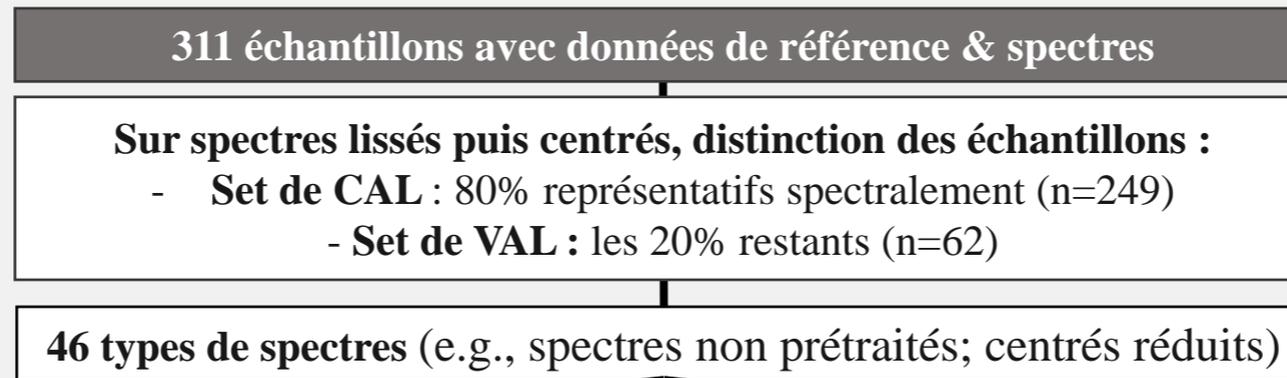
## 8 MATÉRIELS ET MÉTHODES



## 9 MATÉRIELS ET MÉTHODES



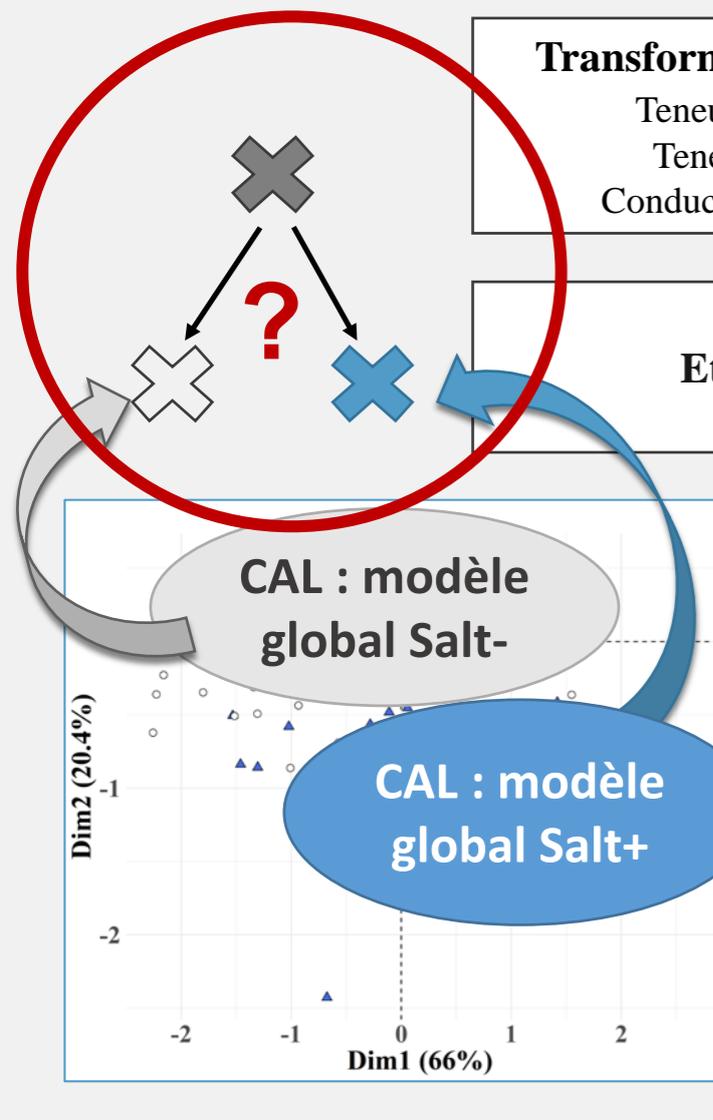
# MATÉRIELS ET MÉTHODES



**CAL** : Construction d'un modèle de prédiction de SOC, N, EC par classe de salinité (Salt- vs. Salt+)

**VAL** : prédiction de la classe de salinité à partir du spectre (modèle préalablement construit sur CAL)

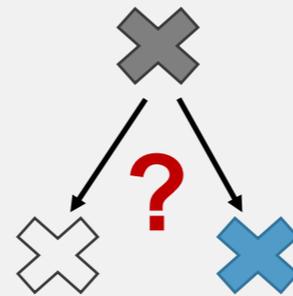
Utilisation du modèle construit dans la classe correspondante pour prédire la variable (SOC, N, EC) de VAL



