HAL
open science

Prédiction des propriétés du sol à l'échelle nationale à partir des données du RMQS : Exemple des Éléments Trace Métalliques

Nicolas N. Saby, Claudy C. Jolivet, Line Boulonne, Céline Ratié, Ben P. Marchant, R. Murray Lark, Jean Thioulouse, Dominique D. Arrouays

To cite this version:

Nicolas N. Saby, Claudy C. Jolivet, Line Boulonne, Céline Ratié, Ben P. Marchant, et al.. Prédiction des propriétés du sol à l'échelle nationale à partir des données du RMQS : Exemple des Éléments Trace Métalliques. 10. Journées d'Etude des Sols, May 2009, Strasbourg, France. hal-02757800

HAL Id: hal-02757800

https://hal.inrae.fr/hal-02757800

Submitted on 4 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prédiction des propriétés du sol à l'échelle nationale à partir des données du RMQS : Exemple des Eléments Trace Métalliques

SABY N.P.A.[1], JOLIVET C.[1], BOULONNE L.[1], RATIE C.[1], MARCHANT B.P.[2], LARK R.M.[2], THIOULOUZE J.[3], et ARROUAYS D.[1]

[1] : INRA INFOSOL, US 1106, 2163 Avenue de la Pomme de Pin, CS 40001 ARDON, 45075 ORLEANS Cedex 2
[2] : Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK
[3] : Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, FRANCE

En 2006, La Commission Européenne a adopté une stratégie thématique et une proposition de directive cadre (European Communities, 2006) visant à assurer le maintien de la qualité des sols européens et leur capacité à supporter l'activité humaine et les écosystèmes. Ce travail a permis d'identifier les principales menaces pesant sur les sols, comme par exemple la contamination des sols par les éléments traces métalliques. La mise en place d'une telle politique de protection des sols et l'identification des zones à risques ont pour conséquence la nécessité de disposer d'informations précises et distribuées sur la variabilité des propriétés des sols dans l'espace et dans le temps. En d'autres termes, ces politiques impliquent la mise en place de dispositifs de mesure de la qualité des sols à forte résolution spatiale. De plus, les données collectées par ce type de réseau doivent être analysées par des méthodes statistiques adaptées à la nature et l'extension spatiale de l'information manipulée afin d'offrir des prédictions fiables sans artefact de calculs.

Ainsi, le nombre d'observations et leur extension spatiale à l'échelle d'un pays se traduisent par des défis à relever pour leur analyse statistique. Le nombre d'observations (supérieur à 2000 pour la France) pose d'abord des difficultés pour la mise en œuvre de techniques géostatistiques demandant beaucoup de ressource en calcul, en particulier les méthodes REML (Lark & Cullis, 2004). L'extension spatiale quant à elle se traduit par la possibilité de rencontrer des contextes pédo-climatiques contrastés conduisant à des structures spatiales hétérogènes des propriétés du sol. En outre, l'impact des pollutions locales peut biaiser l'analyse des structures spatiales à plus petite échelle. Ces données sont souvent nommées « outliers » et conduisent notamment à la surestimation de la variance de la propriété du sol étudiée. Enfin, le nombre de propriétés mesurées en chaque observation offre la possibilité d'étudier l'échelle des relations spatiales entre les propriétés du sol.

L'objectif de ce travail est de proposer des méthodes d'analyses statistiques adaptées aux données issues du programme Réseau de Mesures de la Qualité des Sols (RMQS) permettant de cartographier les pollutions diffuses en ETM au niveau national. Pour cela, il se fonde sur la prise en compte du contexte pédoclimatique et des outliers dans le calcul des modèles de distribution spatiale des ETM.

Le jeu de données extrait de la base de données du RMQS comporte plus de 2000 sites localisés sur le territoire national selon une grille régulière de 16 km de côté. Sur chaque site RMQS, deux échantillons composites (issus de 25 prélèvements) sont prélevés selon un échantillonnage non aligné sur une surface d'échantillonnage de 20 x 20 m. Le premier échantillon est prélevé dans la couche supérieure du sol (couche travaillée ou 0-30 cm) et le second dans la couche de sol sous-jacente (limite inférieure de la couche travaillée jusqu'à 50 cm ou 30-50 cm). Les teneurs totales de 8 ETM (Cd, Cr, Cu, Co, Ni, Mo, Pb, Zn) ont été déterminées après extraction par HF-$HClO_4$ et 5 ETM (Cd, Cr, Cu, Pb, Zn) par extraction par l'EDTA.

Dans ce travail, nous proposons premièrement de modéliser la teneur en ETM à l'échelle nationale à l'aide des modèles linéaires à effets mixtes (MLM) (Lark & Cullis, 2004).

Nous nous fondons sur un modèle à 3 composantes. La première composante correspond au déterminisme du matériau parental ou de l'occupation du sol et, est représentée par l'effet fixe du MLM. La deuxième composante correspond aux variations au sein de chaque modalité de l'effet fixe et est représentée par l'effet aléatoire du MLM. La troisième composante correspond au processus spatial secondaire issu de la présence d'outliers dans le jeu de données. Elle nécessite pour sa modélisation d'employer des techniques robustes (Lark, 2002) permettant d'identifier ces outliers. Ils peuvent ainsi, soit être éliminés du jeu de données (Rawlins et al., 2006), soit « winsorizer » (Hawkins & Cressie, 1984). Cette dernière technique consiste à réévaluer la valeur des outliers en éliminant la 3e composante tout en maintenant les 2 premières.

Dans un deuxième temps, nous avons étudié les relations spatiales entre les éléments trace par la mise en œuvre d'analyses multivariées sous contraintes spatiales (Thioulouse J., 1995). Une relation de voisinage des observations est traduite en termes de pondération (nombre de voisin en chaque site) et introduite par la suite dans le calcul de l'ACP. Elle permet de mettre en évidence des structures globales par le calcul de composantes cartographiables.

Les résultats de ce travail montrent l'intérêt de disposer de bases de données distribuées à l'échelle nationale pour la cartographie des contaminations des sols en ETM. Ils montrent aussi l'importance de s'appuyer sur des techniques statistiques adaptées à la nature de l'information et notamment la présence d'outliers.

Bibliographie

European Communities (2006). Proposal for a directive of the European Parliament and of the Council. Establishing a Framework for the Protection of Soil and Amending Directive 2004/35/EC.COM(2006) 232.2004/35/EC.COM(2006)232. European Commission, Brussels, Belgium.

Hawkins, D.M., Cressie, N. (1984). Robust kriging - A proposal. Journal of the International Association for Mathematical Geology, 16, 3–18.

Lark, R.M., 2002. Modelling complex soil properties as contaminated regionalized variables. Geoderma, 106, 799–813.

Lark, R.M., Cullis B.R. (2004). Model based analysis using REML for inference from systematically sampled data on soil. European Journal of Soil Science, 55, 799–813.

Rawlins, B.G., Lark, R.M., Webster, R., O'Donnell, K.E. (2006). The use of soil survey data to determine the magnitude and extent of historic metal deposition related to atmospheric smelter emissions across Humberside, UK. Environmental Pollution, 143, 416–426.

Thioulouse, J., Chessel, C. and Champely S. (1995). Multivariate analysis of spatial patterns: a unified approach to local and global structure. Environmental and Ecological Statistics, 2, 1-14.