



HAL
open science

Evaluation d'une nouvelle méthode de mesure de la stabilité structurale par une application mobile SLAKES et comparaison à la méthode de référence en France métropolitaine

Thomas Chalaux

► To cite this version:

Thomas Chalaux. Evaluation d'une nouvelle méthode de mesure de la stabilité structurale par une application mobile SLAKES et comparaison à la méthode de référence en France métropolitaine. Science des sols. 2021. hal-03530262

HAL Id: hal-03530262

<https://hal.inrae.fr/hal-03530262v1>

Submitted on 17 Jan 2022

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.

Mémoire en vue d'obtenir le Master M2
Mention
“ **Agrosciences, Environnement, Territoires, Paysages, Forêts** ”
Parcours
“ **Gestion des Sols et Services Écosystémiques** ”

« Evaluation d'une nouvelle méthode de mesure de la stabilité structurale par une application mobile SLAKES et comparaison à la méthode de référence en France métropolitaine »

« Evaluation of a new measuring method to assess aggregate stability with a mobile phone application SLAKES and comparison with the standard method in metropolitan France »

Par : **Thomas CHALAUX**

Stage effectué à INRAE Centre Val de Loire, UR SOLS et US Infosol
2163 Avenue de la pomme de pin, 45075 Ardon
Sous la direction de **LACOSTE Marine et SABY Nicolas**



Soutenu le 02 / 09 / 2021, à Paris, devant le jury composé de :

Rapporteur : **GILLIOT Jean-Marc** (AgroParisTech)

Examineurs : **VAUDOUR Emmanuelle** (AgroParisTech)
COQUET Yves (AgroParisTech)
MICHELIN Joël (AgroParisTech)
Cécile Quentin (Université Paris-Saclay)

Remerciements

Je tiens à remercier vivement mes deux maîtres de stage, Marine Lacoste et Nicolas Saby, pour la qualité de leur encadrement et leur accueil au sein de l'UR SOLS et Infosol à l'INRAe d'Orléans. Je vous remercie pour la pédagogie dont vous avez fait preuve, pour m'avoir offert l'opportunité de présenter mon travail à différentes occasions et de m'avoir permis de rencontrer d'autres chercheurs avec qui j'ai eu un véritable plaisir à échanger.

Je remercie tous les partenaires du projet ECLAT et en particulier Blandine Lermancier, Frédéric Darboux et Didier Michot pour leur expertise en pédologie et l'aide qu'ils m'ont apportée tout au long de mon travail.

Enfin, je remercie l'ensemble des membres de l'INRAe d'Orléans pour leur gentillesse et leur accueil. Je garderai de bons souvenirs de nos discussions et promenades après les repas.

Table des matières

I. Introduction	1
II. Matériel et méthodes	2
A. Méthodes de mesure de la stabilité structurale	3
a. Préparation des agrégats (adapté de la norme AFNOR NF X31-515 juin 2005)	3
b. Méthode Le Bissonnais (ISO 10930-2012)	3
c. Méthode SLAKES (Fajardo et al., 2016).....	4
B. Jeux de données pédologiques	5
a. Réseau de Mesure de la Qualité des sols.....	6
c. Sites OS2 et Villamblain.....	11
d. Bassin versant de Naizin (ORE AgrHyS)	12
e. Sol de Bretagne	13
f. Essai Travail du Sol et Matière Organique (TS-MO) de la plateforme Effluents d'Elevage et Environnement (EFELE) (adapté de Peluchon, 2021)	13
C. Sélection de l'indicateur SLAKES	16
a. Comparaison des méthodes de regroupement au site.....	16
b. Evaluation de la répétabilité de la mesure SLAKES	16
c. Evaluation de la capacité de l'indicateur SLAKES à différencier des pratiques agricoles..	17
d. Comparaison avec le MWD rapide de la méthode Le Bissonnais.....	17
e. Proposition et évaluation du référentiel d'interprétation SLAKES en France métropolitaine	17
E. Etude du lien avec les propriétés pédologiques et prédiction par la modélisation	18
a. Jeux de variables	18
b. Données de textures	18
c. Ensembles de données.....	19
d. Etude des corrélations avec les propriétés physico-chimique et descriptive du sol	19
e. Modèle par apprentissage automatique	19
f. Sélection des meilleurs modèles par validation croisée (10 blocs)	20
F. Prédiction sur le RMQS.....	21
III. Résultats	21
A. Sélection de l'indicateur	21
a. Comparaison des méthodes de regroupement au site.....	21
b. Evaluation de la répétabilité de la mesure SLAKES	22
c. Evaluation de la capacité de l'indicateur SLAKES à différencier des pratiques agricoles	22
d. Comparaison avec la méthode Le Bissonnais	24
e. Proposition d'un référentiel d'interprétation.....	26
B. Modélisation	28
a. Corrélation avec les propriétés physico-chimiques et descriptives du sol.....	28
b. Sélection des modèles descriptifs.....	29
c. Règles de pédotransfert.....	30
g. Modèles prédictifs	32
C. Prédications sur le RMQS	33
IV. Discussion	35

Evaluation de la méthode SLAKES en France métropolitaine

A. Jeux de données.....	35
B. Répétabilité de la mesure et sélection de l'indicateur SLAKES	36
C. Référentiel d'interprétation.....	37
D. Formules de pédotransfert	38
E. Cartes de stabilité structurale	38
V. Conclusion	40
VI. Bibliographie	42
VII. Annexes	1

Liste des figures

- Figure 1* : Poste de travail organisé pour la mesure SLAKES (A) plusieurs postes de travail et (B) poste de travail (Peluchon, 2021).
- Figure 2* : Localisation des sites possédant une mesure de la stabilité structurale par (a) la méthode SLAKES, (b) la méthode Le Bissonnais selon la campagne du RMQS.
- Figure 3* : Localisation des sites du RMQS possédant une mesure de la stabilité structurale uniquement SLAKES, uniquement Le Bissonnais ou les deux.
- Figure 4* : Localisation des sites sélectionnés pour la mesure de la stabilité structurale par la méthode SLAKES dans le bassin versant de Naizin (Tahar, 2021).
- Figure 5* : Localisation des sites étudiés provenant du programme Sol de Bretagne et de l'essai TS-MO (Peluchon, 2021).
- Figure 6* : Organisation des parcelles expérimentales de l'essai TS-MO de la plateforme EFELE réparties aléatoirement en 3 blocs (Peluchon, 2021).
- Figure 7* : Localisation des sites de l'ensemble des jeux de données étudiées dans le cadre du projet ECLAT (349 sites).
- Figure 8* : Schéma de la construction des jeux de données à partir des propriétés pédologiques et des textures.
- Figure 9* : Distribution des différences calculées par site entre les méthodes de regroupement au site pour chaque indicateur *SLAKES* (a) Asym et (b) SI600 (n = 146).
- Figure 10* : Distribution des indicateurs *SLAKES*, Asym et SI600, en fonction de l'effet croisé du travail du sol (L = labour ; TS = travail simplifié) et de la fertilisation (FB = Fumier bovin, MIN = minérale).
- Figure 11* : Distribution des indicateurs *SLAKES*, Asym et SI600 en fonction du (a) travail du sol ou de (b) la fertilisation. p-value du test de Wilcoxon, n = 96.
- Figure 12* : Indicateur *SLAKES* Asym médian en fonction du SI600 médian (n = 349). noir : x = y ; rouge = régression linéaire entre x et y.
- Figure 13* : MWD rapide de Le Bissonnais en fonction des indicateurs *SLAKES* : a. Asym moyen, b. Asym médian, c. Asym *model based*, d. SI600 moyen, e. SI600 médian et f. SI600 *model based*, et représentation des référentiels d'interprétations de Le Bissonnais (1996) et Fajardo et al. (2016).
- Figure 14* : MWD_{rapide} de Le Bissonnais en fonction du SI600 médian *SLAKES* et représentation du référentiel d'interprétation de Le Bissonnais (1996) et du référentiel *SLAKES* proposé (n = 349).
- Figure 15* : Coefficients de corrélation de Pearson (r) entre indicateurs de stabilité structurale *SLAKES* (Asym, SI600) (n = 117) et Le Bissonnais (MWD_{rapide}) (n = 151) et les propriétés de sols.
- Figure 16* : Représentation des individus en analyse en composante principale (ACP) selon l'occupation du sol pour le (a) SI600 médian *SLAKES* et le (b) MWD_{rapide} de Le Bissonnais.
- Figure 17* : Règle de pédotransfert du modèle *Cubist* de l'indicateur de d'instabilité structurale *SLAKES* SI600.
- Figure 18* : Règles de pédotransfert du modèle *Cubist* de l'indicateur de stabilité structurale MWD rapide de Le Bissonnais.

Evaluation de la méthode SLAKES en France métropolitaine

Figure 19 : Distribution du SI600 SLAKES en fonction (1) de la CEC (cmol+/kg), (2) des limons (2-50 μm) (g/ kg), (3) des sables (50-2000 μm) (g/kg), (4) du Ca échangeable (cmol+/kg) et (5) du carbone organique (g/kg) (n = 121).

Figure 20 : Distribution du MWD rapide de Le Bissonnais en fonction (1) du carbone organique (g/kg), (2) de l'Al échangeable (cmol+/kg) pour des teneurs en carbone organique < 31,80 g/kg, (3) de l'Al échangeable (cmol/kg) pour les sites où l'occupation du sol est "Cultures".

Figure 21 : Distribution des valeurs de stabilité structurale (a) SI600 SLAKES et (b) $\text{MWD}_{\text{rapide}}$ Le Bissonnais observées (resp. 132 ; 183) et prédites (resp. 1874 ; 1875) du le RMQS.

Figure 22 : Distribution des sites par classe de stabilité structurale (a) SI600 SLAKES et (b) $\text{MWD}_{\text{rapide}}$ Le Bissonnais, sites observées 132 et 183 respectivement, et sites prédits 1742 et 1692 respectivement, du RMQS selon le référentiel adapté.

Figure 23 : Carte de France de la stabilité structurale des agrégats du sol selon (a) l'indicateur SI600 SLAKES et (b) l'indicateur MWD rapide Le Bissonnais, nombre de sites observés 132 et 182 respectivement et sites prédits 1742 et 1692 respectivement.

Liste des tableaux

Tableau I : Nombre de sites possédant une mesure de la stabilité structurale par la méthode SLAKES ou Le Bissonnais selon la campagne du RMQS (1 ou 2).

Tableau II.a : Résumé des propriétés physico-chimiques du sol des sites RMQS possédant une mesure de la stabilité structurale par la méthode Slakes.

Tableau II.b : Résumé des propriétés physico-chimiques du sol des sites RMQS possédant une mesure de la stabilité structurale par la méthode Le Bissonnais.

Tableau III.a : Répartition des sites selon la classification des sols du référentiel pédologique (Baize et al., 1995 ; Baize & Girard, 2008) des sites du RMQS possédant une mesure SLAKES ou Le Bissonnais.

Tableau III.b : Répartition des sites selon l'occupation du sol des sites du RMQS possédant une mesure SLAKES ou Le Bissonnais.

Tableau IV : Résultats des tests du χ^2 de conformité (p-value) des jeux de données SLAKES et Le Bissonnais du RMQS.

Tableau V : Résumé des propriétés du sol des sites OS2 (Saint-Denis-Des-Puits, Cernay) et Villamblain.

Tableau VI : Résumé des propriétés du sol des sites sélectionnés pour la mesure de la stabilité structurale par la méthode SLAKES dans le bassin versant de Naizin (198 sites).

Tableau VII : Résumé des propriétés du sol des sites du programme Sols de Bretagne (25 sites).

Tableau VIII : Résumés des propriétés du sol pour les sites de l'essai TS-MO de la plateforme EFELE.

Tableau IX : Description des jeux de données de textures.

Tableau X : Médianes et écarts interquartiles des différences calculées entre les méthodes de regroupement au site pour chaque indicateur SLAKES (Asym et SI600) (n = 146).

Tableau XI : Nombre de répétitions de la mesure de 3 agrégats nécessaires pour différencier les sites Saint-Denis-Des-Prés, Cernay et Villamblain selon chaque indicateur SLAKES : Asym et SI600.

Tableau XII : Nombre de répétitions de la mesure de 3 agrégats nécessaires pour différencier les modes de travail du sol : labour (L) et travail simplifié (TS) selon chaque indicateur SLAKES : Asym et SI600.

Tableau XIII : Matrice de corrélation de Spearman entre le MWD rapide de Le Bissonnais et les indicateurs SLAKES (Asym et SI600) (n = 349).

Tableau XIV : Corrélogramme des coefficients de corrélation de Pearson entre les indicateurs SLAKES (Asym et SI600) (n = 349).

Tableau XV : Bonnes attributions des classes d'interprétation du MWD rapide de Le Bissonnais par le référentiel d'interprétation proposé du SI600 médian de SLAKES (en %).

Tableau XVI : Référentiel d'interprétation de la stabilité des agrégats selon la valeur du SI600 médian SLAKES proposé.

Tableau XVII : Statistiques de la validation croisée (10 blocs) des modèles *cupist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

Evaluation de la méthode SLAKES en France métropolitaine

Tableau XVIII : Utilisation des variables présentant des données manquantes par les modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais

Tableau XIX : Statistiques de la validation croisée (10 blocs) des modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

Tableau XX : Statistiques de la validation croisée (10 blocs) des modèles *cubist* prédictif SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

Tableau XXI : Statistiques de la validation croisée (10 blocs) des modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

Evaluation de la méthode SLAKES en France métropolitaine

Résumé

La battance et l'érosion résultent de la désagrégation des agrégats de sol et du détachement de particules sous l'action des pluies. Ils dépendent de la stabilité structurale du sol : un sol stable sera moins sensible à la battance et l'érosion. La mesure de référence de la stabilité structurale en France est la méthode normalisée Le Bissonnais. Cependant, cette mesure est complexe à mettre en œuvre, longue et onéreuse, ce qui représente un frein dans l'acquisition de données. Une nouvelle méthode, développée par Fajardo et al. (2016), accompagnée d'une application téléphone (Fajardo & McBratney, 2019), nommée SLAKES, est *a priori* beaucoup plus simple à mettre en œuvre. L'objectif de cette étude est d'évaluer cette nouvelle méthode dans le contexte de la France métropolitaine. Nous avons procédé en plusieurs temps. Tout d'abord, la robustesse et la précision des indicateurs SLAKES (appelés « Asym » et « SI600 ») sont comparés sur un grand nombre de sols. Leur sensibilité aux pratiques agricoles est comparée, par rapport à la méthode de référence. Dans un deuxième temps, ce travail aborde le lien avec les propriétés pédologiques influençant sa distribution afin de caractériser *in fine* sa distribution spatiale à l'échelle de la France métropolitaine par l'utilisation des sites du Réseau de Mesure de la Qualité des Sols (RMQS). L'évaluation de la répétabilité des indicateurs démontre que l'indicateur proposé dans ce travail, le SI600, est plus précis et robuste que celui proposé par la méthode SLAKES (Asym). Le nombre de 30 à 70 répétitions de la mesure de 3 agrégats permet d'estimer la stabilité avec un niveau de confiance de 10 à 1%. Le SI600 est très corrélés au MWD_{rapide} de Le Bissonnais ($r = 0.78$) et de cette corrélation, un référentiel d'interprétation de SLAKES est proposé. L'étude des corrélations avec les propriétés pédologiques et la modélisation par modèle *cubist* montre des liens attendus entre les indicateurs et les propriétés pédologiques. La distribution des deux indicateurs à l'échelle du territoire est semblable et dessine les grands ensembles géologiques français.