## RÉSULTATS ET DISCUSSION

## Résultats de la prédiction des classes de salinité (VAL)

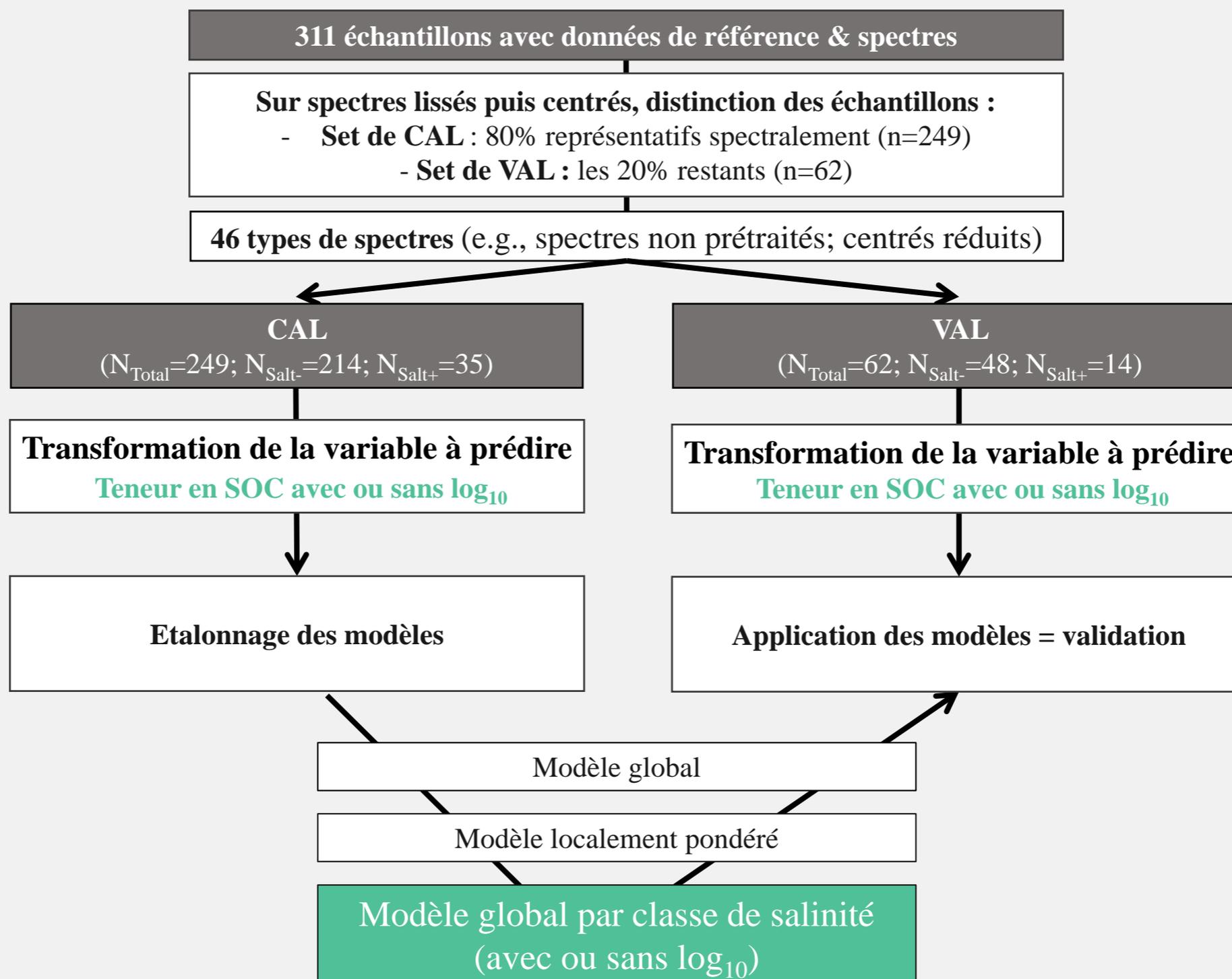
- Testée pour les 46 types de spectres



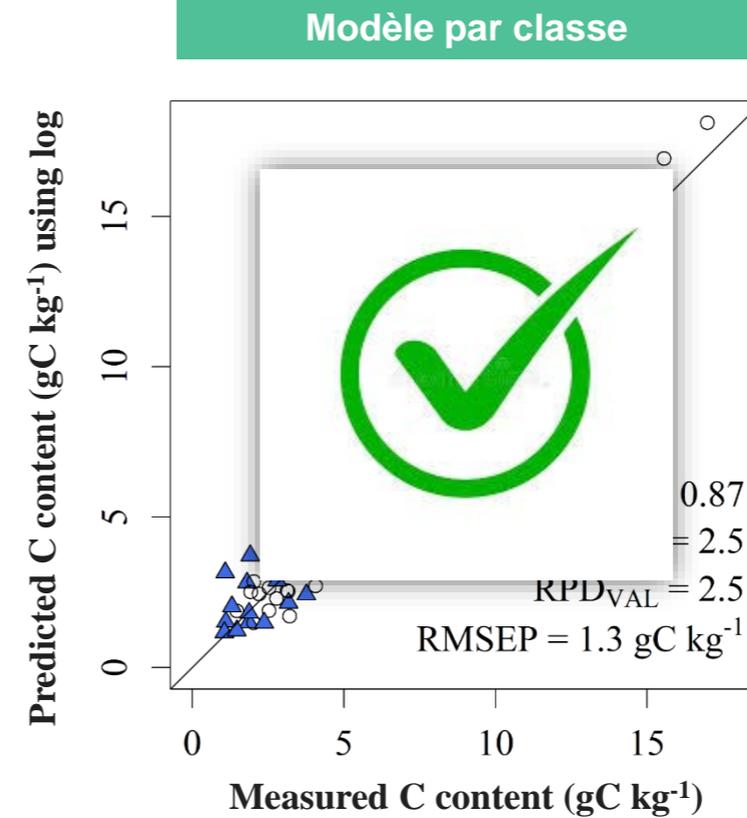
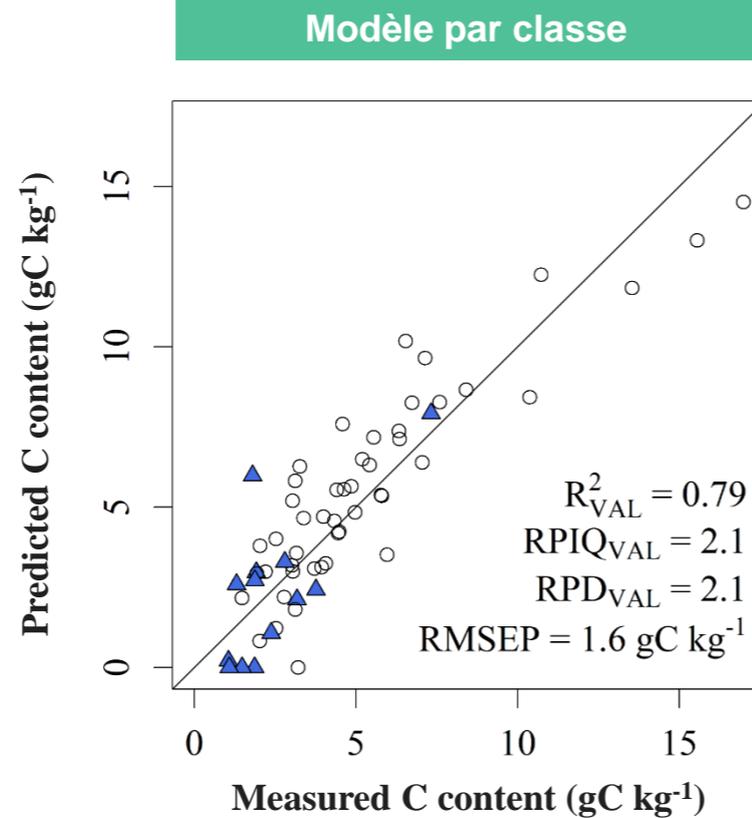
Proportion de types de spectres (TS)	Proportion de VAL bien prédits ( $N_{\text{Total}} = 62$ ; $N_{\text{Salt-}} = 48$ ; $N_{\text{Salt+}} = 14$ )
57 % (=26 TS)	<b>100 %</b> (=62 éch)
33 % (=15 TS)	<b>98 %</b> (=61 éch)
4 % (=2 TS)	<b>97 %</b> (=60 éch)
6 % (=3 TS)	<b>94 %</b> (=58 éch)



