

Mise à jour des références relatives aux diagnostics de fertilisation P des sols en France

Mémoire de master 2 réalisé par *DEROULLEDE Simon*

Encadrants : Lionel JORDAN-MEILLE, et Pascal DENOROY, Enseignant-Chercheur et Ingénieur de Recherche, UMR 1391 ISPA, Bdx Sc Agro, INRAE

Comité de suivi : Sarah RANOARISON (Laboratoire départemental d'analyses et de recherche de l'Aisne), Gildas COTTEN (UNIFA), Khady DIEDHIOU (COMIFER), Daniel HANOCQ (Chambre d'Agriculture de Bretagne), Paul TAUVEL (ITB), et Grégory VERICEL (ARVALIS).

Dates du stage : Juin-Août 2025

RESUME

On peut qualifier les seuils de fertilisation P français comme « sécuritaires » pour les grandes cultures comparés aux autres pays de l'UE. La recommandation en P du COMIFER se base sur une réponse culturale en fonction de types de sols établis. Dans ce travail il est vu que la teneur en éléments fixateurs de P et, sa quantification par statistiques géoréférencées semble être une bonne piste d'amélioration de l'incertitude autour de la dose de P recommandée.

Ce travail s'axe autour de l'amélioration des types de sols dans la recommandation en fertilisation P du COMIFER. L'apparition récente d'essais agronomiques sur le P avec des analyses de sols donnant la quantification en Fe et Al (fixateurs du P dans le sol) nous a fait voir une corrélation entre ces éléments et le seuil de fertilisation en P. Après plusieurs tentatives sur diverses méthodologies, l'interpolation de valeurs pédologiques géoréférencées (disponibles parfois en source ouverte) semble le meilleur moyen de répondre au défi d'amélioration des types de sols (donc de précision de la recommandation), le manque de caractérisation pédologique d'essais agronomiques étant le principal facteur limitant.

INTRODUCTION

Le COMIFER (Comité Français d'Études et de Développement de la Fertilisation Raisonnée) regroupe tous les acteurs de la fertilisation. En France, les seuils de fertilisation du P utilisés pour orienter les recommandations agronomiques datent de la brochure PK COMIFER de 1995 et apparaissent aujourd'hui comme sécuritaires d'un point de vue international (Jordan-Meille et al. 2024), et sans traçabilité. Or, la gestion du P représente un enjeu économique et environnemental majeur, conditionnant à la fois les coûts de production et la préservation des ressources. Dans ce contexte, un travail de révision de ces seuils à partir d'une base de données (BDD) nationale d'essais de fertilisation P du COMIFER est mené dans le but d'en assurer la mise à jour, la fiabilité statistique et la traçabilité méthodologique. L'objectif est d'apporter une meilleure caractérisation des types de sols pour, en sortie, diminuer l'incertitude autour de la recommandation en fertilisation P.

Le P du sol existe sous forme organique et inorganique, cette dernière fraction représentant 35–70 % du P dans le sol (Van Doorn et al. 2024). Dans les sols non calcaires, l'adsorption sur les oxydes de Fe et d'Al est le principal mécanisme de régulation. Toujours dans ces sols non calcaires, la fraction dissoute, directement absorbée par les plantes, est minime (<0,015 %) et constamment renouvelée par dé-

sorption, dissolution minérale et minéralisation. Le taux de désorption peut limiter la nutrition des cultures.

Les seuils de fertilisation sont déterminés par le modèle linéaire-plateau, qui identifie la teneur en P du sol au-delà de laquelle le rendement n'augmente plus. Les seuils de fertilisation utilisés ici sont ceux issus de divers essais en condition plein-champ menés par divers acteurs de recherche agronomique depuis ces 50 dernières années.

L'objectif général de ce travail est de contribuer à redéfinir les valeurs des seuils de fertilisation sur la base de tous les essais P disponibles en France, et recensés par le COMIFER. L'objectif secondaire est d'assurer de la traçabilité sur l'établissement des nouveaux seuils.

Première approche statistique - Random Forest

L'utilisation de Random Forest semble être une solution très efficace dans notre cas de statistiques exploratoires. Elle permet notamment d'identifier les variables avec le plus d'importance dans l'arbre de décision pour la prédiction d'une variable définie (le seuil de fertilisation en l'occurrence).

L'application de la méthode statistique aux données de seuils du COMIFER a nécessité un long travail de nettoyage pour rendre la base exploitable, homogène et traçable. L'objectif était d'explorer, via des outils comme Random Forest et les corrélations (corrplot), les liens entre variables pédologiques et seuils de fertilisation.

Deux sous-bases ont été analysées : un premier sous-groupe de données (TABLE1.1 avec n=323, variables classiques : granulométrie, MO, CEC, P, CaCO_3) et un second (TABLE1.2 avec n=48, enrichie avec quantification de Fe et Al). Les effectifs sont le principal facteur limitant : en dessous de 50 échantillons, les prédictions deviennent aberrantes. Les meilleurs résultats concernent les groupes à fort effectif.