Mots clés : sol ; stabilité structurale ; érosion ; méthode ; SLAKES

Summary

Crusting and erosion result from the disintegration of soil aggregates and the detachment of soil particles by rain. They depend on aggregate stability: a stable soil will be less sensitive to crusting and erosion. In France, the reference measure of aggregate stability is the standardized method Le Bissonnais (1996). However, this method is complex to implement, long and expensive. This is a curb on data acquisition. A new method developed by Fajardo and al. (2016) and followed by a mobile phone application named SLAKES (Fajardo & McBratney, 2019) is *a priori* much easier to implement. The objective of this study is to assess the SLAKES method in the context of the metropolitan France. We proceeded in several steps. First, the robustness and accuracy of the SLAKES indicators (called « Asym » and « SI600 ») are compared on a large set of soils. Their sensitivity to management practices is compared to the reference method. In a second step, this work addresses to determine the influence of soil properties on their values in order to characterize their spatial distributions at metropolitan France scale, using the *Réseau de Mesure de la Qualité des Sols (RMQS)* sites. The evaluation of indicators repeatability shows that the indicator proposed in this work, SI600, is more accurate and robust than the one proposed by the SLAKES method (Asym). The number of 30 to 70 repetitions of the measure of 3 aggregates allows to estimate the stability

Evaluation de la méthode SLAKES en France métropolitaine

with a confidence level of 10 to 1%. The SI600 is highly correlated with Le Bissonnais's MWD_{rapide} ($r = 0.78$) and from this correlation we propose an interpretation scale. Expected links between indicators and soil properties were highlighted by the study of correlations with soil properties and cubist modelling. The distribution of the indicators (SI600 and MWD_{rapide}) at France scale is similar and draws French geological settings.

Key words : soil ; aggregate stability ; erosion ; method ; SLAKES

I. Introduction

L'érosion est une des menaces qui pèse sur les sols (FAO, 2019), pouvant à terme compromettre la fourniture des services écosystémiques attendus. L'érosion correspond à l'ensemble des processus de détachement, de transport et de dépôts de matériaux solides à la surface de la terre. En France métropolitaine, l'érosion est majoritairement hydrique et près de 18% des sols présentent un aléa d'érosion moyen à très fort (Gis Sol : Erosion des sols). L'érosion hydrique a des conséquences à différentes échelles spatiales. A long terme, l'érosion peut localement conduire à une dégradation irréversible des sols par la perte de matériaux, des éléments chimiques et de la couche organique. Cette dégradation peut altérer les services écosystémiques des sols. De plus, l'érosion a des effets à des échelles spatiales plus larges en formant des coulées boueuses et des transferts de pollutions (eutrophisation par transfert d'éléments fertilisants ou de produits phytopharmaceutiques) (Gis Sol : Erosion des sols).

L'érosion hydrique a lieu lorsque l'eau ne peut plus s'infiltrer dans le sol et qu'un ruissellement se forme. Elle découle de deux processus, la battance et l'érodibilité. La battance se définit par la destruction des agrégats à la surface de certains sols sous l'impact des gouttes de pluie (ou de certains dispositifs d'irrigation), conduisant à la formation d'une mince croûte (0,5 à 5 mm) à structure liée, plus ou moins continue, appelée croûte de battance (Baize, 2016). Cette croûte de battance réduit les transferts d'eau et de gaz à la surface du sol suite à l'action mécanique de la pluie. L'érodibilité traduit la vulnérabilité d'un sol à l'érosion, dans des circonstances précises (état de surface, humidité, type de précipitation) (Baize, 2016). Ainsi, une forte sensibilité à la battance et une forte érodibilité sont favorables à l'érosion. Ces deux processus dépendent de la stabilité structurale du sol, c'est-à-dire, la capacité du sol à résister à la dégradation de sa structure sous l'action de contraintes mécaniques dues à l'eau. Donc, un sol présentant une stabilité structurale élevée possèdera un risque de battance et une érodibilité plus faibles.

Un ensemble de méthodes existent pour évaluer la stabilité structurale du sol (Amezketta, 1999 ; Almajmaie et al. 2017) En France, seule la méthode développée par Yves Le Bissonnais en 1996 fait l'objet d'une normalisation ISO (ISO 10930:2012) (AFNOR, 1992). La méthode est dite par tamisage humide, quand d'autres méthodes reposent sur la simulation de pluie (Young, 1984), la vibration ultrasonique (Edwards & Bremner, 1967), l'immersion (Levoday & Pyle, 1973 ; Field et al., 1997), le tamisage sec, la spectroscopie ou la granulométrie laser. La méthode de Le Bissonnais s'inspire de Yoder (1936), Hénin et al. (1958) et Tisdall & Oades (1982). Cette méthode est reconnue comme reliée au comportement réel du sol (Le Bissonnais & Le Souder, 1995 ; Le Bissonnais et al., 1996 ; Amezketa et al., 1996) et a été utilisée dans différents contextes (Saygin et al., 2015 ; Le Bissonnais et al., 2007 ; Annabi et al, 2017 ; Shi et al., 2020 ; Liu et al., 2021). Ayant été cité 812 fois (CrossRef sur Le Bissonnais, 1996), elle est fréquemment utilisée comme point de comparaison (Amézketa, 1999 ; Gomez et al., 2013 ; Erktan et al., 2016 ; Almajmaie et al., 2017 ; Liu et al., 2021), ce qui a valu la publication d'une réflexion par Yves Le Bissonnais sur l'impact de sa méthode sur le sujet de l'évaluation de l'érosion hydrique (Le Bissonnais, 2016). Ses principales limites sont cependant son coût élevé, le temps nécessaire pour effectuer les mesures et la maîtrise de son protocole. Ces limites représentent un frein dans l'acquisition de données sur la stabilité structurale, malgré l'importance de cette propriété au vu des enjeux d'érosion et plus largement de qualité des sols.

Récemment, une nouvelle méthode a été développée par Fajado et al. (2016), puis implémentée au sein d'une application pour téléphone afin de faciliter sa diffusion (Fajardo & McBratney, 2019). Cette approche, appelée « méthode SLAKES », s'inspire de la méthode d'évaluation de la stabilité d'agrégats dans l'eau (Aggregate Stability in Water : ASWAT) de Field et al. (1997). La méthode SLAKES implémentée dans l'application évalue l'étalement d'agrégats de 3 à 5 mm environ préalablement séchés immergés dans de l'eau pendant 10 minutes par analyse d'images acquises automatiquement. La méthode présente l'avantage d'être disponible librement, peu onéreuse et simple et rapide à mettre en œuvre.

L'antériorité de la méthode Le Bissonnais, la normalisation de la mesure, la disponibilité d'un référentiel d'interprétation pour les sols en France (Le Bissonnais, 1996 ; Le Bissonnais & Arrouays, 1997) ainsi que son lien avec le comportement réel du sol face à l'érosion et la battance (Le Bissonnais & Le Souder, 1995 ; Le Bissonnais et al., 1995 ; Amezketa et al., 1996) en font un indicateur robuste. Au contraire, la méthode SLAKES, bien que prometteuse, reste une méthode encore en cours de développement et ne présente que très peu de retours d'expériences (Flynn et al., 2020 ; Jones et al., 2021). Le projet ECLAT, porté par Nicolas Saby (INRAE US Infosol), a pour but de contribuer à son développement et d'évaluer la méthode SLAKES dans le contexte des sols de France métropolitaine en s'appuyant notamment sur une comparaison avec la méthode Le Bissonnais et sur un corpus de données et d'échantillons disponibles au sein du partenariat. Ainsi, les résultats de ce projet sont les premiers retours sur la méthode SLAKES dans le contexte français métropolitain et européen.

S'inscrivant dans le projet ECLAT, les objectifs de ce stage sont les suivants :

- 1) Affiner le calcul d'un indicateur de la stabilité structurale du sol et évaluer sa robustesse et précision sur un grand nombre d'échantillon de sols dans le contexte de la France métropolitaine (section C) ;
- 2) Identifier les déterminants de la distribution des indicateurs SLAKES et Le Bissonnais à l'aide de variables explicatives décrivant les propriétés physico-chimiques et descriptives du sol (section D) ;
- 3) Caractériser la distribution spatiale à l'échelle du territoire de ces mesures à l'aide d'une prédiction sur un ensemble de sites, et définir les zones à risque de stabilité structurale faible sur le territoire de la France métropolitaine à l'aide du référentiel précédemment défini (section E).

II. Matériel et méthodes

La méthode SLAKES et le test rapide de Le Bissonnais évaluent la stabilité structurale face au phénomène d'éclatement (Le Bissonnais, 1996). En effet, lors de la ré-humectation d'un agrégat de sol sec, l'air contenu dans l'agrégat est comprimé par l'eau et cette augmentation de pression cause l'éclatement de l'agrégat (Hénin, 1938 ; Kemper, 1965 ; Emerson, 1967 ; Truman et al., 1990 ; Le Bissonnais, 1996). Ce phénomène est fonction de l'air contenu à l'intérieur de l'agrégat, mais aussi de sa teneur en argile (Le Bissonnais, 1996). Il se rapproche de ce que pourrait subir un agrégat sec lors d'une pluie d'orage.

A. Méthodes de mesure de la stabilité structurale

a. Préparation des agrégats (adapté de la norme NF X31-515 juin 2005)

Le prélèvement de l'échantillon est préconisé en conditions d'humidité modérée (inférieure à la capacité au champ) et réalisé à l'aide d'un transplantoir ou une bêche sur une épaisseur de 10 cm, sans prendre les parties lissées par l'outil. On prélève en différents endroits, minimum 5, sur une surface homogène de l'ordre de 10 m². L'échantillon composite final est d'environ 1 kg de terre. On place l'échantillon dans une boîte plastique rigide hermétique (1-2 L). Le transport est fait dans une glacière à 4°C et les échantillons sont stockés en chambre froide.

Après le prélèvement, les échantillons sont mis à sécher à l'air, dans une atmosphère tempérée (20°C) et ventilée, et étalés délicatement dans des récipients plats posés sur des rayonnages ou d'autres supports. Durant la période du séchage, de un à quelques jours, les plus grosses mottes peuvent être émietées manuellement quotidiennement pour produire un maximum d'agrégats de taille pluri-millimétrique. L'émottage ne doit pas modifier la porosité. La porosité correspond à l'ensemble des volumes d'un échantillon de sol non occupé par la phase solide.

Dans un deuxième temps, les échantillons sont tamisés en appliquant une énergie modérée : seule la fraction 3-5 mm est récupérée. L'agitation ne doit pas être trop intense pour ne pas émousser les agrégats. On procède à un séchage à 40°C pendant 48 h. Les agrégats sont conditionnés dans des boîtes de stockage rigides en chambre froide.

La norme de la méthode Le Bissonnais précise des limites dans son applications (Afnor, 2012). Si le pourcentage de gravier est compris entre 10 % et 40 %, la mesure peut être faussée, il faut alors rajouter une étape pour discerner les graviers dans les agrégats restants supérieurs à 2 mm. Si le pourcentage de gravier est supérieur à 40%, la mesure n'est pas significative. De plus, la méthode ne s'applique pas aux sols non structurés ou dont la cohésion n'est pas suffisante pour former des agrégats (Norme iso). Même si ce n'est pas précisé dans la méthode SLAKES, ces limites sont aussi valables.

b. Méthode Le Bissonnais (ISO 10930-2012)

La méthode Le Bissonnais comporte trois tests différents : humectation rapide, lente et désagrégation mécanique. Chacun de ces tests permet d'évaluer plus précisément le comportement des agrégats face aux phénomènes de désagrégation. Le (A) test d'humectation rapide évalue le phénomène d'éclatement, le (B) test d'humectation lente différencie les sols instables et le (C) test de désagrégation mécanique la stabilité structurale sans le phénomène d'éclatement. Les résultats sont classés selon un référentiel d'interprétation composé de cinq classes : très instable (<0,3), instable (0,3-0,8), moyennement stable (0,8-1,2), stable (1,2-2) et très stable (2-5).

Chaque test se compose de trois étapes. Trois tests sont réalisés à partir des agrégats précédemment préparés et utilise 5 à 10 g d'agrégats (~ 200 à 250 agrégats). Les agrégats secs sont soit (A) rapidement immergés dans de l'eau, soit (B) réhumectés à l'aide d'une table de tension ou (C) saturés par de l'éthanol puis plongés et agités dans de l'eau déionisée pour ensuite être placés sur un tamis de 50 µm immergé dans l'éthanol afin de mesurer leurs dimensions et la masse résiduelle. Pour chaque test, un diamètre moyen pondéré (*Mean Weighted Diameter* = MWD) est calculé. Enfin, la

moyenne pondérée des trois MWD est calculée. Plus le MWD sera élevé, plus le sol sera stable.

$$MWD = \sum_{i=1}^{i=n} x_i * w_i \quad (1)$$

avec i l'indice du tamis, x_i la taille moyenne des agrégats au tamis i ,
 w_i la proportion de masse totale d'agrégats résiduels entre les tamis $i-1$ et i .

Des classes de stabilité et de sensibilité à la battance permettent d'analyser le MWD final. Selon la norme ISO, il est possible de ne réaliser que le test (A) d'humectation rapide (MWD_{rapide}) qui est un bon indicateur du comportement réel du sol (Norme ISO). Il faudra alors réaliser 3 répétitions de la mesure.

c. Méthode SLAKES (Fajardo et al., 2016)

• Principe

La méthode SLAKES utilise également une méthode de désagrégation par humectation rapide, mais elle s'intéresse également à la dynamique temporelle du processus en suivant l'étalement des particules de terre. La mesure proposée jusqu'ici consiste à acquérir des photographies de la surface des agrégats secs plongés dans l'eau à pas de temps progressif, de l'immersion jusqu'à 10 minutes (Fajardo et al., 2016) et de fournir un indicateur à partir des données. Un référentiel d'interprétation est proposé par Fajardo et al. (2016). Il est composé de trois classes : haute stabilité (< 3), moyenne stabilité (3-7) et faible stabilité (> 7). Selon la méthode présentée par Fajardo et al. (2016), la mesure SLAKES présente l'intérêt d'une acquisition simple et rapide des résultats.

Pour améliorer la détection numérique des agrégats par l'algorithme de traitement d'images, les mesures sont réalisées dans une salle privée de lumière naturelle et éclairée à l'aide de lampes pour obtenir une lumière diffuse et constante. Le téléphone est placé à une hauteur de 10 à 13 cm au-dessus de l'échantillon à analyser, sur une potence (fig. 1). Les agrégats sont positionnés dans une boîte de Petri, sous le téléphone. On place une feuille blanche plastifiée (fond blanc) sous la boîte de Petri. Pour chaque mesure, il est recommandé de placer 3 agrégats. Une première photo des agrégats secs permet de mesurer la surface initiale de l'agrégat (A_{t_0}). On retire les agrégats afin de remplir la boîte de Pétri avec 30 ml d'eau. La boîte de Pétri remplie est replacée sous l'objectif du téléphone. Les agrégats sont lâchés dans l'eau à proximité de leur position initiale et simultanément, l'acquisition d'image par l'application est mise en route pour une durée de 10 minutes. Il est recommandé de réaliser 3 répétitions de la mesure par site (Fajardo et al., 2016).

