

## DOCUMENT 1

### Le projet « Clés de Sol » : Contexte et objectifs - argumentaire scientifique

---

Depuis une dizaine d'années, le sol est au cœur de débats sociétaux majeurs : production agricole, défi alimentaire, changements climatiques, artificialisation, gestion de l'eau, santé, biodiversité, etc. Des événements internationaux comme la journée annuelle mondiale des sols du 5 décembre mise en place en 2002 (Union internationale des sciences du sol), l'année internationale des sols en 2015 (ONU), ont souligné l'importance du sol. La notion de service écosystémique s'est imposée récemment. Le sol est un écosystème qui contribue à rendre de multiples services à l'humanité (Global Soil Partnership, FAO, 2013). C'est une ressource indispensable et fragile.

La connaissance des sols et de leurs propriétés est un prérequis pour caractériser les fonctions et les services écosystémiques des sols, éclairer les décisions qu'il faut prendre à l'échelle locale en matière d'agriculture, d'environnement et d'aménagement du territoire (Sanchez et al, 2009). Les bases de données et les cartes des sols auxquelles ces bases conduisent, sont le plus souvent à une résolution spatiale trop faible pour être utilisées dans ce but. Ainsi, la France n'est couverte quasi-exhaustivement que par des cartes départementales ou régionales au 1/250 000, les Référentiels Régionaux Pédologiques (Voltz et al, 2018). Ces cartes, élaborées avec relativement peu d'observation de sol (en moyenne un profil - ou fosse - pour 40 km<sup>2</sup> un sondage pour 400 ha), ne délimitent que de grands ensembles – polygones - (taille moyenne de 3 km<sup>2</sup>). Elles représentent donc les variations des sols et de leurs propriétés avec de trop faibles précisions pour aider à des décisions aux échelles locales (de la parcelle à la commune en passant par l'exploitation agricole ou le petit bassin versant).

Pour répondre à cette demande d'une meilleure connaissance des sols, la communauté scientifique a développé une nouvelle approche de cartographie des sols, appelée en France cartographie des sols à base de modélisation statistique (CSMS) (Lagacherie et al, 2013, Voltz et al, 2018). Le principe consiste à estimer des classes de sol ou leurs propriétés à partir des données disponibles sur la zone à étudier et de données spatiales représentant des éléments du paysage. Des modèles statistiques du type « Scorpan » ( $Sol = f(S : \text{connaissance sol a priori}, C : \text{Climat}, O : \text{Organismes}, \text{Relief}, P. \text{Matériau parental}, A : \text{Age et } N : \text{position géographique})$ ) sont déployés (Mc Bratney, 2003). Les travaux scientifiques récents permettent d'estimer une incertitude et de prendre en compte d'autres données spatiales récentes (base de données du relief, images de télédétection, données de capteurs des propriétés du sol). Ces prédictions sont le plus souvent réalisées par des fonctions de prédiction induites par des

algorithmes d'apprentissage automatique ou des modèles géostatistiques. Le grand intérêt de ces modèles est qu'ils peuvent être constamment améliorés par l'ajout de nouvelles observations du sol et de nouvelles données spatiales au fur et à mesure des avancées technologiques.

Sous l'impulsion d'une communauté scientifique organisée au sein d'un groupe de travail de l'Union Internationale de science du Sol, la CSMS apparaît comme une solution opérationnelle de cartographie des sols. Un projet international GlobalSoilMap (Arrouays et al, 2014) s'est mis en place pour prédire, sur 80% de la planète, les propriétés de sol les plus courantes sur une grille de 90 m de côté et à différentes profondeurs. Ce programme a été appliqué sur de grands bassins versant (ex. Nussbaum et al, 2018), des régions (ex. Vaysse et al, 2015), des pays (ex. Mulder et al, 2016) ou même à l'échelle mondiale (Hengl et al, 2017). A l'examen des résultats, il apparaît clairement que, malgré leur forte résolution spatiale affichée, leurs précisions restent trop faibles pour pouvoir fournir aux décideurs une connaissance suffisante des sols pour aider leur décision. En effet, la part de variabilité des sols reproduites par les produits GSM ( $R^2$ ) restent faibles (moins de 50% pour 80% des propriétés de sols prédites).

L'analyse fine des performances des différents produits GSM obtenus sur la région Languedoc-Roussillon (Vaysse et al, 2015) a montré que la densité spatiale des observations constituait le facteur limitant les performances de ces modèles. Ceci a été confirmé par plusieurs études récentes qui ont testé la sensibilité des performances des modèles aux densités d'échantillonnage (Somarathna et al., 2017 ; Wadoux et al., 2019 ; Lagacherie et al, 2020) mais aussi à la répartition spatiale de ces échantillons (Lagacherie et al, 2020). Toutes ces études montrent que les performances des modèles CSMS s'accroissent avec la densité d'échantillonnage, une répartition régulière dans l'espace et représentative des distributions de la propriété ciblée. Lagacherie et al (2020) ont montré que des performances élevées compatibles avec des prises de décision à l'échelle locale pouvaient être obtenues par ces approches, à condition que les observations soient en quantité et en qualité suffisantes. Ces résultats théoriques ont trouvé confirmation dans un travail récent où des performances élevées ont été enregistrées ( $R^2 = 0.7$  soit 70% de variabilité cartographiée) pour cartographier le Réservoir Utile du sol sur une commune d'un périmètre irrigué de la plaine littorale Languedocienne (Thèse Q. Styc, 2020), utilisant des données pédologiques anciennes très denses (4 profils de sol et 38 sondages pour 1 km<sup>2</sup>). Ainsi, la densité et la régularité des échantillonnages constitue la première condition pour que la CSMS concrétise les espoirs de fournir des cartes des sols et de leurs propriétés utilisables de manière effective pour l'aide à la décision locale.