La variable la plus déterminante est de loin la mesure Joret-Hébert du P_2O_5 , puis d'autres variables fixatrices du P dans le sol (Al, Fe, CEC, MO etc...).

La surperformance de la mesure Joret-Hébert du P_2O_5 peut s'expliquer soit par un biais de redondance avec la recommandation de fertilisation, soit par sa capacité à refléter le pouvoir tampon en P du sol, hypothèse appuyée par la littérature.

Globalement, la démarche met en évidence les limites imposées par les effectifs réduits et les limites de l'imputation par la médiane des

données manquantes.

Seconde approche statistique

Dans cette seconde partie nous compilons deux autres démarches exploratoires ayant pour objectif de mieux caractériser le seuil de fertilisation.

La première procédure explore l'efficacité potentielle d'un indice trouvé dans la littérature (le PSD ou degré de saturation en P) vis-à-vis du lien avec le seuil de fertilisation sur une partie de notre jeu de données.

La seconde procédure cherche à vérifier la pertinence de la variable de type de sol ('sol_COMIFER95') qui est déterminante dans la caractérisation des seuils de fertilisation, mais qui a été établie sans traçabilité.

1- Dans le travail de Van Doorn et. Al. (2024), les auteurs utilisent le PSD (Phosphorus Saturation Degree) en tant que variable de prédiction de Random Forest pour prédire une teneur en P dans le sol avec une bonne performance.

Nous émettons deux hypothèses importante servant à faire le lien entre ce travail et notre BDD COMIFER :

Le P oxalate (utilisé dans le PSD) est comparable au P Joret-Hébert : Les deux méthodes sont en réalité différentes en un point, l'extractant est le même (oxalate) mais le pH d'extraction diffère.

Le Fe oxalate et Al oxalate sont comparables aux Fe Tamm et Al Tamm : Il est stipulé dans la littérature que les méthodes semblent sensiblement les mêmes.

$$PSD = \frac{P_{Joret-Hébert} \times 0.437 \cdot 10^{-4}}{\alpha_{max} \times (Fe_{Tamm} + Al_{Tamm})} \times 100$$

Dans la suite de ce travail exploratoire, nous essayons d'établir la corrélation entre le PSD et le seuil de fertilisation.

La régression linéaire simple entre PSD et seuil de fertilisation (clx_olsen) montre certes une corrélation significative mais faible ($R^2 = 0,27$), et aucune relation robuste n'est trouvée.

Le PSD, et plus largement la création d'un indicateur mathématique du pouvoir tampon du sol apparaît donc comme une piste prometteuse, mais les effectifs et la structure des données empêchent de confirmer sa pertinence.

2- On cherche globalement à évaluer la pertinence de la variable « type de sol » établie sans traçabilité il y a plus de 25 ans vis-à-vis du reste du jeu de données, ceci conforterait la démarche statistique précédemment établie.

Une analyse de classification hiérarchique teste la fiabilité de la variable « sol_COMIFER95 » en fonction des variables numériques De notre BDD. Le nombre de groupes est choisi via la fonction NbClust(), et le résultat montre la création de 3 clusters d'échantillons ayant des types de sols qui suivent bien le gradient pédologique relatif au P (calcaire → limoneux), validant la cohérence du diagnostic et sa pertinence pour les analyses statistiques, malgré le manque de traçabilité.

Approche géoréférencée

Une troisième approche expérimentale autre que la méthode statistique est celle des géostatistiques. La disponibilité en source ouverte de bases de données d'échantillonnages de sols pour la France (RMQS) élargit les possibilités quant à la caractérisation des propriétés du sol (Jolivet, Boulonne, Ratié 2006). Le référencement géographique de la BDD COMIFER et de la BDD RMQS nous donne accès à l'interpolation par le calcul lié à la spatialisation des données.

Notre objectif général pour cette approche est de mettre en œuvre un protocole pour l'interpolation basée sur le géoréférencement de valeurs de Fe (disponibles dans les deux bases de données) en s'appuyant sur la BDD RMQS pour prédire vers la BDD COMIFER.

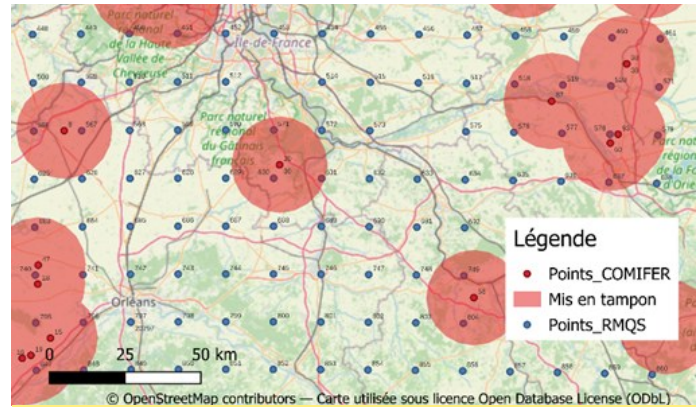


Illustration du protocole d'interpolation du Fe

Enrichir notre BDD de travail suit notre objectif général visant à mieux caractériser les classes de sol.