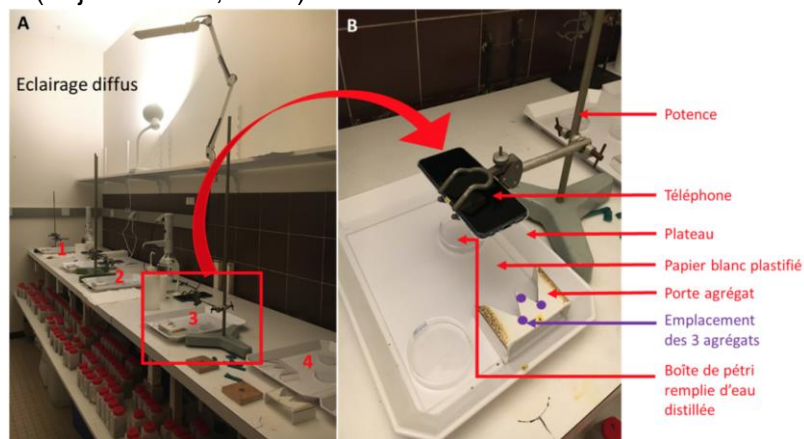


Figure 1. Poste de travail organisé pour la mesure SLAKES (A) plusieurs postes de travail et (B) poste de travail (Peluchon, 2021)

Plusieurs postes de mesure peuvent être mis en œuvre par un opérateur seul, permettant de réaliser des mesures en parallèle. Nous avons testé 8 postes. Cependant, il peut exister un effet du site de poste lié à l'éclairage (Peluchon, 2021) et aussi un effet du téléphone (Jabely, 2021). L'effet téléphone est lié aux différences entre les propriétés du matériel optique (résolution, sensibilité, ouverture) et aux traitements des images par les systèmes d'exploitation.

• Définition des indicateurs

Selon la méthode originale proposée par Fajardo et al. (2016), les mesures brutes du taux d'accroissement de la surface SI au cours du temps (A_t) est calculé : $SI_t = \frac{A_t - A_{t0}}{A_{t0}}$

(2). Une équation de Gompertz est ensuite ajustée sur les données expérimentales :

$SI_t = Asym * e^{-b * e^{-c * \log(t)}}$ (3) par une approche des moindres carrés non linéaires. La courbe de Gompertz permet ainsi de capter la dynamique temporelle. Dans la publication initiale (Fajardo et al., 2016), l'indicateur de la stabilité structurale retenu est le paramètre $Asym$ de l'équation de Gompertz. $Asym$ est un équivalent théorique à l'étalement maximal de l'agrégat au bout d'un temps infini. Cependant, l'ajustement sur l'équation de Gompertz peut ne pas se faire correctement lorsque les données brutes sont biaisées ou bruitées (Sakande, 2020 ; Peluchon, 2021 ; Mousset, 2021) et les valeurs estimées $Asym$ sont alors inutilisables. Ensuite, Flynn et al. (2020) ont proposé de retenir l'indicateur SI à 600 secondes ($SI600$) qui est décrit comme plus sensible aux pratiques agricoles. De plus, il présente l'avantage de ne pas nécessiter d'ajuster une équation de Gompertz. Les deux indicateurs sont calculés pour cette étude. Les indicateurs sont ensuite regroupés par site à l'aide soit d'une moyenne, d'une médiane ou d'un modèle linéaire mixte (mb) prenant en compte le téléphone ayant réalisé la mesure. A la différence du MWD, plus le SI et le $Asym$ seront élevés, plus la stabilité structurale sera faible.

Fajardo et al. (2016) met en place un maximum théorique pour le $Asym$ à 7,8 à partir de la corrélation entre le $Asym$ et la teneur en carbone organique du sol. Nous avons décidé d'augmenter le seuil à 20, de manière à étudier une plus large gamme de résultats. Dans la publication initiale (Fajardo et al., 2016), l'équation de Gompertz ajustée d'un agrégat est conservée si le carré du coefficient de détermination (R^2) est satisfaisant, soit proche ou supérieur à 0,80. Il est aussi précisé que pour les agrégats stables le coefficient de corrélation est plus faible que pour des sols instables. De manière à éviter de rejeter les agrégats stables tout en minimisant les erreurs d'ajustement sur la Gompertz, on fixe $R^2 > 0,60$.

B. Jeux de données pédologiques

Dans le cadre de ce travail, des échantillons provenant de plusieurs ensembles de sites ont été utilisés afin de maximiser la disponibilité des échantillons à budget réduit. Cela a donc consisté à mobiliser des échantillons disponibles chez les partenaires du projet ECLAT, des phases de terrain n'étant pas envisageable pendant la durée du stage.

a. Réseau de Mesure de la Qualité des sols

Le principal jeu de données de ce stage provient du programme Réseau de Mesures de la Qualité des Sols (RMQS) piloté par l'unité Infosol (INRAe Orléans). Les sites du RMQS sont répartis uniformément selon une maille carrée de 16 km de côté (Arrouays et al., 2002). La taille de la grille fait du RMQS un réseau statistiquement représentatif du territoire français (Arrouays et al., 2003). Deux campagnes de collecte d'échantillons sont disponibles. La première campagne réalisée de 2000 à 2009 (RMQS-1) correspondant à des prélèvements effectués dans le cadre de travaux de thèse (Algayer, 2012) et la 2nd de 2016 à 2021 (RMQS-2) correspondant aux échantillons prélevés pour les mesures physiques de pF (Swiderski et al., 2017).

D'une part, des mesures SLAKES sont réalisées sur une sélection de 132 sites issus du Réseau de Mesure de la Qualité des Sols (RMQS) de France métropolitaine (2145 sites) (Fig. 2.a). Pour des raisons opérationnelles, les sols analysés se répartissent entre les deux campagnes de suivi des sols du RMQS. Les 132 mesures acquises pour le projet ont donc été réalisées selon la disponibilité des échantillons provenant des deux campagnes. D'autre part, 200 mesures de la stabilité structurale par la méthode Le Bissonnais étaient présentes dans la base de données du RMQS. Ces mesures sont en réalité réalisées sur 183 sites du RMQS (Fig 2.b). Certains sites possèdent une mesure de stabilité sur les deux campagnes et sont notés "doublon". Le nombre de sites pour chaque mesure est précisé dans le Tableau I. La figure 3 positionne les sites possédant une mesure SLAKES et/ou une mesure Le Bissonnais en France métropolitaine. La nuance entre les deux campagnes du RMQS (1 et 2) n'a pas été prise en compte dans le cadre de ce travail.

Tableau I. Nombre de sites possédant une mesure de la stabilité structurale par la méthode SLAKES ou Le Bissonnais selon la campagne du RMQS (1 ou 2).

	Slakes	Le Bissonnais
Sites par campagne du RMQS		
RMQS-1	91	134
RMQS-2	55	66
Total	146	200
Sites uniques	132	183
Sites en doublon	14	17

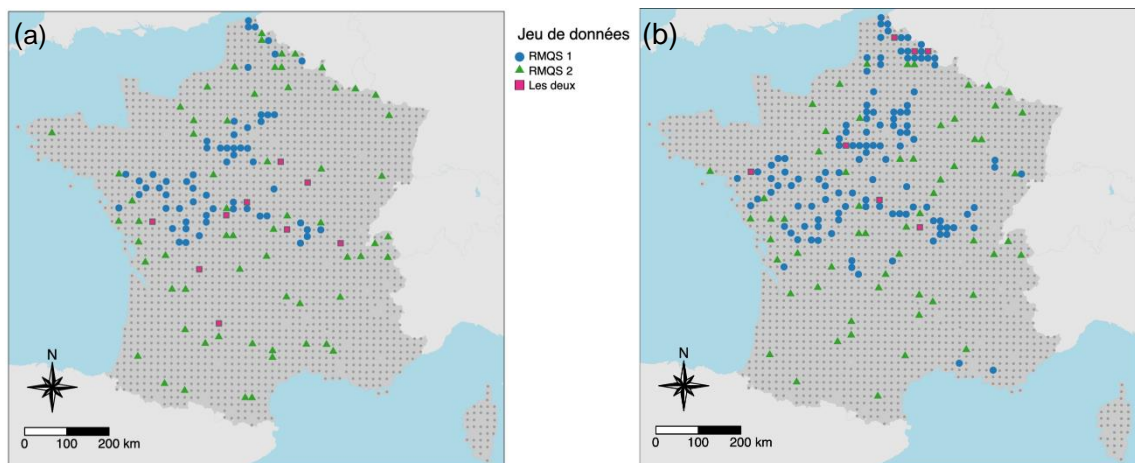


Figure 2. Localisation des sites possédant une mesure de la stabilité structurale par (a) la méthode SLAKES, (b) la méthode Le Bissonnais selon la campagne du RMQS. rond bleu : RMQS-1, triangle vert : RMQS-2, carré rose: les deux.

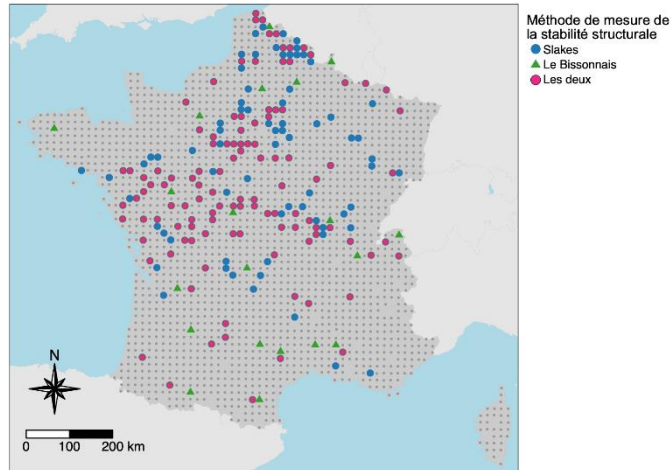


Figure 3. Localisation des sites du RMQS possédant une mesure de la stabilité structurale uniquement SLAKES, uniquement Le Bissonnais ou les deux. rond bleu : SLAKES, triangle vert : Le Bissonnais, carré rose: les deux méthodes.

Les propriétés physico-chimiques et descriptives des sols pour les sites RMQS mobilisées dans le cadre du projet ECLAT sont résumés dans les Tableaux II (a, b) et III (a, b, c) sans distinction de la campagne du RMQS. La représentativité des propriétés des sites sélectionnés par rapport au reste du RMQS a été évaluée l'aide de test de Kolmogorov-Smirnoff pour les variables quantitatives (Tableau II.a et II.b) et un test du Chi 2 de conformité pour les variables qualitatives (Tableau IV).

Tableau II.a Résumé des propriétés physico-chimiques du sol des sites RMQS possédant une mesure de la stabilité structurale par la méthode SLAKES.

d : descriptif ; p : prédictif.

	Unité	Effectifs	Minimum	Médiane	Moyenne	Maximum	Ecart-type	Test de Kolmogorov - Smirnov	Utilisation par les modèles
Carbone organique (CO)	g/kg	141	5.37	16.15	19.54	69.90	12.07	0	d p
N total	g/kg	141	0.54	1.56	1.78	5.44	0.90	0	d p
C/N		141	7.87	10.15	10.84	27.24	2.58	0	d p
Carbonate	g/kg	141	0.50	1.00	35.62	646.00	91.84	0	d p
pH eau		141	3.9	6.8	6.7	8.3	1.2	0.01	d p
Densité apparente	g/cm3	141	0.62	1.38	1.36	1.84	0.22	0.02	d p
CO/argile		141	0.03	0.07	0.08	0.34	0.05	0	d p
CEC/argile	cmol+/g	141	0.02	0.06	0.06	0.15	0.02	0	d p
Texture									
Argile (< 2 µm)	g/kg	141	84	232	261	648	121	0.21	d p
Limons (2-50 µm)	g/kg	141	101	451	458	785	168	0.05	d p
Limons fins (2-20 µm)	g/kg	141	62	244	240	428	74	0.08	d p
Limons grossiers (20-50 µm)	g/kg	141	34	184	217	518	126	0	d p
Sables (50-2000 µm)	g/kg	141	13	247	280	816	182	0.01	d p
Sables fins (50-200 µm)	g/kg	141	8	114	130	564	101	0.31	d p
Sables grossiers (200-2000 µm)	g/kg	141	4	118	149	688	136	0	d p
Limons fin/grossier		141	0.366	1.299	1.437	8.417	0.948	0	d p
Sables fin/grossier		141	0.118	0.941	2.675	34.400	4.801	0.04	d p
ALR limons		141	0.004	0.012	0.012	0.025	0.005	0	d p
ALR sables		141	0.002	0.010	0.011	0.035	0.006	0	d p
Éléments majeurs par extraction totale									
Al total	g/100g	141	1.00	4.40	4.69	9.30	1.63	0.04	p
Fe total	g/100g	141	0.68	2.30	2.65	10.60	1.56	0.33	p
Ca total	g/100g	141	0.50	1.00	34.92	646.00	88.14	0	p
K total	g/100g	141	0.30	1.39	1.49	3.49	0.62	0.15	p
Mg total	g/100g	141	0.05	0.35	0.49	7.49	0.82	0.01	p
Na total	g/100g	141	0.06	0.33	0.45	2.43	0.39	0.16	p
Mn total	mg/100g	141	104.0	581.0	675.0	2140.0	421.1	0.45	p
Mn extractible	mg/100g	141	0.00	0.02	0.05	0.44	0.06	0.58	p
Élément libre									
Fe libre	g/100g	131	0.04	0.27	0.33	1.27	0.21	0.28	d p
Cations échangeables									
CEC Cobalti	cmol+/kg	141	2.18	12.81	15.18	45.00	9.69	0.03	d p
Ca échangeable	cmol+/kg	141	0.19	11.00	13.45	42.30	9.54	0.01	d p
K échangeable	cmol+/kg	141	0.09	0.37	0.41	1.59	0.23	0.03	d p
Na échangeable	cmol+/kg	141	0.01	0.05	0.06	0.16	0.03	0.55	d p
Mg échangeable	cmol+/kg	141	0.12	0.73	1.05	12.00	1.51	0.5	d p
Al échangeable	cmol+/kg	141	0.01	0.04	0.29	7.21	0.90	0	d p
Fe échangeable	cmol+/kg	141	0.00	0.00	0.01	0.06	0.01	0	d p
Na ech/Mg ech		141	0.003	0.070	0.086	0.375	0.059	0.23	d p
S/CEC		141	0.03	0.11	0.12	0.41	0.07	0.35	d p
ESP		141	0.00	0.02	0.03	0.11	0.02	0.43	d p
Na ech/CEC		141	0.001	0.004	0.005	0.023	0.004	0.19	d p
Composition microbiologique									
Biomasse microbienne	µg ADN/g	140	3.10	34.70	48.67	302.60	48.92	0.02	d p
Diversité bactérienne	taxon/g	128	1168	2168	2161	2975	280	0.12	d p

Tableau II.b Résumé des propriétés physico-chimiques du sol des sites RMQS possédant une mesure de la stabilité structurale par la méthode Le Bissonais.

d : descriptif ; p : prédictif.