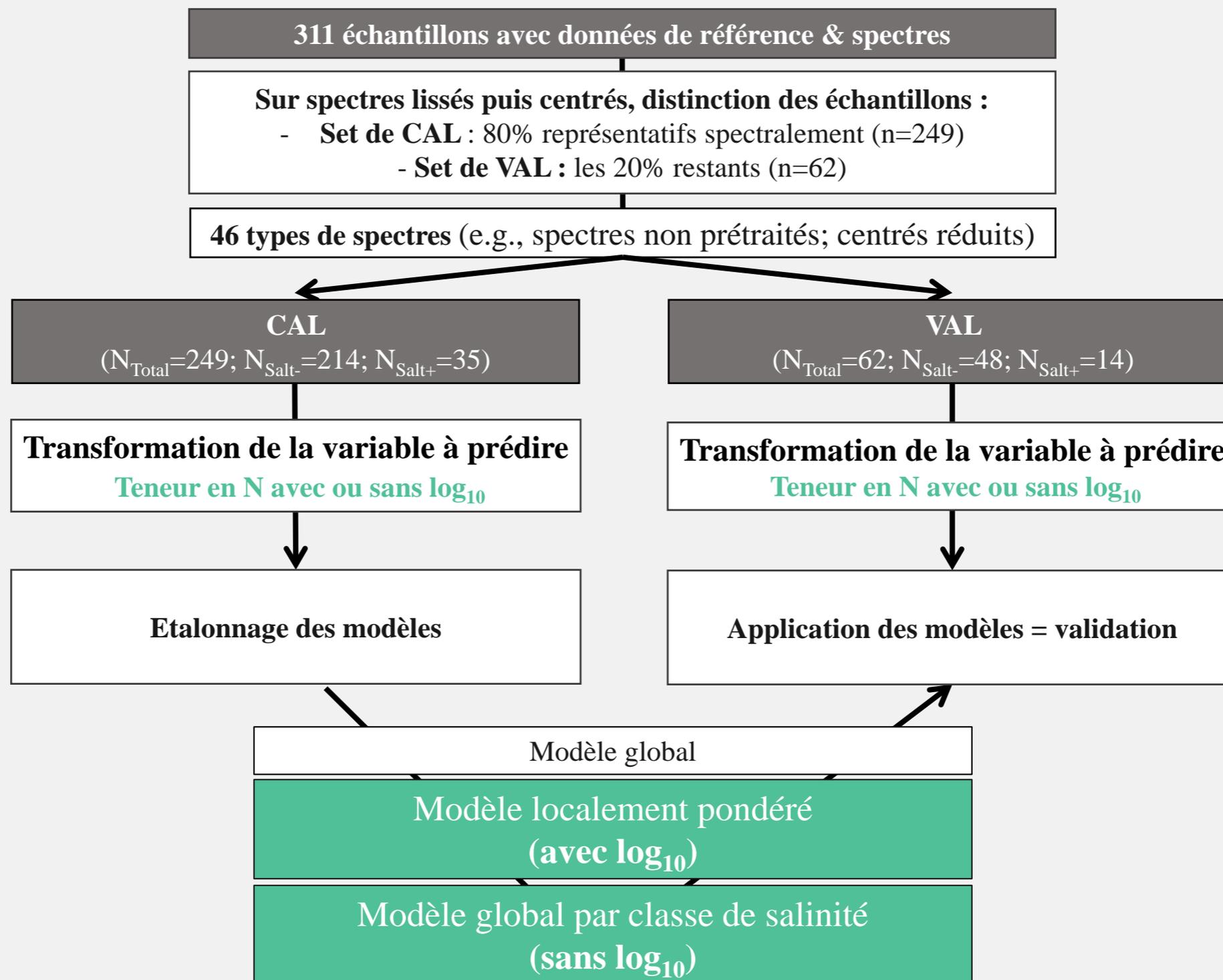
## Résultats de prédiction (VAL) – SOC content



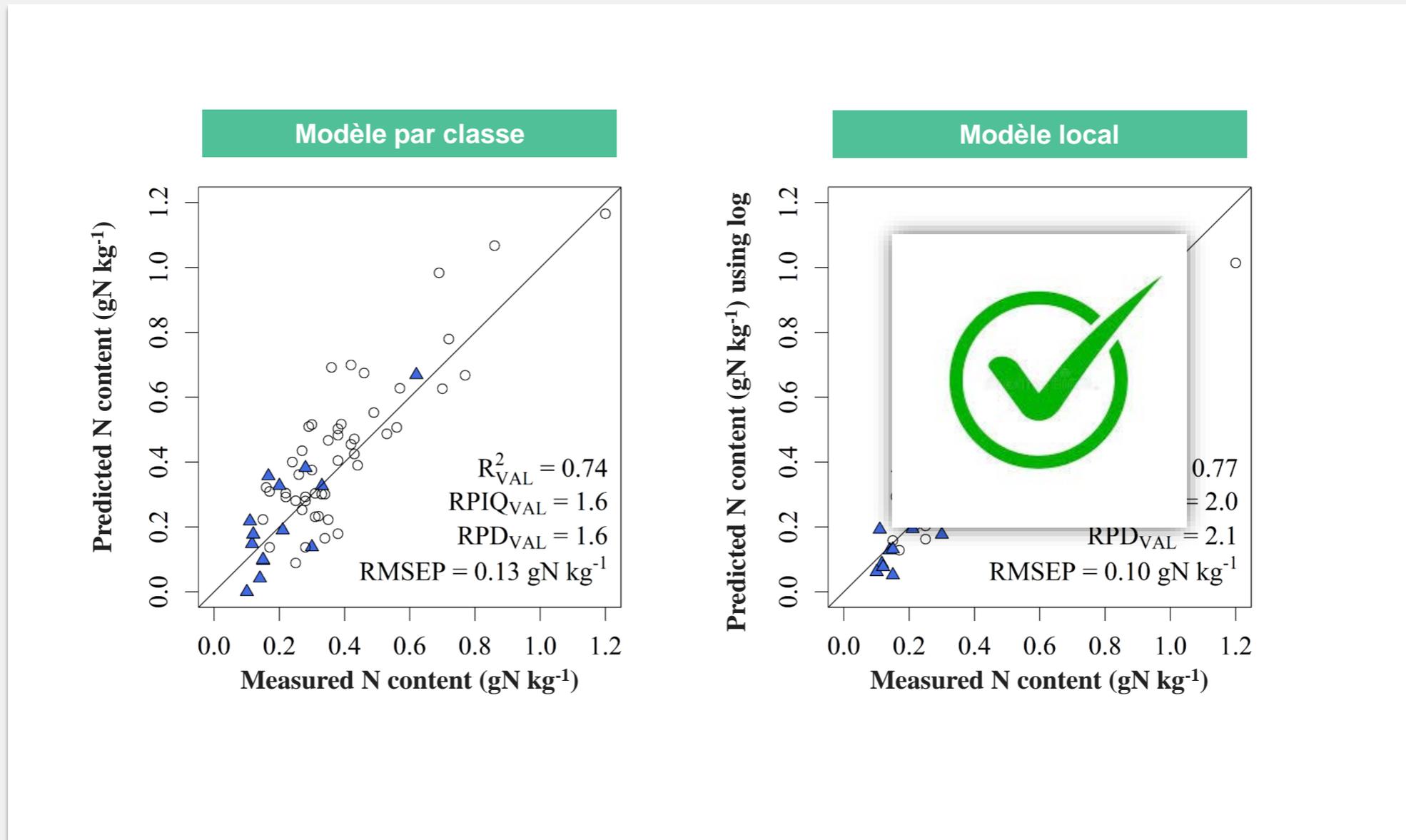
## Résultats de prédiction (VAL) – SOC content