Notre approche liminaire se base uniquement sur les teneurs en Fe (de la BDD RMQS) les plus proches, pondérées par la distance à nos essais agronomiques (BDD COMIFER). Elle donne des résultats prometteurs avec des corrélations fortes et motivent la poursuite de la démarche. Pour la suite, nous mettons alors en place une démarche similaire, en rajoutant le facteur du référentiel pédologique (ou type de sol) associé à chaque point.

Ainsi pour chaque essai agronomique (Points_COMIFER) on sélectionne sur un rayon de 16km tous les points RMQS ayant le même référentiel pédologique (type de sol) que ce point COMIFER, pour effectuer l'interpolation de la moyenne pondérée par la distance. Si aucun point RMQS du cercle n'a le même référentiel, on sélectionne celui ayant le même référentiel, qu'importe la distance.

Le résultat de cet algorithme est pertinent mais néanmoins très perfectible, surtout comparé à la performance statistique obtenue dans l'étape exploratoire de cette méthode. Deux points d'amélioration sont relevés :

- L'algorithme sélectionne pour le calcul les points ayant le même référentiel pédologique dans son périmètre. A défaut, il prend la valeur du point RMQS le plus proche ayant la même valeur de référentiel pédologique. Cette consigne induit potentiellement une forte variabilité sur la distance entre le point d'extrapolation et la valeur Fe COMIFER extrapolée.

- La précision du référentiel pédologique et de la spatialisation est un facteur limitant dans l'extrapolation des valeurs. L'accès à des bases de données existantes ayant une meilleure précision quant au diagnostic pédologique augmenterait la précision des données d'extrapolation pour l'algorithme imaginé.

CONCLUSION/PERSPECTIVES

Ce travail de stage s’inscrit dans la continuité des études menées sur les seuils de fertilisation P et la caractérisation pédologique des sols à partir des bases de données expérimentales du COMIFER. L’objectif était double : d’une part, comprendre les facteurs de notre BDD COMIFER (notamment pédologiques) déterminant les seuils de fertilisation en P, et d’autre part, explorer des méthodes permettant d’améliorer la qualité, la robustesse et la complétude des données existantes.

Plusieurs méthodes ont été menées mais on retiendra que l’approche géostatistique fondée sur le géoréférencement des bases COMIFER et RMQS qui a été mise en place pour pallier le manque de données en Fe et Al est la plus performante des approches. Les premiers essais d’interpolation spatiale ont montré une cohérence encourageante entre valeurs mesurées et estimées, confirmant la faisabilité du protocole. Cependant, la précision limitée du référentiel pédologique, le manque de robustesse de l’algorithme et la variabilité spatiale des points restreignent encore la robustesse des résultats.

Dans l’ensemble, ces travaux démontrent l’intérêt des approches combinées — statistiques et géostatistiques — pour valoriser des jeux de données hétérogènes et renforcer la compréhension des mécanismes régissant les seuils de fertilisation.

La poursuite du travail passera notamment par l’exploration complète de la méthode dite de géoréférencement qui jusqu’ici a donné des résultats prometteurs mais a manqué d’optimisation dans la méthode.

Bibliographie

JORDAN-MEILLE, L., RUBÆK, G. H., EHLERT, P. A. I., GENOT, V., HOFMAN, G., GOULDING, K., RECKNAGEL, J., PROVOLO, G. et BARRACLOUGH, P., 2012. An overview of fertilizer-P recommendations in Europe: soil testing, calibration and fertilizer recommendations. Soil Use and Management. décembre 2012. Vol. 28, n° 4, pp. 419-435. DOI 10.1111/j.1475-2743.2012.00453.x. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

VAN DOORN, Maarten, van Rotterdam ,Debby, Ros ,Gerard, Koopmans ,Gerwin F., Smolders ,Erik et AND DE VRIES, Wim, 2024. The phosphorus saturation degree as a universal agronomic and environmental soil P test. Critical Reviews in Environmental Science and Technology. 3 mars 2024. Vol. 54, n° 5, pp. 385-404. DOI 10.1080/10643389.2023.2240211.

JOLIVET, Claudy C., BOULONNE, Line et RATIÉ, Céline, 2006. Manuel du Réseau de Mesures de la Qualité des Sols (RMQS) [en ligne]. Disponible à l’adresse : <https://hal.inrae.fr/hal-02818011> [Consulté le 10 octobre 2025]

L’auteur

S.Deroullede est étudiant en 2ème année d’école d’ingénieur agronome a Bordeaux Sciences Agro. Il est spécialisé dans l’informatique pour l’agronomie.

Remerciements

Lionel Jordan-Meille

Soutien financier

GIS Grandes Cultures, COMIFER, INRAE