	Unité	Effectifs	Minimum	Médiane	Moyenne	Maximum	Ecart-type	Test de Kolmogorov-Smirnov	Utilisation par les modèles
Carbone organique (CO)	g/kg	196	5.37	16.91	20.20	79.70	11.84	0	d p
N total	g/kg	196	0.54	1.64	1.87	5.66	0.98	0	d p
C/N		196	7.85	10.13	10.77	27.24	2.40	0	d p
Carbonate	g/kg	196	0.50	1.00	43.53	866.00	125.27	0	d p
pH eau		196	4.0	6.8	6.8	8.4	1.1	0.01	d p
Densité apparente	g/cm3	193	0.77	1.36	1.33	1.84	0.19	0.02	d p
CO/argile		196	0.02	0.07	0.08	0.28	0.04	0	d p
CEC/argile	cmol+/g	196	0.01	0.06	0.06	0.12	0.02	0	d p
Texture									
Argile (< 2 µm)	g/kg	196	81	239	263	648	120	0.21	d p
Limons (2-50 µm)	g/kg	196	101	471	471	789	171	0.05	d p
Limons fins (2-20 µm)	g/kg	196	62	247	245	428	76	0.08	d p
Limons grossiers (20-50 µm)	g/kg	196	34	187	226	518	131	0	d p
Sables (50-2000 µm)	g/kg	196	13	210	265	816	189	0.01	d p
Sables fins (50-200 µm)	g/kg	196	8	103	123	564	95	0.31	d p
Sables grossiers (200-2000 µm)	g/kg	196	4	81	141	688	148	0	d p
Limons fin/grossier		196	0.409	1.260	1.440	8.417	0.957	0	d p
Sables fin/grossier		196	0.118	1.027	2.951	34.400	4.903	0.04	d p
ALR limons		196	0.004	0.011	0.012	0.029	0.005	0	d p
ALR sables		196	0.002	0.010	0.011	0.036	0.006	0	d p
Éléments majeurs par extraction totale									
Al total	g/100g	196	0.60	4.40	4.68	11.60	1.70	0.04	p
Fe total	g/100g	196	0.34	2.30	2.59	10.60	1.48	0.33	p
Ca total	g/100g	196	0.50	1.00	43.06	866.00	124.04	0	p
K total	g/100g	196	0.14	1.41	1.53	3.72	0.65	0.15	p
Mg total	g/100g	196	0.04	0.34	0.45	7.49	0.68	0.01	p
Na total	g/100g	196	0.04	0.40	0.48	2.43	0.37	0.16	p
Mn total	mg/100g	196	60.0	589.0	682.6	2789.0	447.1	0.45	p
Mn extractible	mg/100g	196	0.00	0.02	0.05	0.44	0.06	0.58	p
Élément libre									
Fe libre	g/100g	190	0.04	0.28	0.34	1.27	0.21	0.28	d p
Cations échangeables									
CEC Cobalti	cmol+/kg	196	0.80	13.46	15.41	54.10	9.80	0.03	d p
Ca échangeable	cmol+/kg	196	0.24	11.25	13.58	52.70	9.83	0.01	d p
K échangeable	cmol+/kg	196	0.09	0.38	0.42	1.60	0.24	0.03	d p
Na échangeable	cmol+/kg	196	0.01	0.05	0.06	0.17	0.03	0.55	d p
Mg échangeable	cmol+/kg	196	0.10	0.71	1.02	12.00	1.31	0.5	d p
Al échangeable	cmol+/kg	196	0.01	0.04	0.26	5.48	0.77	0	d p
Fe échangeable	cmol+/kg	196	0.00	0.00	0.01	0.06	0.01	0	d p
Na ech/Mg ech		196	0.006	0.077	0.086	0.367	0.053	0.23	d p
S/CEC		196	0.02	0.11	0.12	0.41	0.08	0.35	d p
ESP		196	0.00	0.03	0.03	0.11	0.02	0.43	d p
Na ech/CEC		196	0.001	0.004	0.005	0.023	0.004	0.19	d p
Composition microbiologique									
Biomasse microbienne	µg ADN/g	196	1.00	33.55	50.24	382.50	53.96	0.02	d p
Diversité bactérienne	taxon/g	167	1168	2149	2158	2975	276	0.12	d p

Tableau III.a Répartition des sites selon la classification des sols du référentiel pédologiques (Baize et al., 1995 ; Baize & Girard, 2008) des sites du RMQS possédant une mesure SLAKES ou Le Bissonnais.

Nom du sol	RMQS	Slakes	Le Bissonnais	Nom du sol	RMQS	Slakes	Le Bissonnais	Nom du sol	RMQS	Slakes	Le Bissonnais
ALOCRI SOL	113		3	HISTOSOL	3			PSEUDO.LUVISOL	7		1
ALUANDOSOL	3			LITHOSOL	9			QUASI.LUVISOL	1		
ANDOSOL	5			LITHOVERTISOL	2			RANKOSOL	76	2	
ANTHROPOSOL	8		1	LUVISOL	255	21	29	REDOXISOL	58	5	8
ARENOSOL	15			MAGNESISOL	2		1	REDUCTISOL	35	2	2
BRUNISOL	556	34	53	NEOLUVISOL	111	15	18	REGOSOL	3		
CALCISOL	134	16	16	ORGANOSOL	2			RENDISOL	30		1
CALCOSOL	319	26	31	PALEOFERSIALSOL	2			RENDOSOL	75	4	4
COLLUVIOSOL	93	5	6	PELOSOL	7			SALISOL	1		
DOLOMITOSOL	5			PEYROSOL	11			THALASSOSOL	9		1
FERSIALSOL	17		1	PLANOSOL	24	3	4	TOPOVERTISOL	1		1
FLUVIOSOL	83	8	15	PODZOSOL	63			VERACRISOL	5		

Tableau III.b Répartition des sites selon la nature du matériau parental des sites du RMQS possédant une mesure SLAKES ou Le Bissonnais.

Nom du sol	RMQS	Slakes	Le Bissonnais
inconnu	85	2	
Alluvions indifferenciées	473	27	42
Roches calcaires	418	30	38
Matériaux argileux	159	18	20
Matériaux sableux	188	7	7
Matériaux limoneux	290	32	55
Formations détritiques	77	4	4
Roches cristallines et migmatites	402	17	26
Roches volcaniques	31	3	3
Autres roches	22	1	1

Tableau III.a Répartition des sites selon l'occupation du sol des sites du RMQS possédant une mesure SLAKES ou Le Bissonnais.

Nom du sol	RMQS	Slakes	Le Bissonnais
Cultures	878	92	126
Prairies	532	31	51
Bois et forêts	583	14	16
Verger et vigne	59	4	3
Friche	20		
Milieux naturels	66		
Parc	6		

Tableau IV. Résultats des tests du chi² de conformité (p-value) des jeux de données SLAKES et Le Bissonnais du RMQS.

	Slakes	Le Bissonnais
Nom Référentiel Pédologique (1995 & 2008)	0.17	0.01
Domaine géologique du matériau parental	0.11	0.01
Occupation du sol	< 1e-6	< 1e-10

c. Sites OS² et Villamblain

Les sols étudiés proviennent de deux sites du projet de OS² et sont des parcelles agricoles, situées sur les communes de Saint-Denis-Des-Puits (SDDP) (28333), Cernay (28067) et une parcelle agricole de Villamblain (45337) (France). La localisation des sites est représentée sur la figure 4. Ces trois parcelles font partie de sites d'étude suivis par l'UR SOLS. Les trois sites sont étudiés car ils présentent une classe de texture identique : limon moyen (triangle de l'Aisne), mais dont la répartition entre les fractions texturales diffère entre les sites (Tableau V). Ces sites sont respectivement deux Luvisols et un Calcisol. Ce sont des parcelles de grandes cultures, SDDP et Villamblain en rotation blé tendre d'hiver, orge d'hiver et colza tandis que Cernay est majoritairement en culture de maïs.

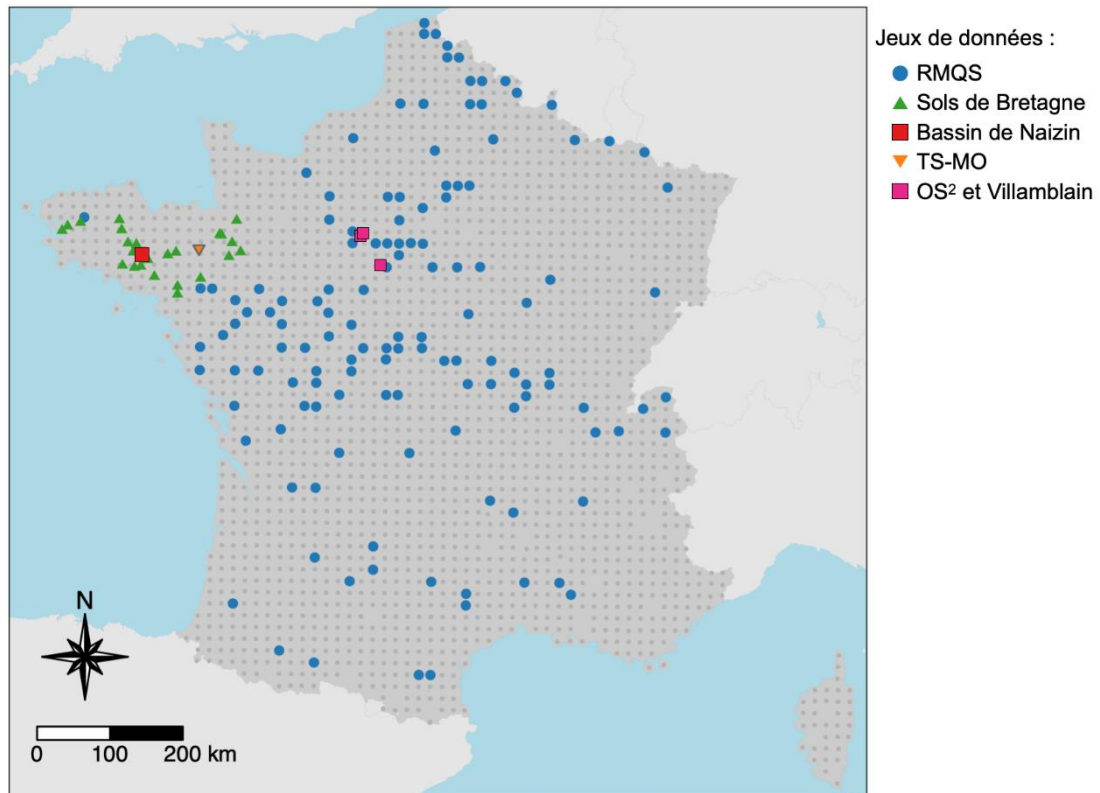


Figure 4. Localisation des sites de l'ensemble des jeux de données étudiées dans le cadre du projet ECLAT (349 sites). rond bleu : RMQS ; triangle vert : Sol de Bretagne ; carré rouge : Bassin de Naizin ; triangle retourné orange : TS-MO ; carré rose bordé : OS² et Villamblain

Tableau V. Résumé des propriétés du sol des sites OS² (Saint-Denis-Des-Puits, Cernay) et Villamblain. Texture selon le triangle de l'Aisne (Delaunois & Maucorps, 1980) et nom du sol selon le référentiel pédologique (Baize & Girard, 2008).

	Argile (< 2 µm)	Limon (2-50 µm)	Sable (50-2000 µm)	Texture	Nom du sol
Saint-Denis-Des-Puits	118	829	53	LM	Luvisol
Cernay	140	750	110	LM	Luvisol
Villamblain	275	600	125	LM	Calcisol

d. Bassin versant de Naizin (ORE AgrHyS)

Le bassin versant de Naizin (ORE AgrHyS) supporte un réseau de mesures géré par UMR SAS INRAE institut Agro. Il représente une zone d'environ 26 km². L'activité agricole est dominante et représente 86,5% de la surface totale. Les mesures SLAKES ont été réalisées par Olfa Tahar dans le cadre de son stage de Master 2 (Tahar, 2021). Un total de 198 sites a été sélectionné de manière à couvrir la zone d'étude de manière homogène (Fig. 5) et analysés en 2013 pour la stabilité structurale selon la méthode Le Bissonnais. Les propriétés pédologiques des sols sont résumées dans le Tableau VI.

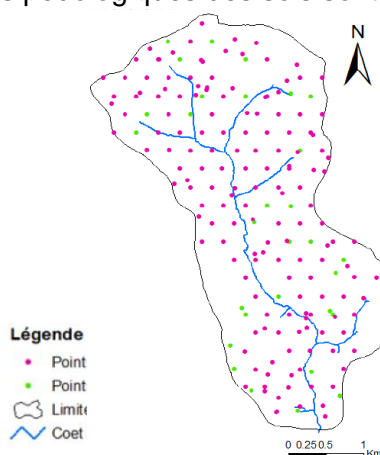


Figure 5. Localisation des sites sélectionnés pour la mesure de la stabilité structurale par la méthode SLAKES dans le bassin versant de Naizin (Tahar, 2021).

Tableau VI. Résumé des propriétés du sol des sites sélectionnés pour la mesure de la stabilité structurale par la méthode SLAKES dans le bassin versant de Naizin (198 sites).

	Unité	Effectifs	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Carbone organique	g/kg	172	14.60	26.35	28.96	99.50	10.92
N total	g/kg	196	1.42	2.31	2.59	14.20	1.19
C/N		171	9.67	11.28	11.33	18.66	0.98
pH eau		198	3.90	6.04	5.96	7.38	0.52
Texture							
Argile (< 2 µm)	g/kg	198	135	174	179	356	32
Limons fins (2-20 µm)	g/kg	198	270	370	372	472	26
Limons grossiers (20-50 µm)	g/kg	198	135	283	277	372	48
Sables fins (50-200 µm)	g/kg	198	39	75	75	139	17
Sables grossiers (200-2000 µm)	g/kg	198	30	88	94	255	35
Éléments majeurs par extraction totale							
Al total	g/100g	198	0.082	0.304	0.307	0.635	0.085
Fe total	g/100g	198	0.116	0.510	0.551	2.190	0.208

e. Sol de Bretagne

Les sols étudiés proviennent de 24 sites prélevés (Fig. 6) dans le cadre d'un projet de recherche en visant à améliorer et diffuser les connaissances pédologiques et mesurer l'évolution de la qualité des sols bretons dans l'objectif de protéger les ressources naturelles et d'adapter les pratiques agricoles (Peluchon, 2021 ; UMR SAS & UMR Ecobio, 2015). Les propriétés pédologiques du sol sont résumées dans le Tableau VII. Le projet sol de Bretagne est financé par le Conseil régional de Bretagne et s'intègre dans le programme Inventaire, Gestion et Conservation des sols (IGCS).

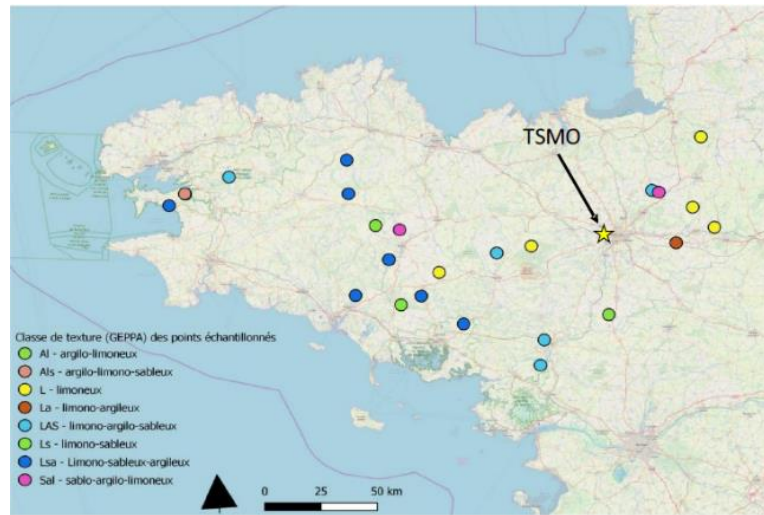


Figure 6. Localisation des sites étudiés provenant du programme Sol de Bretagne et de l'essai TS-MO (Peluchon, 2021)

Tableau VII. Résumé des propriétés du sol des sites du programme Sols de Bretagne (25 sites).

