

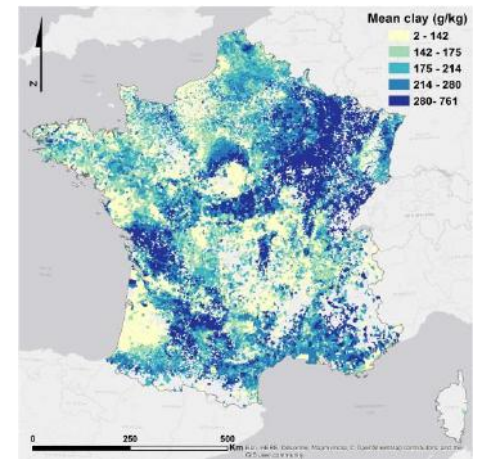
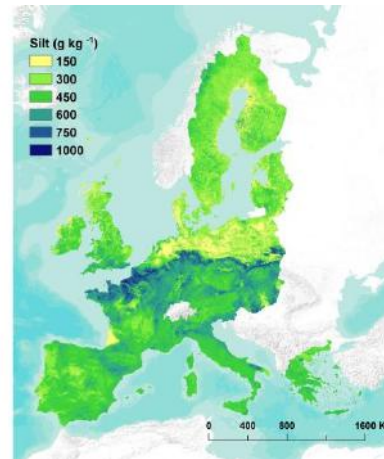


Intérêt du couplage des prédictions spatiales issues de différentes sources de données disponibles en France



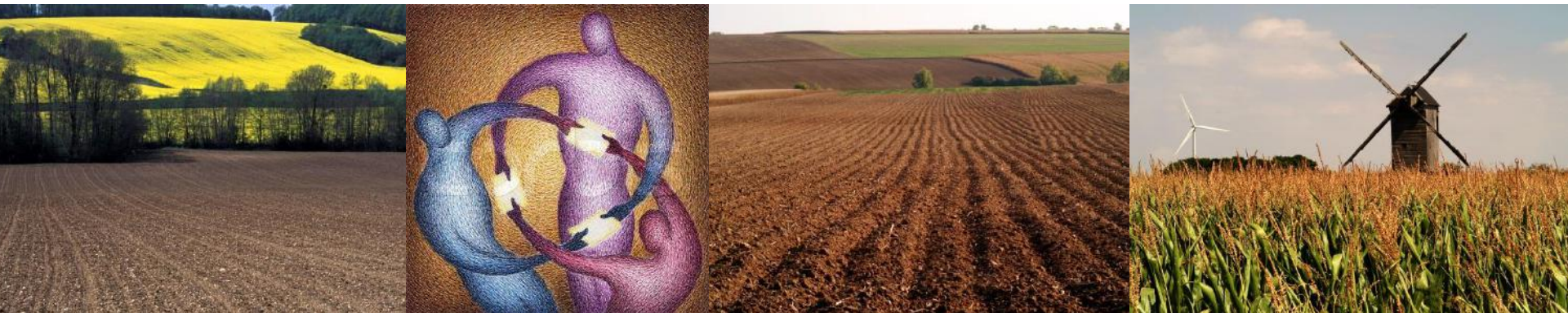
Introduction

- ❖ L'essor de la cartographie numérique des sols a permis de mettre à disposition de nombreuses cartes avec différentes caractéristiques (résolution, étendue, incertitude, etc.)
- ❖ En France, des initiatives dans les régions et au niveau national ont conduit à l'existence de plusieurs produits construits de façon indépendante.
- ❖ Comment peut-on améliorer la précision des prédictions sans faire d'échantillonnage plus intensif et conjuguer les efforts des différents groupes de recherche?