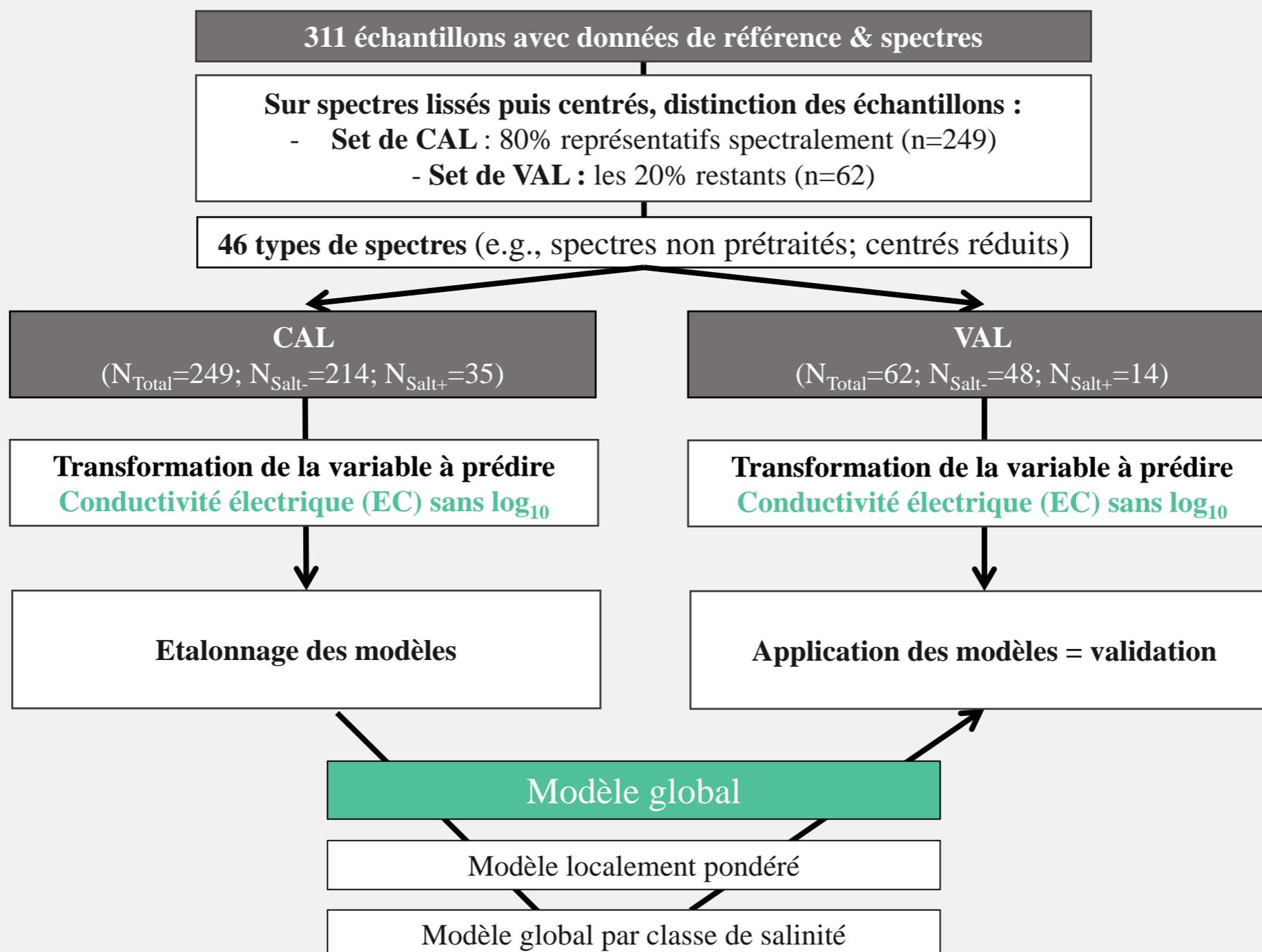
## Résultats de prédiction (VAL) – N content