	Unité	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Argile (< 2 µm)	g/kg	60	171	184	380	79
Limons (2-50 µm)	g/kg	280	520	523	775	138
Sable (50-2000 µm)	g/kg	95	262	291	576	139
Carbone organique	g/kg	0.45	2.11	2.27	5.08	1.19
pH eau		3.9	5.7	5.4	6.9	0.8

f. Essai Travail du Sol et Matière Organique (TS-MO) de la plateforme EFluents d'Élevage et Environnement (EFELE) (adapté de Peluchon, 2021)

Le site expérimental de la plateforme EFELE du Système d'Observation et d'Expérimentation au long terme pour la Recherche en Environnement des Produits Résiduels Organiques (SOERE PRO) est situé sur la commune de Le Rheu, en Ille-et-Vilaine (Fig. 7). L'essai TS-MO (travail du sol-matière organique) a pour objectif d'étudier l'impact des effluents d'élevage sur l'environnement et la production agricole. Dans le cadre de l'essai TS-MO, 12 parcelles sont cultivées depuis 2012 pour évaluer l'impact de deux modalités de travail du sol, travail simplifié au Compil (TS) et labour traditionnel

(L), et de la fertilisation, fumier de bovin (FB) et azote minérale (MIN), sur l'environnement et la productivité. La rotation culturale mise en place est blé-maïs avec une interculture de moutarde blanche. Les 12 parcelles sont réparties en 3 blocs de 4 parcelles. Chaque bloc contient les quatre conditions expérimentales de travail du sol et de fertilisation (TS-FB, TS-MIN, L-FB, L-MIN) (Fig. 6). Les échantillons ont été prélevés en mars 2017 pour chaque parcelle dans l'horizon 0-15 cm. Les propriétés pédologiques des sols sont résumées dans le Tableau VIII.



Figure 7. Organisation des parcelles expérimentales de l'essai TS-MO de la plateforme EFELE réparties aléatoirement en 3 blocs. TS : travail simplifié ; L : labour ; MIN : fertilisation minérale ; FB : Fumier de Bovin (Deluchon, 2021)

Tableau VIII. Résumés des propriétés du sol pour les sites de l'essai TS-MO de la plateforme EFELE.

	Unité	Effectifs	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Carbone organique	g/kg	12	10.5	11.3	11.4	12.4	0.6
N total	g/kg	12	1.12	1.20	1.19	1.29	0.05
C/N		12	9.4	9.6	9.6	9.8	0.2
Carbonate	g/kg	12	1.095	1.216	1.211	1.336	0.066
pH eau		12	5.59	5.82	5.82	6.10	0.16
Elements libres							
Al libre	cmol+/kg	12	0.055	0.117	0.113	0.168	0.033
Fe libre	cmol+/kg	4	0.005	0.006	0.007	0.011	0.003
Cations échangeables							
CEC Cobalti	cmol+/kg	12	5.13	5.75	5.79	6.56	0.37
Ca échangeable	cmol+/kg	12	3.91	4.34	4.33	4.77	0.24
K échangeable	cmol+/kg	12	0.152	0.266	0.274	0.431	0.087
Mg échangeable	cmol+/kg	12	0.483	0.580	0.596	0.791	0.111
Na échangeable	cmol+/kg	12	0.033	0.040	0.039	0.045	0.004
Mn échangeable	cmol+/kg	12	0.077	0.106	0.103	0.135	0.019

C. Sélection de l'indicateur SLAKES

Sur chaque site étudié les deux indicateurs SLAKES, Asym et SI600, sont calculés. Les valeurs issues du suivi de la désagrégation des agrégats captée par les différents téléphones (7) pour un site sont regroupées afin d'obtenir une valeur unique. Un site correspond à un lieu physique de prélèvement de sol. Trois méthodes de regroupement sont étudiées : moyenne, médiane et modèle linéaire mixte avec un effet aléatoire correspondant au téléphone. L'objectif est de sélectionner l'indicateur et sa méthode de regroupement, fiable, répétable et sensible aux variations. Pour cela sont évaluées : la répétabilité de l'indicateur, la sensibilité de l'indicateur à la méthode de regroupement et aux différences entre smartphones ainsi que le lien avec le MWD rapide de Le Bissonnais et la transposition de son référentiel d'interprétation.

a. Comparaison des méthodes de regroupement au site

Afin de simplifier la transformation des données issues de l'application, le package « *slakescreation* », automatisant l'analyse des données brutes issues des smartphones, a été développé dans le cadre de ce stage (Annexe 1).

L'indicateur SLAKES final pour un site est obtenu par regroupement des mesures individuelles des agrégats mesurés. Le protocole de la méthode SLAKES implique la mesure de 3 d'agrégats à chaque mesure. Ainsi, le nombre d'agrégats mesuré pour un site est un multiple de 3. Le regroupement des mesures individuelles est réalisée selon trois modalités : une moyenne, une médiane et un modèle mixte linéaire (*model based* : mb). Cette dernière méthode à l'avantage de considérer explicitement dans son calcul l'impact du téléphone utilisé en appliquant à un effet aléatoire lié au téléphone.

On cherche à savoir si la méthode de regroupement des mesures individuelles des agrégats a un impact sur l'indicateur SLAKES final. Les méthodes de regroupement sont comparées statistiquement pour chaque indicateur SLAKES à l'aide d'un test de Wilcoxon de signes et rangs sur la médiane de la différence entre deux méthodes. Pour cela on calcule la différence (D) entre deux méthodes (X_i) pour chaque site (j) : $D = X_{j,1} - X_{j,2}$. Le jeu de données utilisé se compose de 146 sites du RMQS mesurés par l'application SLAKES.

b. Evaluation de la répétabilité de la mesure SLAKES

La répétabilité correspond à la capacité d'un outil de mesure à fournir des résultats homogènes. Dans le cas présent, il n'est pas possible d'utiliser un élément unique pour répéter la mesure et évaluer sa dispersion. On considère équivalent la répétition de la mesure d'agrégats issus d'un site unique. La répétabilité de la mesure SLAKES a été évaluée par l'étude de la dispersion de la mesure SLAKES finale (regroupement des mesures individuelles d'agrégats) en comparant plusieurs sites. On cherche le nombre de répétitions de la mesure de 3 agrégats nécessaire pour que les intervalles de confiance ne se recoupent pas entre les 3 sites.

Dans le cadre de son stage d'IUT, Paul Mousset a réalisé un grand nombre mesures SLAKES, 160 mesures de 3 agrégats (480 agrégats) par site pour 3 sites : Saint-Denis-Des-Puits, Cernay, et Villablain (Mousset, 2021). A partir de ces mesures, on évalue numériquement le nombre de répétitions nécessaires pour que la dispersion de l'indicateur SLAKES final permette de différencier les trois sites. Pour cela, on réalise une simulation numérique qui consiste à tirer aléatoirement k fois $n*3$ valeurs des indicateurs (mesure d'un agrégat), n variant de 1 à 160. La simulation est répétée un

grand nombre de fois : $k = 1000$, pour chaque mesure de n agrégat. Deux sites sont différenciés lorsque, pour un nombre n de répétitions, l'intervalle de confiance à α (90, 95 et 99%) des valeurs de la moyenne de l'indicateurs SLAKES ne se recoupent plus. Le nombre de répétition minimal pour un indicateur est atteint lorsque les trois sites sont différenciés deux-à-deux.

c. Evaluation de la capacité de l'indicateur SLAKES à différencier des pratiques agricoles

Lors de son stage au sein du projet ECLAT, Margot Peluchon (2021) a évalué la capacité de l'indicateur SLAKES Asym à différencier des pratiques agricoles. Cependant, l'utilisation du SI600 et des méthodes de regroupement par site n'ont pas été évaluées, car l'utilisation de cet indicateur n'a été évoquée que plus tard par Flynn et al. (2021). La capacité du SI600 à différencier les pratiques ainsi que l'influence des méthodes de regroupement par site pour les deux indicateurs SLAKES sont ici évaluées.

Les échantillons proviennent des placettes de l'essai TS-MO. Margot Peluchon a réalisé 15 mesures de 3 agrégats (45 agrégats) pour chaque site avec un unique modèle de téléphone. Afin de déterminer le nombre de mesures nécessaire pour différencier pratiques agricoles, on procède de la même manière que pour la répétabilité (section C.b).

d. Comparaison avec le MWD rapide de la méthode Le Bissonnais

Les indicateurs SLAKES, Asym et SI600, sont comparés au MWD_{rapide} de Le Bissonnais afin de déterminer quel indicateur est le plus représentatif de cette mesure de référence. Les corrélations de Spearman entre le MWD_{rapide} et les indicateurs SLAKES, Asym et SI600, sont calculés. Dans un second temps, on représente la distribution du MWD_{rapide} en fonction des indicateurs SLAKES, et le référentiel d'interprétation du MWD_{rapide} de Le Bissonnais est représenté. L'indicateur SLAKES le mieux corrélés au MWD_{rapide} et pour lequel les classes d'interprétations de Le Bissonnais se transposent le mieux est retenu. Les données issues du RMQS, du bassin versant de Naizin, de l'essai TS-MO et de Sol de Bretagne sont utilisées.

L'indicateur SLAKES le moins sensible à la méthode de regroupement, à la plus forte répétabilité, différenciant le plus efficacement les pratiques agricoles et correspondant le mieux au MWD_{rapide} est retenu pour le reste de l'étude.

e. Proposition et évaluation du référentiel d'interprétation SLAKES en France métropolitaine

Un référentiel d'interprétation de cet indicateur *SLAKES* retenu est défini à partir de la distribution du MWD_{rapide} en fonction de l'indicateur *SLAKES* par transposition du référentiel d'interprétation de Le Bissonnais. Afin d'évaluer la précision du nouveau référentiel d'interprétation, on évalue le nombre de sites bien classés par le référentiel proposée par rapport au référentiel Le Bissonnais. On cherchera à minimiser les erreurs de sous-estimation du risque et plus particulièrement pour les catégories « instable » et « très instable ».

E. Etude du lien avec les propriétés pédologiques et prédiction par la modélisation

Deux voies sont explorées. D'une part, (A) une approche explicative à l'aide d'un modèle facilement interprétable afin d'étudier les liens éventuels avec les variables édaphiques et de gestion modulant le comportement des indicateurs. D'une autre part, (B) une approche prédictive à l'aide d'un modèle plus complexe pouvant servir à prédire l'indicateur sur l'ensemble des sites du RMQS afin d'étudier la distribution de l'indicateur en France métropolitaine.

a. Jeux de variables

Tout d'abord, un jeu de données destinés aux (A) modèles descriptifs est défini. Ce jeu de données est construit pour évaluer quelles sont les caractéristiques du sol qui influencent les indicateurs *SLAKES* et Le Bissonnais. Les variables sont sélectionnées à dire d'experts (Frédéric Darboux, INRAe et Didier Michot, Institut Agro) par rapport à leur capacité à moduler la stabilité structurale du sol. Le jeu de données comprend des variables pédologiques classiques (carbone organique, pH eau, carbonates, N total et densité apparente), l'éléments libre (Fe libre), les cations échangeables et la composition microbiologique. La liste des variables physico-chimiques est précisée dans les Tableaux II.a et II.b à l'aide de la notation « d ». De plus, le jeu de données comprend des variables de classification comme le nom du sol selon le référentiel pédologique français (Baize & Girard, 2008) (Tableau III.a), le matériaux parental (Tableau III.b) et l'occupation du sol (Tableau III.c).

Ensuite, est défini un jeu de données destinés aux (B) modèles prédictifs. Il est composé de la totalité du (A) jeu de données descriptif, enrichi des éléments majeurs par extraction totale. La liste de ces variables physico-chimiques est précisée dans les Tableaux II.a et II.b à l'aide de la notation « p ».

b. Données de textures

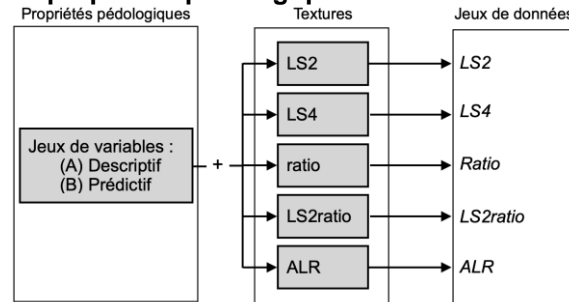
L'utilisation des textures dans un modèle statistique n'est pas simple. En effet, la texture se définissant par la décomposition de la terre fine en fractions granulométriques, ces fractions sont très fortement corrélées entre elles. Plusieurs approches adaptées à la modélisation existent pour définir la texture et certaines ont été explorées (Tableau IX), dont l'approche par *additive log ratio* (ALR) proposée par Lark & Bishop (2007) pour modéliser la texture d'une manière compatible avec des modèles statistiques.

Tableau IX. Description des jeux de données de textures.

Dénomination du jeu de textures	Abréviation	Composition
Deux fractions	LS2	Limons total, sables total
Quatre fractions	LS4	Limons fins, limons grossiers, sables fins, sables grossiers
Ratios limons et sables fins/grossiers	ratio	$\frac{\text{limonfin}}{\text{limongrossier}}$, $\frac{\text{sablefin}}{\text{sablegrossier}}$
LS2 + ratio	LS2ratio	Limons total, $\frac{\text{limonfin}}{\text{limongrossier}}$, sables total, $\frac{\text{sablefin}}{\text{sablegrossier}}$
<i>Additive log ratio</i>	ALR	$\frac{\log_{10}(\text{limontotal})}{\text{argile}}$, $\frac{\log_{10}(\text{sabletotal})}{\text{argile}}$

Ces différents jeux de texture s'ajoutent indépendamment aux jeux de variables descriptifs et prédictifs définis dans la sous-partie précédente (IV.D.a). Les modèles sont donc entraînés sur 5 jeux de données différents par leur jeu de texture (Fig. 8).

Figure 8. Schéma de la construction des jeux de données à partir des propriétés pédologiques et des textures.



c. Ensembles de données

Le jeu de données issu du RMQS est complexe et certaines variables possèdent des données manquantes. L'absence de données peut avoir plusieurs origines. Lors de la première campagne du RMQS, certaines variables n'ont pas été mesurées, comme le Fe libre, et l'acquisition a posteriori n'a pas été possible par impossibilité technique ou manque d'échantillons. Les jeux de données sont filtrés pour retirer les sites aux données manquantes. Le nombre de sites des jeux de données SLAKES et Le Bissonnais réduits sont alors à 100 et 151 sites respectivement, contre 146 et 200 sites initialement. L'intérêt de ces variables pour la modélisation a été étudié, afin de ne pas conserver les variables non utilisées. Dans un premier temps, des modèles sont entraînés sur les jeux de données réduits puis sur des jeux de données sans certaines de ces variables pour évaluer leur intérêt pour le modèle. On cherche à conserver le plus de sites possible car les jeux de données sont assez limités.