Objectifs

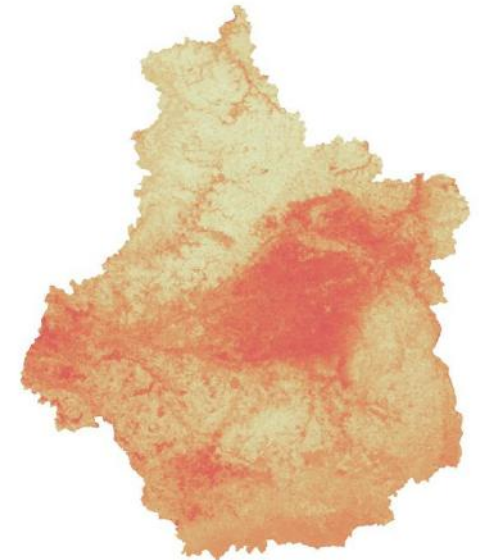
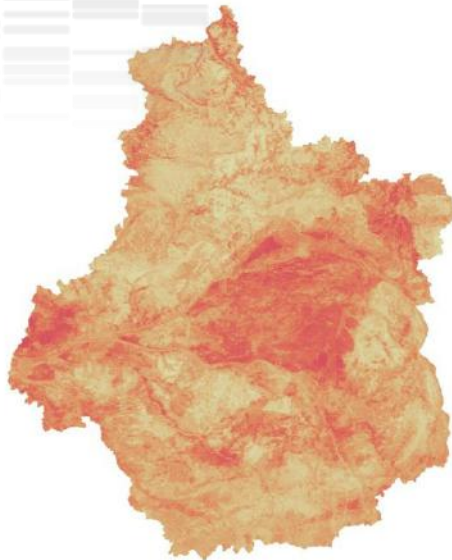
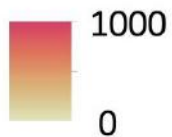
- ❖ Notre objectif était d'étudier l'intérêt et la manière de coupler les différentes prédictions spatiales en vue d'obtenir un produit homogène et d'améliorer la qualité des prédictions disponibles pour le public
- ❖ Prédictions de la texture pour les horizons superficiels des sols cultivés en Région Centre produites avec des données du GIS Sol et européennes



Photos: Jean Weber © INRA

Données – cartes primaires

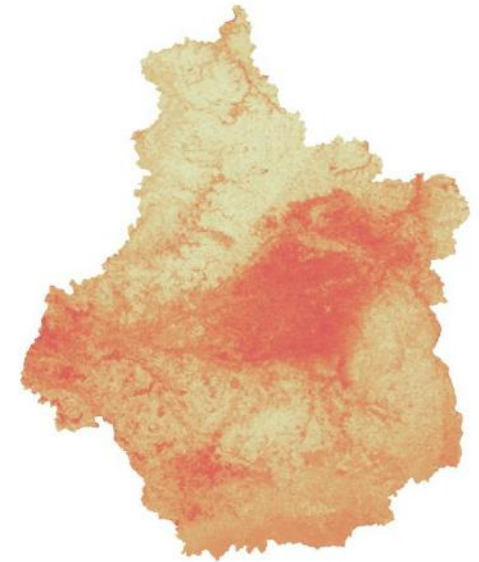
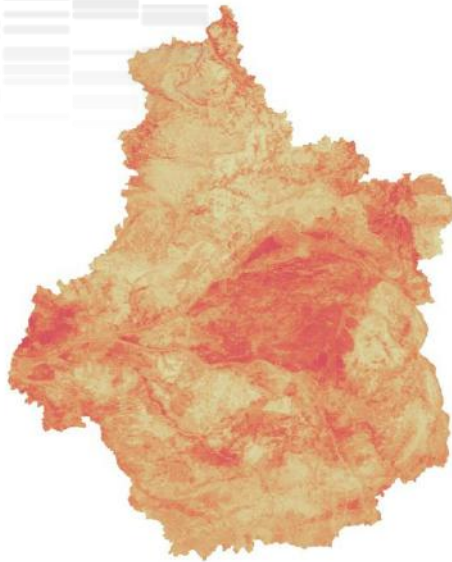
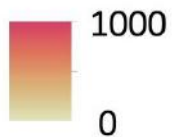
Sable (g kg⁻¹)