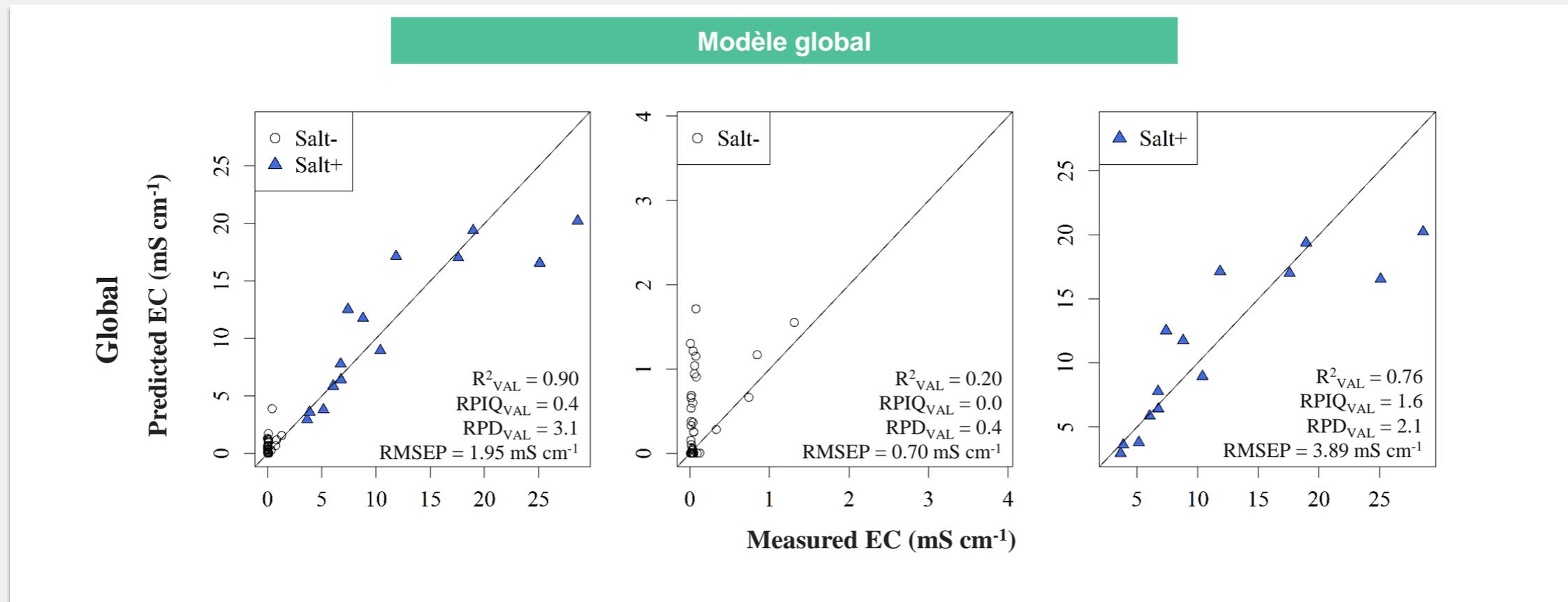
## Résultats de prédiction (VAL) – N content