Deux propriétés qualitatives ont été écartées pour la modélisation : le matériau parental et le nom du sol. En effet, un trop grand nombre de leurs catégories ne possédaient pas suffisamment de sites, ce qui faussait les résultats de la validation croisée. De plus, pour les variables qualitatives, comme l'occupation du sol et le domaine géologique du matériau parental, les sites des catégories possédant moins de 3 observations sont retirés des jeux de données. Ainsi, les jeux de données ne possèdent plus de sites d'occupation du sol en « vignes et vergers ».

d. Etude des corrélations avec les propriétés physico-chimiques et descriptives du sol

Les corrélations entre les propriétés physico-chimiques du sol ainsi qu'avec les indicateurs de stabilité structurale SLAKES (Asym et SI600) et Le Bissonnais (MWD_{rapide}) sont calculées. De plus, une analyse en composantes principales (ACP) entre les propriétés physico-chimiques et descriptives du sol sont réalisées. Les indicateurs de stabilité structurale sont ajoutés en variables quantitatives illustratives.

e. Modèle par apprentissage automatique

La modélisation est réalisée à l'aide de modèle *cubist* (*Cubist 0.3.0 Kuhn & Quinlan, 2021*). Il s'agit d'un modèle d'apprentissage automatique à arbre de décision pour l'application de modèles linéaires, adapté de Quinlan (1992) par Wang et Witten (1997).

Les programmes d'apprentissage automatique permettent la reconnaissance automatique de schémas dans un jeu de données. Le modèle *cubist* se construit comme la juxtaposition de sous-modèles linéaires séparés selon des règles de décisions (Quilan, 1992 ; Wang & Witten, 1997). L'établissement des règles de décision par le modèle se base sur une réduction de l'erreur de prédiction (Quinlan, 1992). Le modèle *cubist* possède un schéma d'optimisation, appelé *committees*. Ce schéma crée des arbres de décisions de manière itérative selon le nombre de *committees*. Cette approche permet d'ajuster plus précisément le modèle aux données (Cubist Regression Models Cran).

Les modèles descriptifs (A) sont calibrés avec 1 *committee*, ce qui limite la multiplication des règles mises en place par le modèle, et conserve donc une bonne interprétabilité des règles. Tandis que les modèles prédictifs (B) sont calibrés avec 15 *committees*, de manière à permettre un ajustement sur les données sans pour autant sur-ajuster les modèles.

f. Sélection des meilleurs modèles par validation croisée (10 blocs)

La validation croisée à K-blocs (K = 10 dans le cas de la présente étude) est une méthode d'évaluation de la fiabilité d'un modèle statistique préconisé en l'absence d'un jeu de données de validation dédié. Elle permet de tirer plusieurs ensembles de validation d'un même jeu et ainsi d'obtenir une estimation plus robuste de la performance de validation du modèle. Pour cela, on découpe aléatoirement le jeu de données en K-blocs de dimensions similaires. A chaque itération K, on isole un bloc, le jeu de validation et on utilise le reste des blocs pour le jeu d'apprentissage. A partir de ce modèle, on prédit l'indicateur sur le jeu de données d'évaluation que l'on conserve. Une fois les K combinaisons réalisées, on évalue la qualité des prédictions sur l'ensemble des prédictions réalisées grâce aux indicateurs suivants : erreur moyenne (EM) (1), erreur quadratique moyenne (EQM) (2), le carré du coefficient de corrélation de Pearson (r^2) (3) et le coefficient de détermination de Pearson (R^2) (4), définis selon les équations suivantes :

$$EM = \frac{\sum_{i=1}^n pred_i - obs_i}{n} \quad (4) ;$$

$$EQM = \sqrt{\frac{\sum_{i=1}^n (pred_i - obs_i)^2}{n}} \quad (5) ;$$

$$r^2 = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \right)^2 \quad (6) ;$$

$$R^2 = 1 - \frac{SSE}{SST} \quad (7)$$

$$\text{Avec } SSE = \sum_{i=1}^n (pred_i - obs_i)^2 \quad (7.a) \text{ \& } SST = \sum_{i=1}^n (obs_i - \bar{obs})^2 \quad (7.b)$$

soit SSE la somme des carrés des résidus et SST la somme totale des carrés.

Lors de la comparaison des statistiques de modélisation entre les différents jeux de variables, le modèle minimisant les valeurs de l'EM et l'EQM et maximisant les valeurs du r^2 et R^2 et le nombre de sites est retenu. Le nombre de sites ainsi que le R^2 sont considérés comme des statistiques de décision si deux modèles sont très proches. Un modèle entraîné sur un plus grand nombre de site est préféré, car il est considéré comme ayant exploré une plus grande diversité de conditions.

F. Prédiction sur le RMQS

A partir des modèles prédictifs, les valeurs des indicateurs de stabilité SLAKES et Le Bissonnais, sont estimées sur l'ensemble des sites non mesurés par ces méthodes dans le RMQS et correspondant au domaine d'application du modèle. Le RMQS étant un réseau statistiquement représentatif des sols de la France métropolitaine, cette prédiction est une première approche de la stabilité structurale intrinsèque des sols français métropolitains pour les sites correspondant au domaine d'application du modèle. Un jeu de données national est produit en rassemblant les données observées et les données prédites.

L'ensemble des traitements statistiques ont été réalisés sur R 4.0.4 (2021-02-15).

III. Résultats

A. Sélection de l'indicateur

a. Comparaison des méthodes de regroupement au site

La comparaison entre les trois méthodes de regroupement à l'aide d'un test Wilcoxon de signes et rangs sur la médiane de la différence entre deux méthodes, avec $\mu = 0$ et $\alpha = 0.05$, montre que l'ensemble des méthodes de regroupement ont statistiquement un effet sur la valeur final par site de l'indicateur SLAKES (Fig. 9). Les différences les plus importantes sont entre la médiane et les deux autres méthodes de regroupement : moyenne et mb.

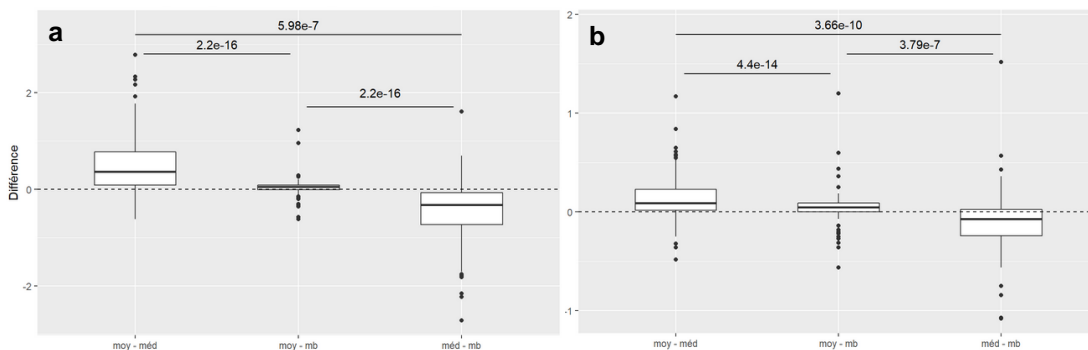


Figure 9. Distribution des différences calculées par site entre les méthodes de regroupement au site pour chaque indicateur SLAKES (a) Asym et (b) SI600 (n = 146).

La médiane de la différence (Tableau X) entre la moyenne et le mb du Asym est environ 6,5 fois inférieures à la différence calculée entre la médiane et la moyenne et le mb en valeur absolue. Pour le SI600, la différence est de 1,4 à 1,8. Cependant, la valeur de cet écart est très faible pour le SI600, lorsque l'on se réfère à la signification réelle d'une variation de l'indice de $\pm 0,1$ reste plus acceptable que le $\pm 0,4$ du Asym. De même pour les écarts interquartiles qui sont beaucoup plus étendus pour le Asym que pour SI600.

La faible différence entre la moyenne et le mb illustre un impact limité du téléphone sur la valeur des indicateurs SLAKES. La valeur au site du SI600 est alors moins sensible à la méthode de regroupement que le Asym (Tableau IX), même si cette différence reste statistiquement significative (Fig. 9).

Tableau X. Médianes et écarts interquartiles des différences calculées entre les méthodes de regroupement au site pour chaque indicateur SLAKES (Asym et SI600) (n = 146).

Différence	Médiane	Ecart interquartile
Asym		
moyenne - médiane	0,36	0,69
moyenne - model based	0,05	0,09
médiane - model based	-0,32	0,67
SI600		
moyenne - médiane	0,09	0,21
moyenne - model based	0,05	0,09
médiane - model based	-0,07	0,27

b. Evaluation de la répétabilité de la mesure SLAKES

L'indicateur nécessitant le moins de répétitions de la mesure de 3 agrégats pour différencier les trois sites est le SI600 médian. Ainsi 30, 40 et 70 répétitions permettent d'espérer une estimation du SI600 avec des intervalles de confiance de 90 %, 95 % et 99% (Tableau XI). Le nombre maximal de répétition étudié, soit 160, ne permet pas au Asym d'atteindre un nombre de répétitions suffisant pour différencier les trois sites. La distribution des indicateurs selon le nombre de répétitions est détaillée en Annexe 2.

Tableau XI. Nombre de répétitions de la mesure de 3 agrégats nécessaires pour différencier les sites Saint-Denis-Des-Prés, Cernay et Villamblain selon chaque indicateur SLAKES : Asym et SI600. "impossible" signifie que l'on ne peut pas différencier les sites avec 160 répétitions.

Intervalle de confiance	Asym			SI600		
	moyen	médian	model based	moyen	médian	model based
90 %	impossible	impossible	impossible	40	30	40
95 %	impossible	impossible	impossible	50	40	60
99 %	impossible	impossible	impossible	90	70	100

c. Evaluation de la capacité de l'indicateur SLAKES à différencier des pratiques agricoles

L'évaluation de l'effet croisé (n = 48) du travail du sol (L = labour et TS = travail simplifié) et du mode de fertilisation (FB = fumier bovin et MIN = minérale) dans des conditions pédologiques proches sur la valeur des indicateurs SLAKES ne donne pas de résultat significatif (Fig 10).

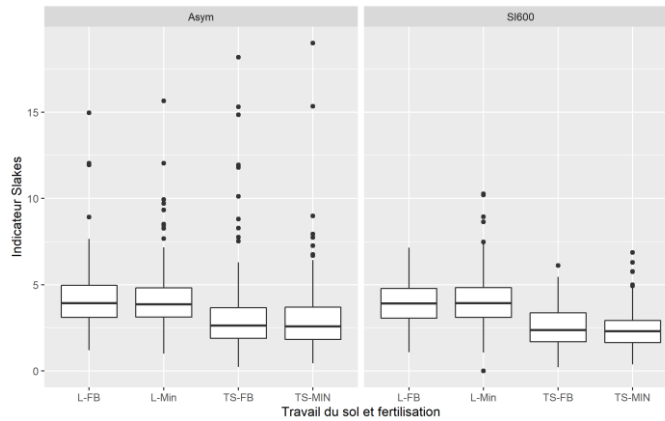


Figure 10. Distribution des indicateurs SLAKES, Asym et SI600, en fonction de l'effet croisé du travail du sol (L = labour ; TS = travail simplifié) et de la fertilisation (FB = Fumier bovin, MIN = minérale)

Les mesures sont regroupées (n = 96) d'un côté selon le travail du sol (L et TS) et d'un autre côté selon le mode fertilisation (FB et MIN) (Fig. 11). La comparaison deux-à-deux à l'aide d'un test de Wilcoxon montre que les modes de travail du sol sont significativement différents (Asym : p-values = $2.29 \cdot 10^{-20}$; SI600 : p-values = $7.55 \cdot 10^{-41}$) tandis que les modes de fertilisation ne sont pas significativement différents (Asym : p-values = 0,65 ; SI600 : p-values = 0,36). Ainsi, les indicateurs SLAKES SI600 et Asym sont significativement différents entre le « labour » et le « travail simplifié ». Le labour augmentant la valeur des indicateurs SLAKES, diminuant donc la stabilité structurale.

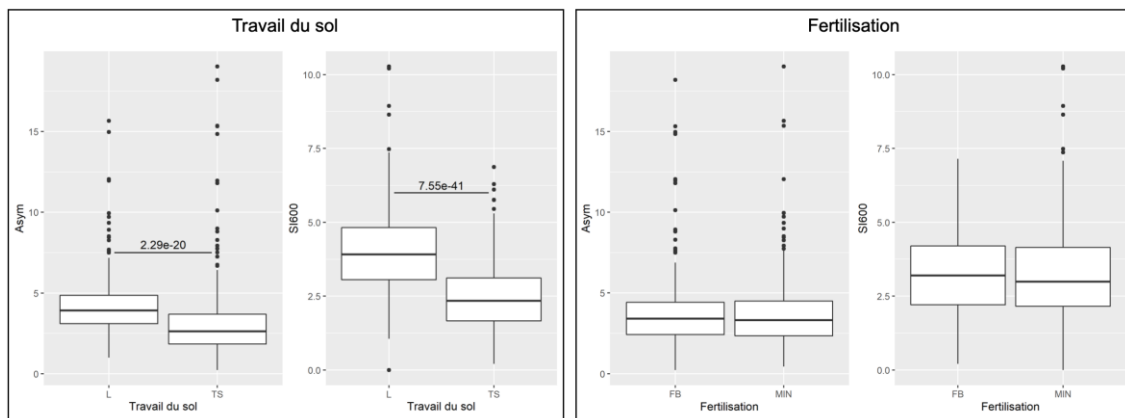


Figure 11. Distribution des indicateurs SLAKES, Asym et SI600 en fonction du (a) travail du sol ou de (b) la fertilisation. p-value du test de Wilcoxon, n = 96.

L'indicateur nécessitant le moins de répétitions de la mesure de 3 agrégats pour différencier les deux modes de travail du sol est le SI600 médian. Ainsi, 10 et 20 répétitions permettent d'espérer une estimation du SI600 avec un intervalle de confiance de 95 % et 99% respectivement (Tableau XII). Le Asym atteint un nombre de répétitions suffisant mais qui est 2 à 6 fois supérieur aux répétitions nécessaires au SI600. La distribution des indicateurs selon le nombre de répétitions est détaillée en Annexe 2.

Tableau XII. Nombre de répétitions de la mesure de 3 agrégats nécessaires pour différencier les modes de travail du sol : labour (L) et travail simplifié (TS) selon chaque indicateur SLAKES : Asym et SI600. “impossible” signifie que l’on ne peut différencier les sites avec 96 répétitions.

Intervalle de confiance	Asym		SI600	
	moyen	médian	moyen	médian
90 %	20	60	10	10
95 %	35	80	15	10
99 %	60	impossible	30	20

d. Comparaison avec la méthode Le Bissonnais

L’indicateur SLAKES le mieux corrélé au MWD_{rapide} de Le Bissonnais est le SI600, r de Spearman = -0,78 (Tableau XIII), avec peu de différences entre les méthodes de rassemblement au site. Le Asym moyen et *model based* sont moins bien corrélé au MWD_{rapide} que le Asym médian, avec $r = -0,68$ contre $r = -0,73$.

Tableau XIII. Matrice de corrélation de Spearman entre le MWD_{rapide} de Le Bissonnais et les indicateurs SLAKES (Asym et SI600) (n = 349).