	<i>GlobalSoilMap</i> (IGCS)	BDAT	LUCAS
Données de texture	2487 profils / 8718 horizons	Moyenne par commune (M = 1582)	19857 échantillons de surface
Résolution	3-arc second (~ 90 m)	180 m	500 m
Étendue	Région Centre	Région Centre	Europe
Profondeur	5-15 cm	0-25 cm	0-20 cm
	Ciampalini et al., 2014	Román Dobarco et al., 2016	Ballabio et al., 2016

Données – cartes primaires

Sable (g kg⁻¹)

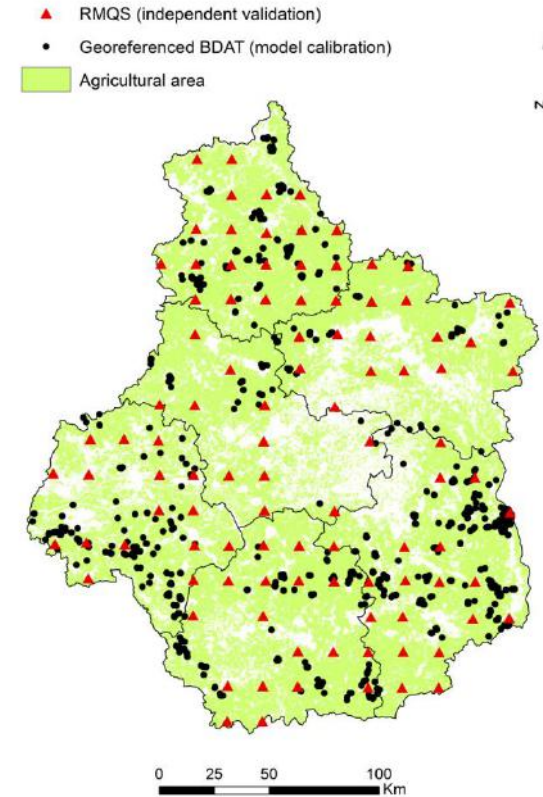


	<i>GlobalSoilMap</i> (IGCS)	BDAT	LUCAS
Occupation des sols	Agricole + forets	Agricole	Agricole + forets
Méthode	GBM + cokriging	Area-to-Point regression cokriging	Multivariate MARS
Incertitude	Modélisé avec GBM	Universal Cokriging variance	Modélisé avec régression (IRLS)
	Ciampalini et al., 2014	Román Dobarco et al., 2016	Ballabio et al., 2016

Méthodes

- ❖ **Hypothèse** : La précision des nouvelles cartes sera au moins comme la meilleure des précisions des cartes primaires
 - ❖ **Deux méthodes de couplage des prédictions spatiales** :
 - ❖ **Granger-Ramanathan** : coefficients calculés avec MCO et données BDAT géoréférencées (n = 717) comme référence. Modèle global, non biaisé.
- $$Y_{geoBDAT} = \beta_0 + (\beta_1 \cdot X_{IGCS}) + (\beta_2 \cdot X_{BDAT}) + (\beta_3 \cdot X_{LUCAS})$$
- ❖ **Bates-Granger ou variance weighed** : pour chaque pixel les coefficients dépendent de la certitude des prédictions.