## Résultats de prédiction (VAL) – EC



## Résultats de prédiction (VAL) – EC



- Variable hétérogène et clusterisée : limite des modèles de prédiction

## RÉSULTATS ET DISCUSSION

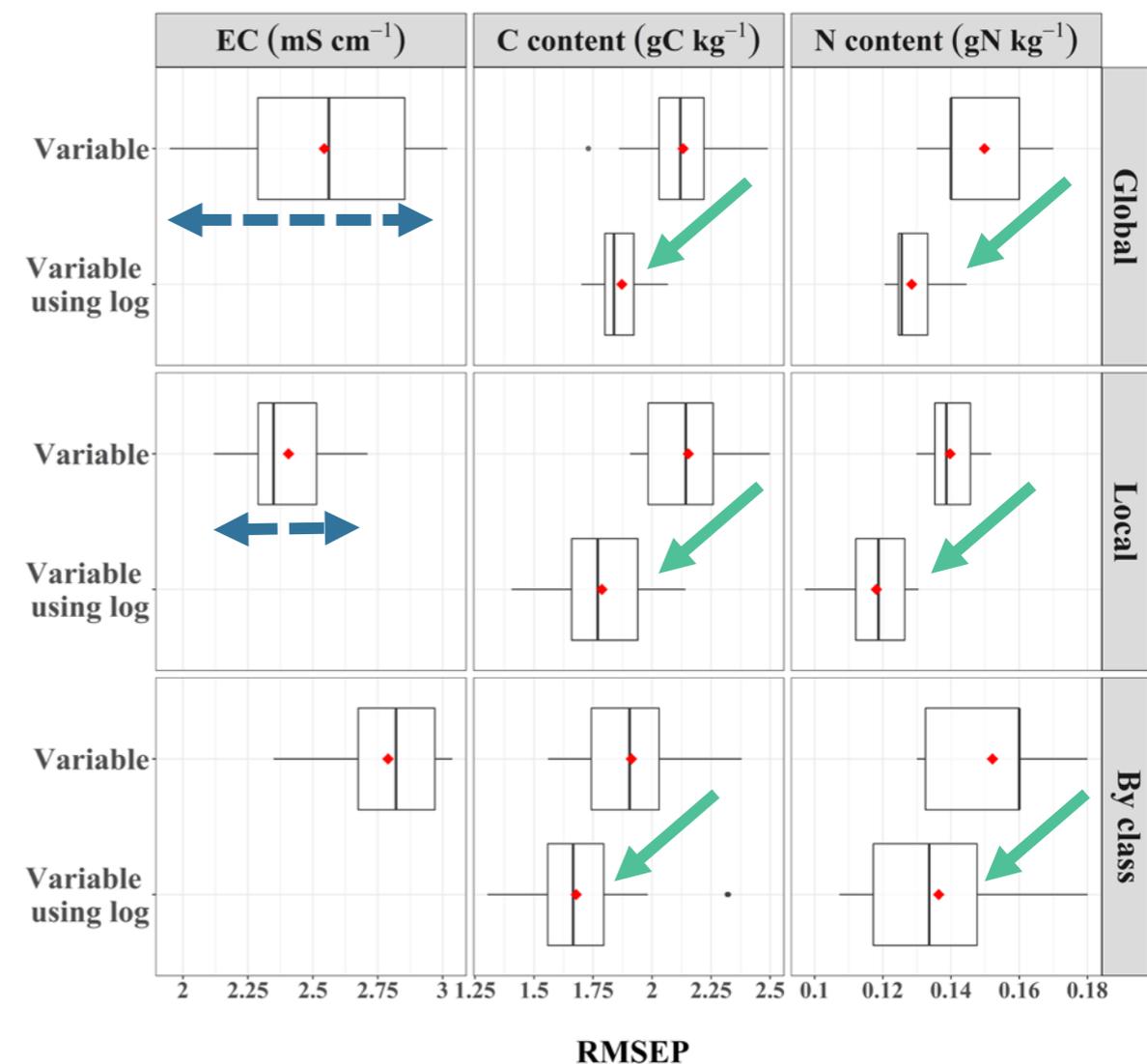
# Résultats de validation, tous prétraitements confondus (spectres bruts + 45 prétraitements)

- Plus RMSEP (=erreur de prédiction) est faible, plus le modèle est précis
- Bilan :

L'enjeu du choix du prétraitement spectral dépend du type de modèle et de la variable étudiée

Gain systématique de précision grâce à la transformation  $\log_{10}$  des teneurs en SOC et N (pas le cas pour EC)

Distribution de l'erreur de prédiction (RMSEP) pour chaque type de modèle et chaque variable – tous les 46 prétraitements spectraux confondus



# 19 CONCLUSION ET PERSPECTIVES

- **Prédictions des teneurs en SOC et N :**
- Résultats améliorés par la transformation des variables
- Modèles par classe et local : les plus intéressants pour la prédiction de ces deux variables
  
- **Prédictions de EC**
- Variable très hétérogène et assez clusterisée. Pour ce type de variable, les modèles prédictifs sont limités
- Le modèle sur la population VAL totale semble être précis mais perte de la précision par classe :
  - pour Salt- car population trop homogène
  - pour Salt+ car population aux effectifs trop faibles

## Pour aller plus loin :

**Poursuivre la recherche d'un compromis optimal de combinaison  
[prétraitement spectral / de la variable & type de modèle]**