	Asym			SI600		
	moyen	médian	bodel based	moyen	médian	bodel based
MWD_{rapide}	-0,68	-0,73	-0,68	-0,78	-0,78	-0,79

Si l’on compare la corrélation entre les différents indicateurs SLAKES et leurs méthodes de regroupement (Tableau XIV), les SI600 sont très fortement corrélés entre eux : $r = 0,98$ à $0,99$. Pour le Asym, les Asym moyen et *model based* ont un $r = 1$, tandis que la corrélation de ces deux méthodes avec la médiane est plus faible $r = 0,92$. Le SI600 semble donc moins sensible aux méthodes de regroupement que le Asym. La corrélation la plus élevée entre Asym et SI600 est entre le Asym médian et le SI600 médian ($r = 0,92$). Lorsque l’on observe la distribution du Asym médian en fonction du SI600 médian (Fig. 12), la corrélation linéaire entre les deux indicateurs est très proche de $y = x$, le coefficient directeur étant de 0,96. L’ordonnée à l’origine est de 0,43, soit un décalage de la relation vers Asym médian. Ce qui est cohérent avec la définition des indicateurs. En effet, le Asym correspond à l’étalement à l’infini (Fajardo et al., 2016), il est donc naturellement supérieur à celui observé à 10 minutes (SI600).

La figure 13 représente l’ensemble des relations entre le MWD_{rapide} de Le Bissonnais en fonction des indicateurs SLAKES. La relation entre les indicateurs n’est pas linéaire, mais semble moins bruitée pour le SI600. Le référentiel d’interprétation proposé pour SLAKES (Fajardo et al., 2016 ; Fajardo & McBratney, 2019) ne correspond pas aux classes définies par le référentiel d’interprétation de Le Bissonnais. Néanmoins, l’indicateur SLAKES permettant de retrouver au mieux ces classes de Le Bissonnais est le SI600 (Fig. 13 d, e & f).

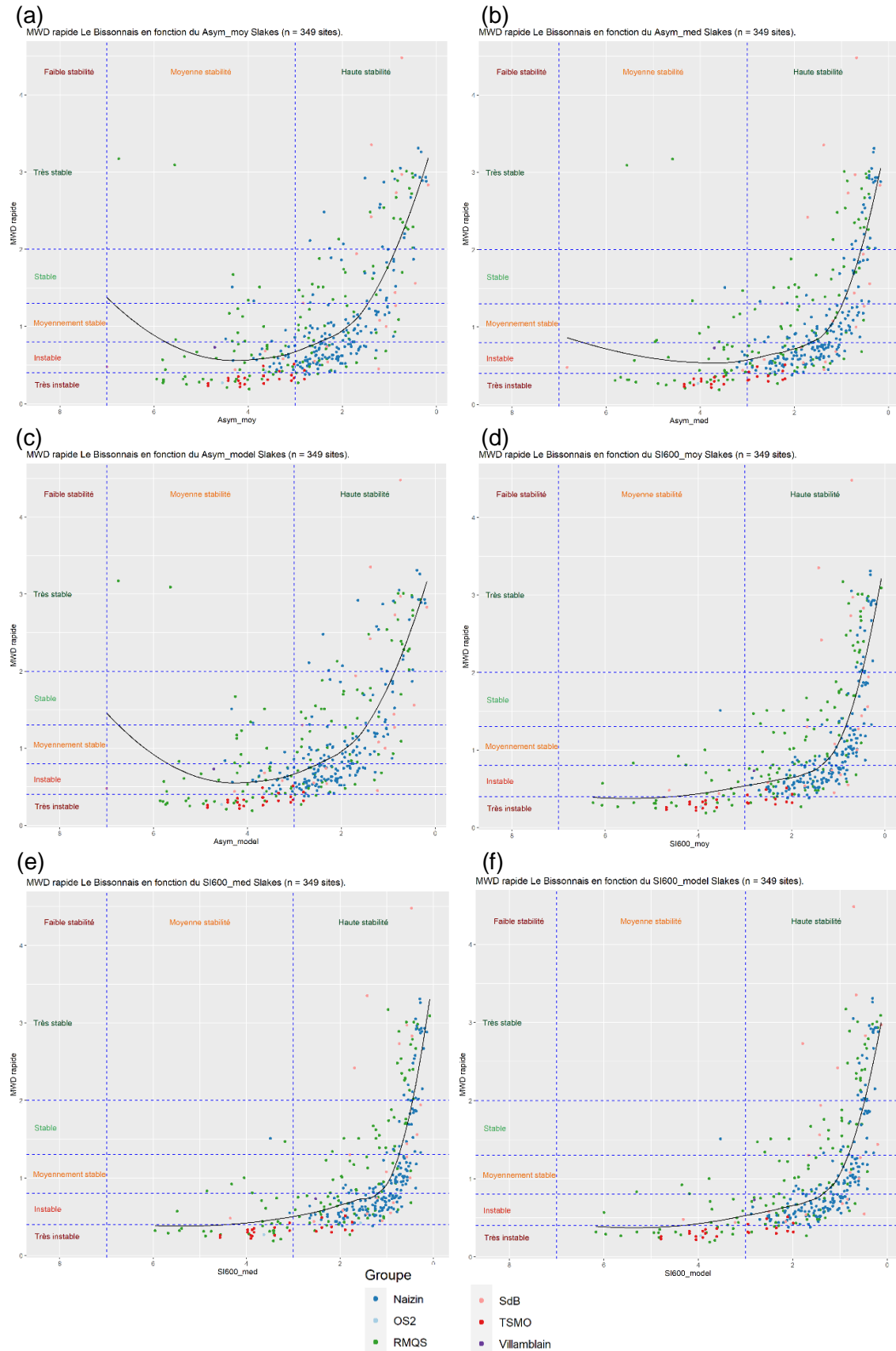


Figure 13. MWD rapide de Le Bissonnais en fonction des indicateurs SLAKES : a. Asym moyen, b. Asym médian, c. Asym *model based*, d. SI600 moyen, e. SI600 médian et f. SI600 *model based*, et représentation des référentiels d'interprétations de Le Bissonnais (1996) et Faiardo et al. (2016).

Tableau XIV. Corrélogramme des coefficients de corrélation de Pearson entre les indicateurs SLAKES (Asym et SI600) (n = 349).

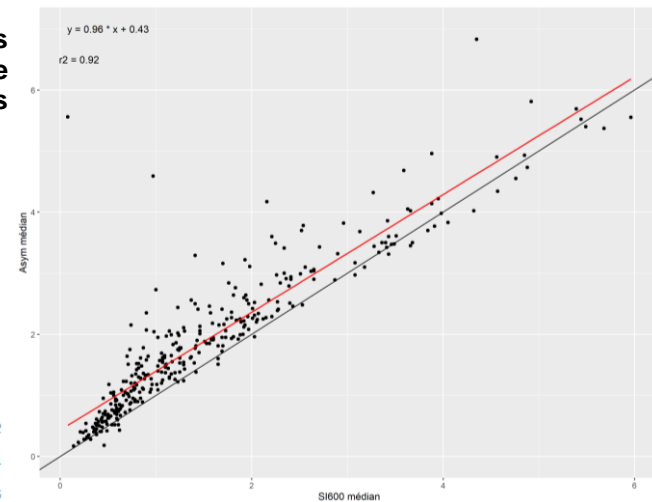
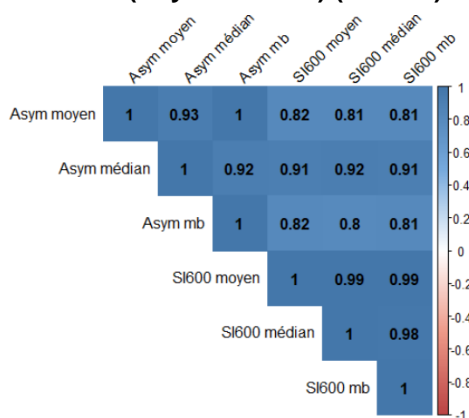


Figure 12. Indicateur SLAKES Asym médian en fonction du SI600 médian (n = 349). noir : $x = y$; rouge = régression linéaire entre x et y .

Au vu des 3 étapes précédentes, le SI600 semble être l'indicateur SLAKES le mieux adapté aux besoins du projet ECLAT. C'est-à-dire que, le SI600 est l'indicateur SLAKES le moins sensible à la méthode de regroupement et à l'effet téléphone (Tableau X), la corrélation avec le MWD_{rapide} est élevée ($r = 0,78$) (Tableau XIII), tout en permettant de retrouver au mieux les classes de référentiel d'interprétation de Le Bissonnais. Enfin, le SI600 médian possède la meilleure répétabilité (Tableau XI) des deux indicateurs SLAKES mais aussi la meilleure capacité de différenciation du travail du sol (Tableau XII). Pour le reste de l'étude, on utilisera le **SI600 médian**.

e. Proposition d'un référentiel d'interprétation

On définit un référentiel d'interprétation du SI600 SLAKES à partir de la projection du SI600 médian en fonction du MWD_{rapide} de Le Bissonnais et de son référentiel d'interprétation (Fig. 14). Les catégories « stable » et « très stable » du référentiel d'interprétation de Le Bissonnais ne sont pas différenciables dans les valeurs du SI600 médian (Fig 14). Les sites de ces deux classes sont rassemblés dans la classe « stable ». On définit le référentiel d'interprétation du SI600 médian en cherchant à minimiser l'erreur d'attribution (Tableau XV). Cependant, il est préféré de surestimer le risque plutôt que de le sous-estimer. Le référentiel d'interprétation de la valeur du SI600 SLAKES est défini dans la table XVI.

Le référentiel proposé permet d'obtenir un taux de bonnes attributions de 89% pour la classe « très instable » et 81% pour « instable » (Tableau XV). Cependant, le taux diminue pour les classes « moyennement stable » et « stable », respectivement 46% et 36%. Malgré tout, lors d'une erreur d'attribution le risque est plutôt surestimé que sous-estimé. La surestimation est 14% pour « instable » et 51% pour « moyennement stable » contre une sous-estimation de 5% et 3% respectivement. La catégorie « stable » est majoritairement surestimée, à 64%. Néanmoins, cette surestimation se traduit majoritairement par un glissement d'une seule classe, vers « moyennement stable » à 40%.

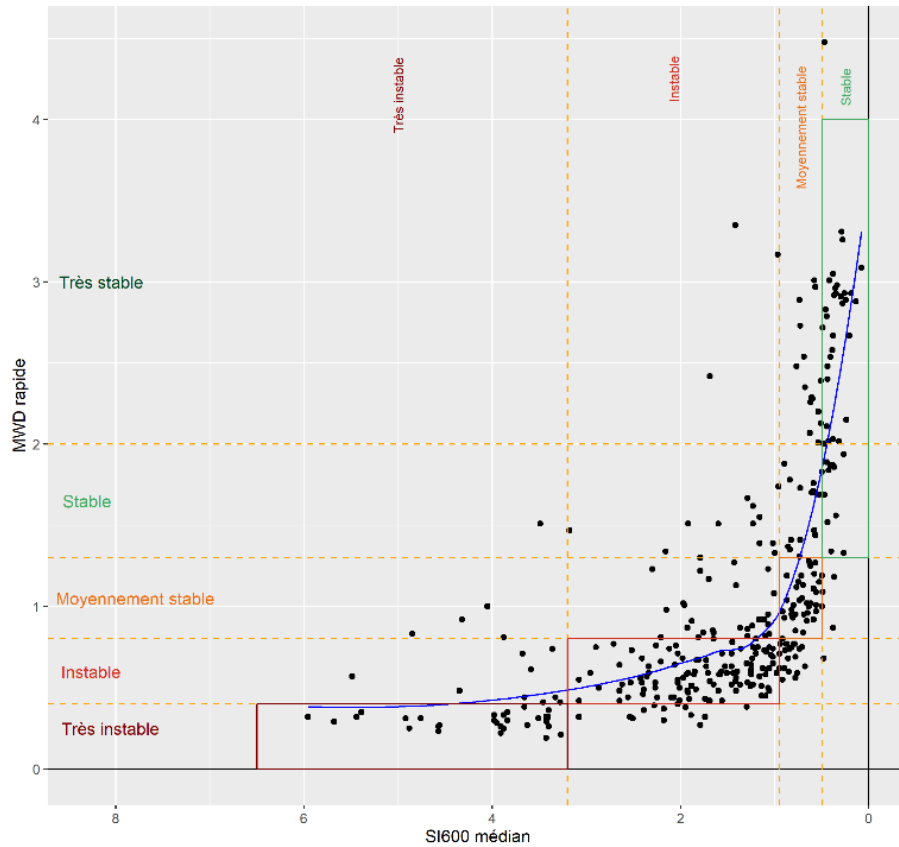


Figure 14. MWD_{rapide} de Le Bissonais en fonction du SI600 médian SLAKES et représentation du référentiel d'interprétation de Le Bissonais (1996) et du référentiel SLAKES proposé (n = 349).

Tableau XV. Bonnes attributions des classes d'interprétation du MWD rapide de Le Bissonais par le référentiel d'interprétation proposé du SI600 médian de SLAKES (en %).

Référentiel d'interprétation Le Bissonais	Référentiel d'interprétation SLAKE proposé			
	Très instable	Instable	Moyennement stable	Stable
Très instable	89	11	0	0
Instable	14	81	5	0
Moyennement stable	6	45	46	3
Stable	1	23	40	36

Tableau XVI. Référentiel d'interprétation de la stabilité des agrégats selon la valeur du SI600 médian de SLAKES

Classe	Valeur du SI600 (%)	Stabilité
1	< 0,50	Stable
2	0,50 - 0,95	Moyennement stable
3	0,95 - 3,20	Instable
4	> 3,20	Très instable

B. Modélisation

a. Corrélation avec les propriétés physico-chimiques et descriptives du sol

L'indicateur SLAKES SI600 et le MWD rapide de Le Bissonnais sont inversement corrélés donc leurs corrélations avec les propriétés physico-chimiques du sol sont elles aussi inversées. Les deux indicateurs sont corrélés significativement ($p > 0,05$) avec des propriétés similaires par ordre décroissant d'importance : limons (total, grossiers, ratio fins grossiers), carbone organique, N total, biomasse microbienne, argiles, CO/argile sables fins/grossiers, C/N, Fe échangeable, Al échangeable, Fe libre, la diversité bactérienne et Mg échangeable (Fig 15). Le MWD rapide de Le Bissonnais est aussi corrélés avec à la densité apparente, Ca échangeables, pH, ALR sables, CEC, K échangeable et le CEC/argile. La majorité des corrélations du MWD_{rapide} sont plus importantes que pour le SI600, de 38% en moyenne, mis-à-part les limons, limons grossiers et les sables auxquels le SI600 est mieux corrélé, de 33% en moyenne.