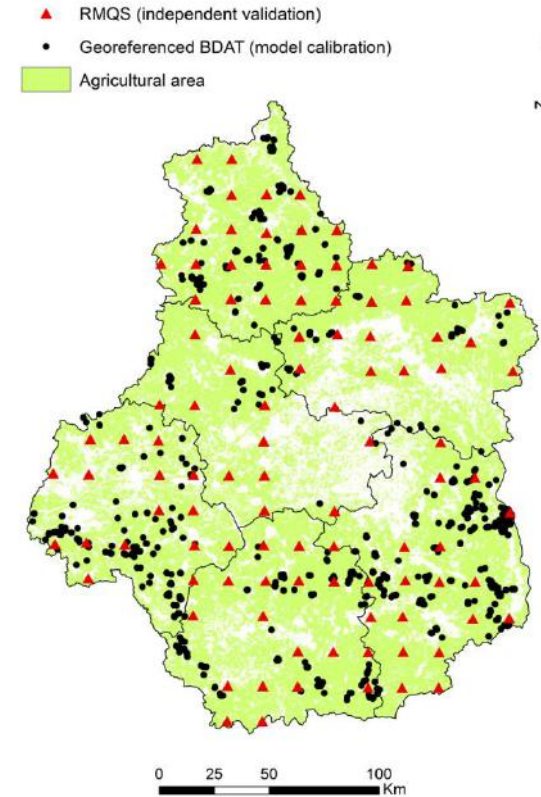
$$W_{ij} = \frac{\frac{1}{\hat{\sigma}_{ij}^2}}{\sum_{j=1}^k \frac{1}{\hat{\sigma}_{ij}^2}}$$



Méthodes

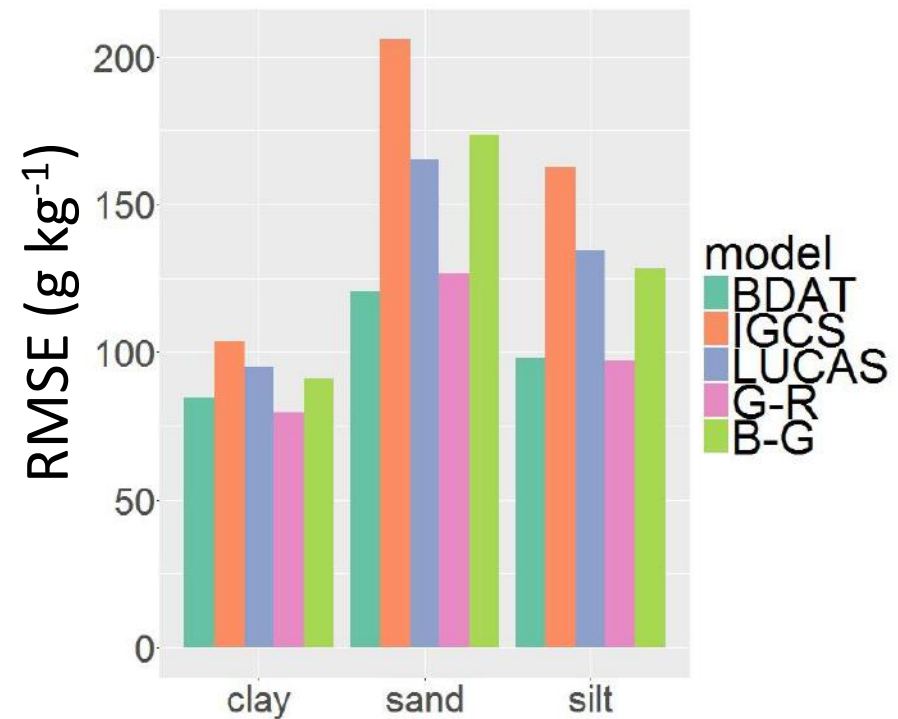
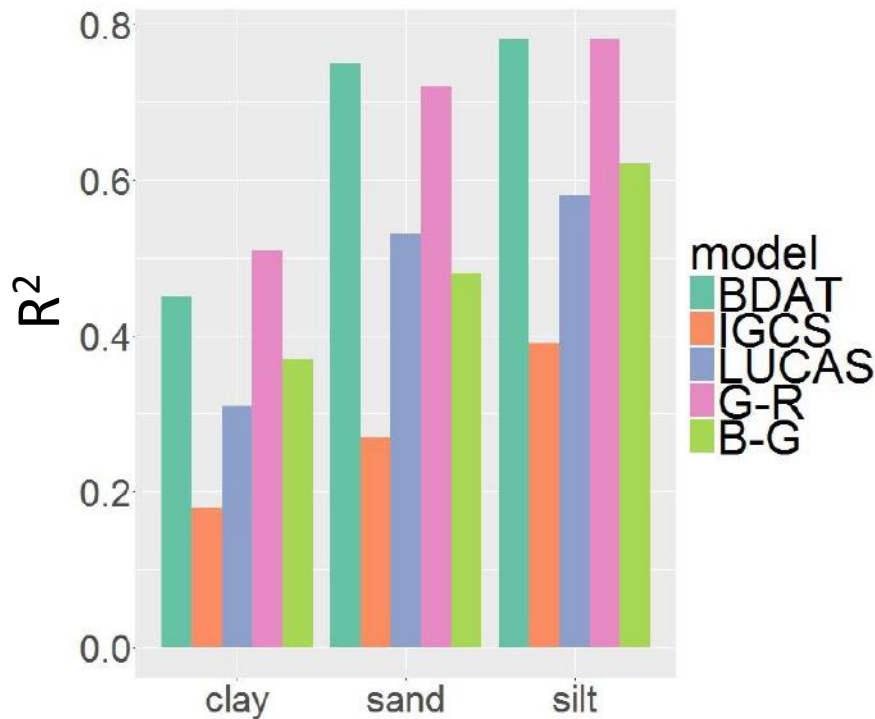
- ❖ Validation indépendante avec données RMQS (n=100)
- ❖ Grille systématique randomisée 16 x 16 km qui permet de calculer des statistiques non biaisées
- ❖ Prediction interval coverage probability (PICP) pour évaluer la validité de l'incertitude
- ❖ Pour un intervalle de prédiction de 90 %, le PICP devrait être aussi de 90 %