A côté de la nécessaire récupération des données pédologiques anciennes et du développement des capteurs permettant d'estimer à coûts limités certaines propriétés de sol, une approche de collecte des données sur les sols par des citoyens apparaît comme une voie complémentaire et prometteuse pour alimenter les modèles de CSMS (« CSMS citoyenne »). Cette nouvelle voie a pour avantage de permettre au citoyen de participer à la construction de bases de données et de cartes de sol, gage également de meilleure appropriation des documents et des décisions futures qui peuvent en découler. Il n'existe à ce jour dans la littérature aucun exemple de mise en œuvre d'une CSMS

citoyenne. Des questions méthodologiques nouvelles se posent pour une CSMS participative : comment bâtir des protocoles de collecte pour des données fiables sur les sols ? comment intégrer ces données nouvelles en tenant compte de leurs incertitudes ? comment orienter le choix de sites et d'échantillonnages pour obtenir les meilleures performances avec le moins d'observations possible ?

Enfin se pose la question de savoir sur quelles propriétés de sol peut et doit porter l'effort d'une CSMS participative. Dans un premier temps, le choix est fait de se focaliser sur des propriétés de sol à la fois importantes pour la prise de décision mais aussi pour lesquelles il existe déjà des bases de données qu'il convient d'enrichir. Ceci amène à privilégier des propriétés pérennes du sol traduisant un potentiel invariant dans le temps. Le groupe de propriétés sélectionnées dans le projet GlobalSoilMap constitue une première base. Le projet Clés de Sol permettra d'affiner cette sélection en identifiant les propriétés qui peuvent effectivement être collectées par les citoyens avec une bonne fiabilité. Les travaux pourront être dans un second temps être étendus à d'autres propriétés comme la biodiversité du sol.

Au début de chaque fiche protocole regroupée dans le document 3, une explication sur l'intérêt de la propriété étudiée est présentée.

### **Pour en savoir plus**

Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B. M., Hong, S. Y., Lagacherie, P., Lelyk, G., McBratney, A. B., McKenzie, N. J. (Auteur de correspondance), Mendonca-Santos, M. D. L., Minasny, B., Montanarella, L., Odeh, I. O. A., Sanchez, P. A., Thompson, J. A., Zhang, G.-L. (2014). GlobalSoilMap : toward a fine-resolution global grid of soil properties. *Advances in Agronomy*, 125, 93 – 134

Hengl, T., De Jesus, J.M., Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* 12, 1–40.

Lagacherie, P., Walter, D.A.C., 2013. Cartographie numérique des sols : principe, mise en œuvre et potentialités. *Etude et Gestion des Sols* 20, 83–98. [http://www.afes.fr/wp-content/uploads/2017/09/EGS\\_20\\_1\\_EGS\\_20\\_1\\_CN\\_Lagacherie.pdf](http://www.afes.fr/wp-content/uploads/2017/09/EGS_20_1_EGS_20_1_CN_Lagacherie.pdf)

Lagacherie, P., Arrouays, D., Bourennane, H., Gomez, C. Nkuba Kasanda, L., 2020. Analysing the impact of soil spatial sampling on the performances of Digital Soil Mapping models and their evaluation: a numerical experiment on Quantile Random Forests using clay contents obtained from Vis-NIR-SWIR hyperspectral imagery. A paraître dans *Geoderma*.

McBratney, A.B., Santos, M.L.M and Minasny B., (2003). *On digital soil mapping*. *Geoderma* 117: 3-52

- Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Arrouays, D., 2016. *GlobalSoilMap* France: high-resolution spatial modelling the soils of France up to two meter depth. *Sci. Total Environ.* 573:1352-1369. <http://dx.doi.org/10.1016/j.scitotenv.2016.07.066.tb01630.x>
- Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M.E., Papritz, A., 2018. Evaluation of digital soil mapping approaches with large sets of environmental covariates. *Soil* 4, 1–22.
- Sanchez, P.A., Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., De Lourdes Mendonça-Santos, M., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vågen, T.-G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., Zhang, G.-L., 2009. Digital soil map of the world. *Science* (80-. ). 325.
- Somarathna, P.D.S.N., Minasny, B., Malone, B.P., 2017. More Data or a Better Model? Figuring what Matters Most for the Spatial Prediction of Soil Carbon. *Soil Sci. Soc. Am. J.* 81, 1413–1426.
- Styc, K. Cartographie numérique du RU à partir de données pédologiques anciennes. Thèse de doctorat. A soutenir .
- Vaysse, K., Lagacherie, P., 2015. Evaluating Digital Soil Mapping approaches for mapping *GlobalSoilMap* soil properties from legacy data in Languedoc-Roussillon (France). *Geoderma Reg.* 4, 20–30.
- Voltz, M., Arrouays, D., Bispo, A., Lagacherie, P., Laroche, B., Lemerrier, B., Richer-de-Forges, A.C., Sauter, J., Schnebelen, N., 2018. La cartographie des sols en France: Etat des lieux et perspectives, INRA, Paris, France, 112 pages.
- Wadoux, A.M.J., Brus, D.J., Heuvelink, G.B.M., 2019. Sampling design optimization for soil mapping with random forest. *Geoderma* 355.