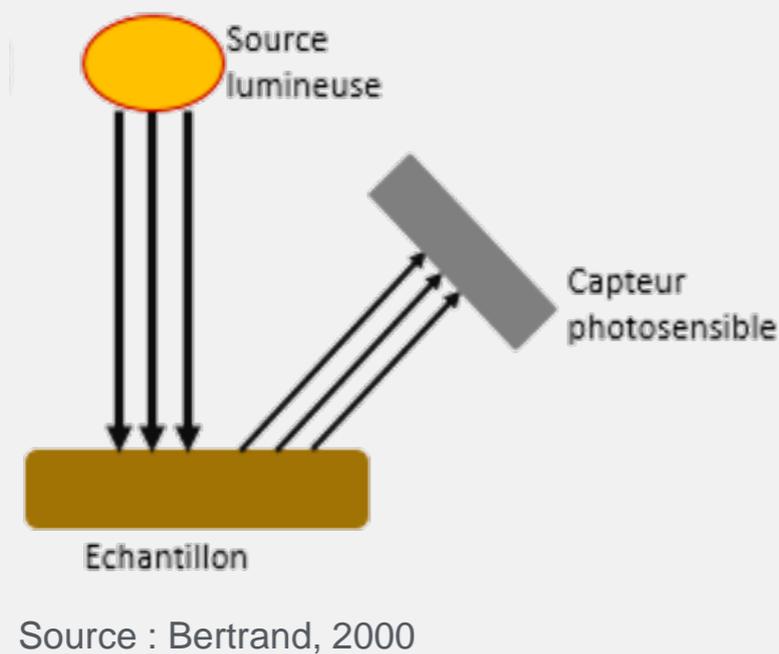
- ➔ Effet du travail en amont (hétérogénéité des sols, nombre d'échantillons en fonction de l'échelle d'étude) ?
- ➔ Résultats sur d'autres variables (tendances ?)

Merci de votre attention



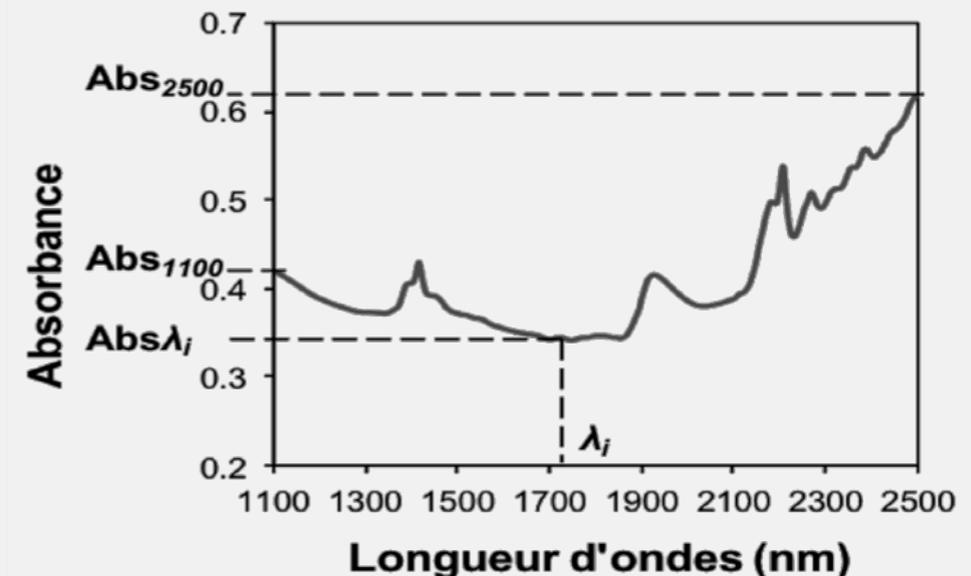


## Principe de la spectroscopie : obtenir des spectres d'absorbance



$$A = \log(1/R)$$

avec A : absorbance  
et R : réflectance



Source : B. G. Barthès

→ Construction de modèles statistiques pour prédire une variable d'intérêt à partir des données spectrales :

étape de calibration primordiale (divers prétraitements spectraux ; divers types de modèles)

## Calibration de modèles

- Jeu d'échantillons caractérisé par :

Des données spectrales

	Abs <sub>350nm</sub>	Abs <sub>352nm</sub>	...	Abs <sub>2500nm</sub>
Ech 1	X <sub>1-350</sub>	X <sub>1-352</sub>	...	X <sub>1-2500</sub>
Ech 2	X <sub>2-350</sub>	X <sub>2-352</sub>	...	X <sub>2-2500</sub>
...	...	...	...	...
Ech n	X <sub>n-350</sub>	X <sub>n-352</sub>	...	X <sub>n-2500</sub>

Une variable analysée selon les méthodes conventionnelles

	SOC content
Ech 1	SOC <sub>1</sub>
Ech 2	SOC <sub>2</sub>
...	...
Ech n	SOC <sub>n</sub>

- Par des régressions multivariées (= PLSR), on résout le système d'équations :

$$SOC_1 = a_{350} \times X_{1-350} + a_{352} \times X_{1-352} + \dots + a_{2500} \times X_{1-2500} + \varepsilon_1$$

$$SOC_2 = a_{350} \times X_{2-350} + a_{352} \times X_{2-352} + \dots + a_{2500} \times X_{2-2500} + \varepsilon_2$$

$$SOC_n = a_{350} \times X_{n-350} + a_{352} \times X_{n-352} + \dots + a_{2500} \times X_{n-2500} + \varepsilon_n$$

Prédiction de SOC sur de nouveaux échantillons

- Pour optimiser les modèles :

Prétraiter les données spectrales (≠ types de spectres)

Transformer la variable analysée, surtout si non-normalité (e.g., log<sub>10</sub>)

Adapter les régressions pour limiter l'effet de l'hétérogénéité (système de pondération / classer les échantillons en groupes homogènes)