

Impact de différents schémas d'échantillonnage sur la qualité prédictive de fonctions de pédotransfert entre deux réseaux de mesures de la qualité des sols

LOUIS Benjamin¹, SABY Nicolas¹, ARROUAYS Dominique¹, JOLIVET Claudy¹, BOULONNE Line¹, LACARCE Eva¹, ORTON Thomas¹, RATIE Céline¹

¹ : US 1106 INFOSOL, INRA Centre de recherche d'Orléans, 2163 avenue de la Pomme de Pin, CS 40001 ARDON, F45075 ORLEANS CEDEX 2

Bien qu'ils s'inscrivent dans un contexte de décisions de plus en plus internationales, les réseaux de mesures de la qualité des sols sont souvent mis en place à des échelles nationales voire locales. Ces réseaux n'utilisant pas, bien souvent, les mêmes dispositifs d'échantillonnage, il y a une forte demande d'harmonisation des données *a posteriori* dans le but de pouvoir répondre aux questions soulevées par les réglementations à l'échelle globale. Certaines études sur l'harmonisation des données, tel que le projet ENVASSO, ont mis en avant l'importance de la mise en place de sites de mesures où ces différents dispositifs sont installés en parallèle afin de calibrer des fonctions de pédotransfert entre les indicateurs mesurés selon les différents protocoles. Cependant, le choix du nombre et de la localisation de ces sites de calibration reste une question clé dans la méthodologie d'harmonisation des données *a posteriori*.

Les sols forestiers français ont été échantillonnés selon deux dispositifs d'échantillonnage et protocoles analytiques différents sur une même grille de maille 16*16 km, au travers du Réseau de Mesures de la Qualité des Sols (RMQS) (Jolivet et al., 2006) à l'échelle française et du réseau européen de surveillance des sols forestiers (ICP Forest niveau 1, projet BioSoil). Dans ce travail, nous comparons ces deux dispositifs à l'aide de certaines variables mesurées (*e.g.* carbone organique, potassium total, plomb total versus extractible et pH) en traitant la question du nombre et de la localisation des sites de calibration par leur impact sur la stabilité des fonctions de pédotransfert en terme de qualité de prédiction. Après avoir construit les fonctions de pédotransfert en utilisant une méthode de sélection de modèle sur l'ensemble des sites des deux réseaux, plusieurs stratégies d'échantillonnage sont testées selon un nombre décroissant de sites échantillonnés telles que l'échantillonnage aléatoire simple (RS) ou l'échantillonnage par hypercube latin conditionné (cLHS et DLHS). Cette dernière est une méthode qui assure que les variables prédictives des fonctions de pédotransfert dans l'échantillon, aient une distribution proche de leur distribution sur l'ensemble des sites (Minasny et McBratney, 2006 et 2010). Les fonctions de pédotransfert sont de nouveau calibrées en utilisant les échantillons sélectionnés et un critère de qualité prédictive, l'erreur quadratique moyenne de prédiction (RMSEP), est ensuite calculé. Les résultats montrent qu'à des tailles d'échantillons élevées, le DLHS permet d'augmenter la qualité prédictive des fonctions de pédotransfert. De plus, bien qu'en moyenne la qualité prédictive des fonctions calibrées à l'aide des échantillons issus du cLHS ne soit que très légèrement meilleure, le nombre de répétitions conduisant à une mauvaise qualité prédictive (RMSEP très élevé) est quant à lui largement diminué, en particulier avec la diminution de la taille de l'échantillon où le DLHS tend à donner les mêmes résultats que le cLHS voir de plus mauvais (figure 1).

Mots clefs : harmonisation a posteriori, réseaux de mesures, RMQS, BioSoil, qualité des sols, échantillonnage, RMSEP

Carbone organique

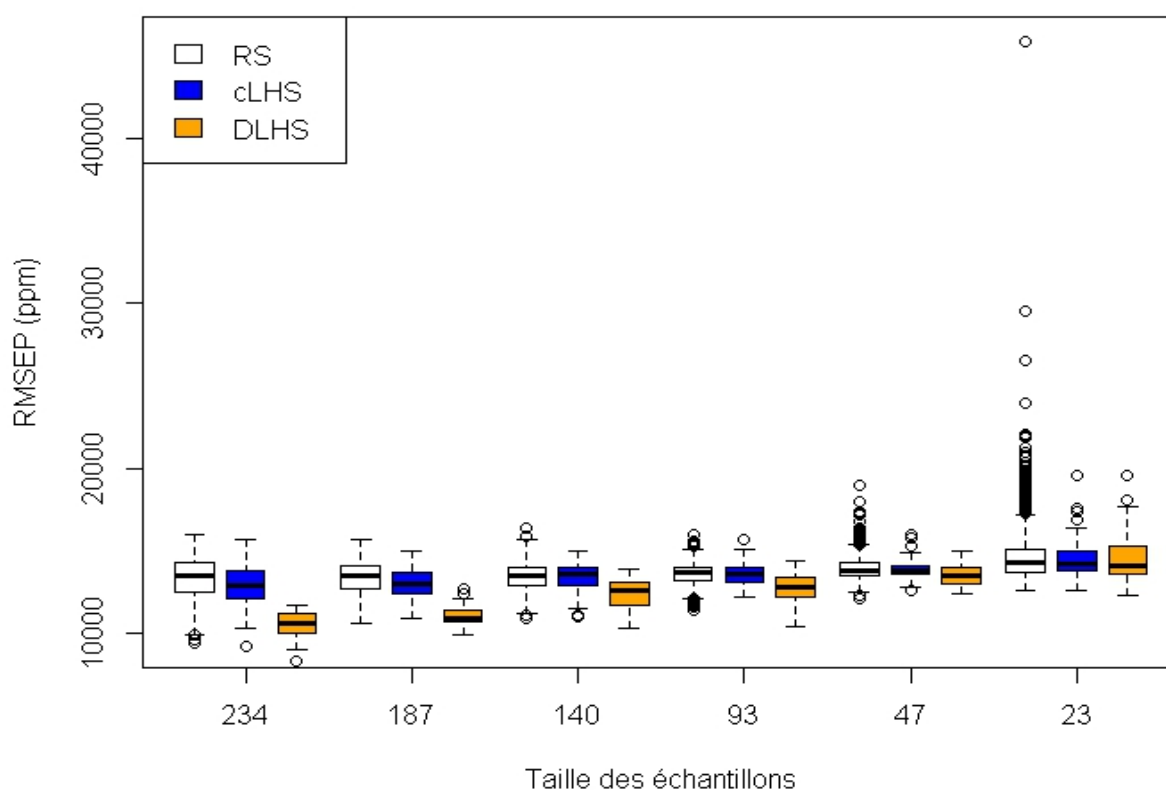


Figure 1 : Distribution de l'erreur quadratique moyenne de prédiction en fonction de la stratégie d'échantillonnage et de la taille de l'échantillon

Références :

Jolivet, C., Arrouays, D., Boulonne, L., Ratié, C. & Saby, N. 2006. Le Réseau de Mesures de la Qualité des Sols de France (RMQS). Etat d'avancement et premiers résultats. Etude et Gestion des Sols, 13, 149-164.

Minasny B, McBratney AB (2006) A conditioned Latin Hypercube method for sampling in the presence of ancillary information. Computational Geosciences 32:1378-1388