$$PICP = \frac{\sum_{i=1}^n (LPL_i < RMQS_i < UPL_i)}{n} \times 100$$



Résultats

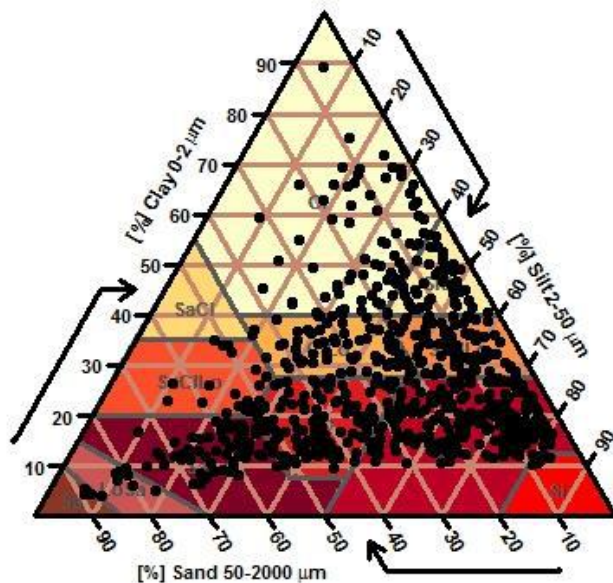
- ❖ BDAT is the best !
- ❖ Les différences entre l'erreur de prédiction globale pour les cartes primaires et les modèles Granger-Ramanathan et Bates-Granger ne sont pas significatives



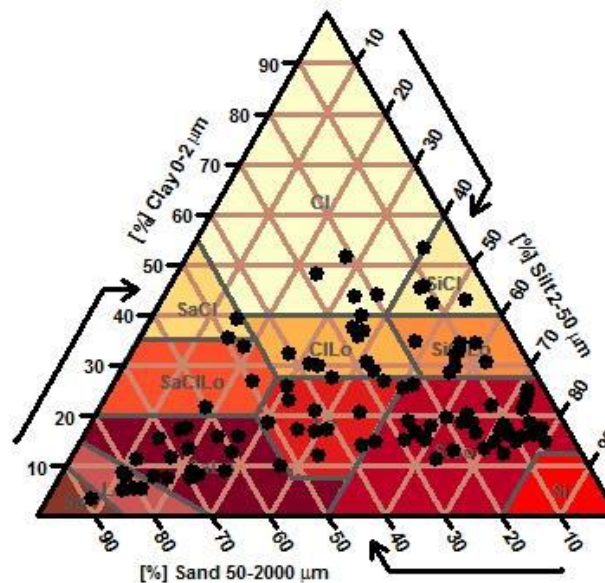
Résultats

- ❖ Poids pour Granger-Ramanathan : $\text{BDAT} > \text{LUCAS} \approx \text{IGCS}$
- ❖ Les données BDAT géoréférencées utilisées pour Granger-Ramanathan couvrent bien les différentes classes de texture, mais leur distribution (un peu agglomérée) ne capture peut-être pas la variabilité spatiale, et donc les coefficients ne sont pas optimaux

BDAT données calibration

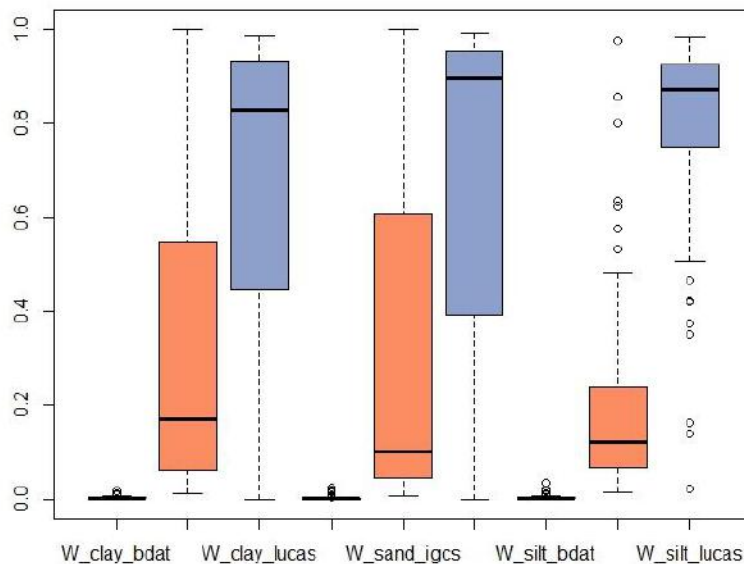


RMQS données validation



Résultats

- ❖ Poids pour Bates-Granger : LUCAS > IGCS >> BDAT
- ❖ Mais LUCAS et IGCS sous-estiment l'incertitude et BDAT a une incertitude « conservative »



	PICP (%)		
	BDAT	IGCS	LUCAS
Argile	96	25	15
Sable	95	14	10
Limons	96	20	2

Conclusions

- ❖ Nous avons produit un cadre statistique pour la combinaison des prédictions spatiales.
- ❖ Malgré sa résolution, la BDAT représente une source d'information essentielle en zone agricole pour l'horizon de surface.
- ❖ La méthode Granger-Ramanathan donne de meilleurs résultats que la méthode Bates-Granger, cependant, aucune méthode de couplage n'a amélioré significativement la précision des prédictions.

Conclusions

- ❖ Une condition nécessaire pour appliquer la méthode Bates-Granger est que les variances pour les cartes primaires soient disponibles et valides !
- ❖ On pourrait calibrer les variances avec une partie des données de validation indépendante ou avec une validation croisée.
- ❖ Le jeu de données pour calculer les coefficients Granger-Ramanathan devrait représenter l'amplitude des textures et sa variabilité spatiale, et suivre la même méthode d'échantillonnage que le jeu de données de validation pour éviter le biais.
- ❖ Une autre façon d'utiliser des données GIS Sol ensemble serait d'utiliser différentes prédictions cartographiques comme variables dans la régression.

Merci de votre attention!



Photo: Jean Weber © INRA