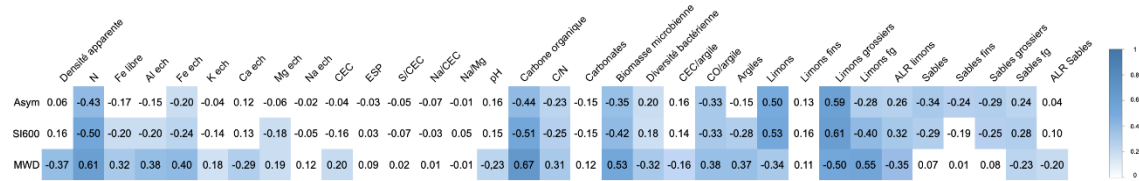


Figure 15. Coefficients de corrélation de Pearson (r) entre indicateurs de stabilité structurale SLAKES (Asym, SI600) ($n = 117$) et Le Bissonnais (MWD_{rapide}) ($n = 151$) et les propriétés de sols. Les cases colorées présentent une corrélation significative p -valeur < 0.05 et l'intensité correspond à l'importance de la valeur absolue de la corrélation. éch = échangeable ; CO = carbone organique ; fg = ratio fins/grossiers ; ALR = additive log ratio.

Pour les deux indicateurs SLAKES et Le Bissonnais, on observe une graduation selon l'occupation du sol (Fig 16). La stabilité structurale étant croissante selon l'ordre : cultures, bois et forêts et prairies.

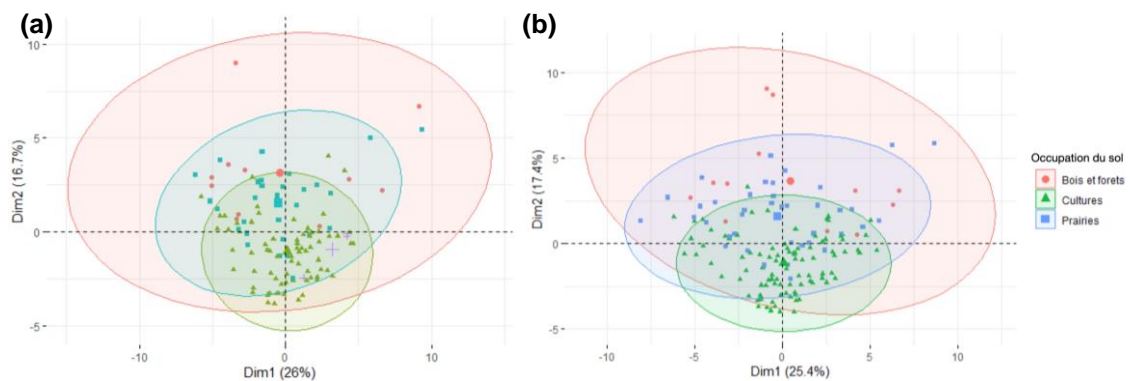


Figure 16. Représentation des individus en analyse en composante principale (ACP) selon l'occupation du sol pour le (a) SI600 médian SLAKES et le (b) MWD_{rapide} de Le Bissonnais.

Cependant, aucune influence nette des matériaux parental, du domaine géologique du matériaux parental ou du nom du sol n'est visible. Le reste l'ACP est disponible en Annexe 3.

b. Sélection des modèles descriptifs

La sélection des sites sans données manquantes réduit le nombre de sites SLAKES et Le Bissonnais à 100 et 151 respectivement, contre 146 et 200 sites initialement. Un modèle *cubist* est entraîné sur un jeu de variables descriptif *modulo* un jeu de textures. Les statistiques d'évaluation des modèles sont meilleures pour le MWD_{rapide} que pour le SI600 (Tableau XVII). Le jeu de texture le plus adapté pour les deux indicateurs est le LS4 (4 fractions : limons fins, limons grossiers, sables fins et sables grossiers).

Tableau XVII. Statistiques de la validation croisée (10 blocs) des modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

(a) Texture set	Effectif	EM	EQM	r ²	R ²	(b) Texture	Effectif	EM	EQM	r ²	R ²
LS2	100	-0.02	1.13	0.34	0.33	LS2	151	-0.01	0.58	0.52	0.48
LS4	100	0.00	1.08	0.39	0.39	LS4	151	-0.01	0.56	0.54	0.51
ratio	100	-0.07	1.37	0.17	0.01	ratio	151	-0.03	0.57	0.54	0.50
LS2ratio	100	-0.01	1.15	0.32	0.30	LS2ratio	151	-0.02	0.57	0.53	0.49
ALR	100	-0.08	1.29	0.20	0.13	ALR	151	-0.01	0.57	0.53	0.50

Ensuite, afin d'améliorer le nombre de sites dans les jeux de données, l'utilisation des variables aux données manquantes est évaluée (Tableau XVIII). La diversité bactérienne n'est utilisée par aucun des modèles du SI600. Les jeux de données sont adaptés selon l'utilisation des variables et les variables non utilisées sont retirées du jeu de données. Par exemple, le jeu de données LS2 (2 fractions : limons et sables) du SI600 n'utilise que la biomasse microbienne. Les variables Fe libre et diversité bactérienne sont retirées. Dans le nouveau jeu de données, les sites sans données sur la biomasse microbienne sont retirés. On note ce nouveau jeu de données « bm » pour biomasse microbienne.

Tableau XVIII. Utilisation des variables présentant des données manquantes par les modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais

(a)	LS2	LS4	ratio	LS2ratio	ALR	(b)	LS2	LS4	ratio	LS2ratio	ALR
Fe libre					X	Classification BioSiSol					
Diversité bactérienne						Densité apparente				X	
Biomasse microbienne	X		X	X	X	Fe libre				X	
Classification BioSiSol						Diversité bactérienne		X			

Des modèles *cubist* sont entraînés sur les nouveaux jeux de données pour chaque indicateur de stabilité structurale SLAKES et Le Bissonnais. Les statistiques de validation sont meilleures que pour les jeux de données précédents (Tableau XIX). Les jeux de données possédant les meilleures statistiques de validation croisée pour le SI600 sont LS2 et ALR. Cependant, le jeu de données LS2 possède plus de sites que l'ALR, 121 contre 111 soit 9% de plus, pour des statistiques proches. Le modèle entraîné

sur le jeu de données LS2 est donc retenu. Pour le MWD_{rapide}, les jeux de données LS2 et ratio ont des statistiques équivalentes. Cependant, le jeu ratio possède 31 sites supplémentaire (+ 20%) que le jeu LS4, on retient donc le jeu de données ratio.

Tableau XIX. Statistiques de la validation croisée (10 blocs) des modèles *cubist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais. Filtrés : nf = non filtré ; bm = biomasse microbienne ; bmf = bm + Fe libre ; db = diversité bactérienne ; daf = densité apparente + Fe libre.

(a) Filtre							(b) Filtre						
Filtre	Effectifs	Texture	EM	EQM	r ²	R ²	Filtre	Effectifs	Texture	EM	EQM	r ²	R ²
bm	121	LS2	-0.06	1.08	0.42	0.42	nf	189	LS2	-0.03	0.60	0.47	0.42
nf	115	LS4	-0.06	1.14	0.37	0.36	db	158	LS4	-0.03	0.60	0.50	0.46
bm	121	ratio	-0.07	1.25	0.26	0.23	nf	189	ratio	-0.07	0.59	0.49	0.45
bm	121	LS2ratio	-0.08	1.13	0.38	0.37	daf	180	LS2ratio	-0.06	0.62	0.42	0.37
bmf	111	ALR	-0.09	1.08	0.44	0.44	nf	189	ALR	0.00	0.63	0.48	0.37

c. Règles de pédotransfert

Le modèle *cubist* du SI600 médian SLAKES avec pour jeu de texture LS2, ne présente qu'une unique règle (Fig. 17). Le carbone organique et le Ca_{ech} réduisent la valeur du SI600 tandis que la CEC, les limons et les sables l'augmentent.

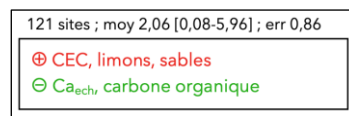


Figure 17. Règle de pédotransfert du modèle *Cubist* de l'indicateur de d'instabilité structurale SLAKES SI600.
ech = échangeable

Le modèle *cubist* du MWD rapide de Le Bissonnais avec pour jeu de texture LS4, présente deux groupes de règles (Fig 18). D'un côté (A) la matière organique et de l'autre (C) l'occupation du sol en "Cultures". Dans le premier cas (A), les valeurs du MWD rapide vont de 0,76 à 2,48 en moyenne ce qui correspond à "instable" à "très stable". Tandis que dans le second cas (C), elles vont de 0,58 à 1,80 soit "instable" à "stable". Dans le premier bloc, la teneur en carbone organique joue un rôle avec des effets seuils. Au-dessus de 51,60 g/kg, la stabilité calculée par le modèle est fixe à 2,34 soit "très stable". En dessous du seuil de 35,80 g/kg de carbone organique (B), l'Al_{ech} joue un rôle de seuil à 0,068 cmol+/kg. Pour les sites avec des valeurs supérieures à ce seuil, on distingue l'occupation du sol "Prairies" et "Bois et Forêts". Dans le second bloc (C), si l'occupation du sol est "Cultures", la teneur en Al_{ech} joue un rôle de seuil à 0,080 cmol+/kg. Au-dessus de ce seuil, les sols ont pour moyenne 0,58, soit "instable", et en dessous, mais supérieur à 0,68 cmol+/kg d'Al_{ech} la moyenne est de 1,80, soit "stable".

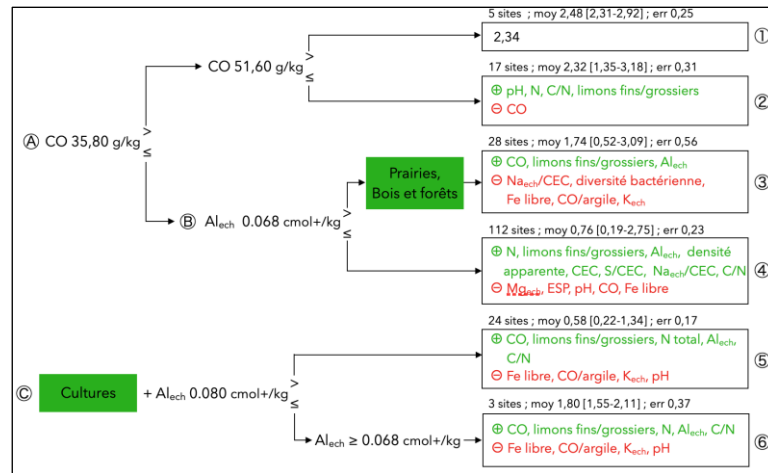


Figure 18. Règles de pédotransfert du modèle *Cubist* de l'indicateur de stabilité structurale MWD rapide de Le Bissonais. CO = carbone organique, ech = échangeable.

Les distributions du SI600 en fonction des propriétés de sol utilisées par le modèle *cubist* sont représentées (Fig. 19). On remarque que le Ca_{ech} (Fig 19.4) et le carbone organique (Fig 19.5) sont corrélés négativement au SI600, ce qui correspond à l'utilisation faite par le modèle (Fig 17). De même pour les limons qui sont corrélés positivement au SI600 (Fig 19.2). Cependant, on observe une corrélation négative avec la CEC (Fig 19.1) et les sables (Fig 19.3), ce qui va à l'inverse du modèle *cubist*.

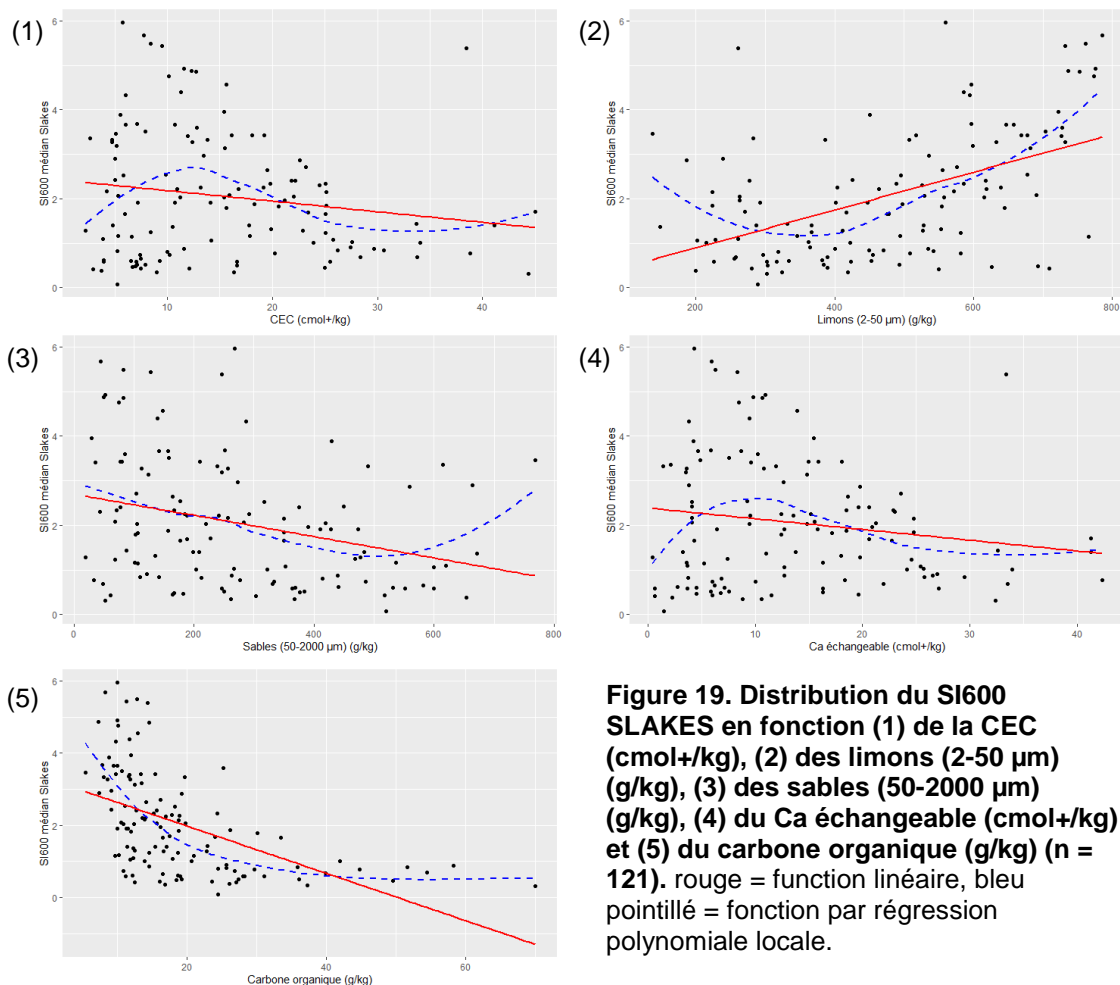


Figure 19. Distribution du SI600 SLAKES en fonction (1) de la CEC (cmol+/kg), (2) des limons (2-50 μm) (g/kg), (3) des sables (50-2000 μm) (g/kg), (4) du Ca échangeable (cmol+/kg) et (5) du carbone organique (g/kg) (n = 121). rouge = fonction linéaire, bleu pointillé = fonction par régression polynomiale locale.

On représente, le MWD_{rapide} en fonction des propriétés du sol pour lesquelles le modèle *cubist* a mis en place des conditions (Fig. 20). La distribution du MWD_{rapide} en fonction du carbone organique a été découpée en trois parties par le modèle. La première et la 2nd partie présentent une corrélation positive et la 3^{ème} une corrélation négative, ce qui est l'inverse du modèle *cubist* . La corrélation négative de la 2nd partie est observable lorsque cette partie est considérée indépendamment (Fig. 20.1). Les seuils fixés pour l' Al_{ech} se comprennent par l'accumulation de données très faibles et quelques sites aux données supérieures (Fig. 20.2 et 20.3).

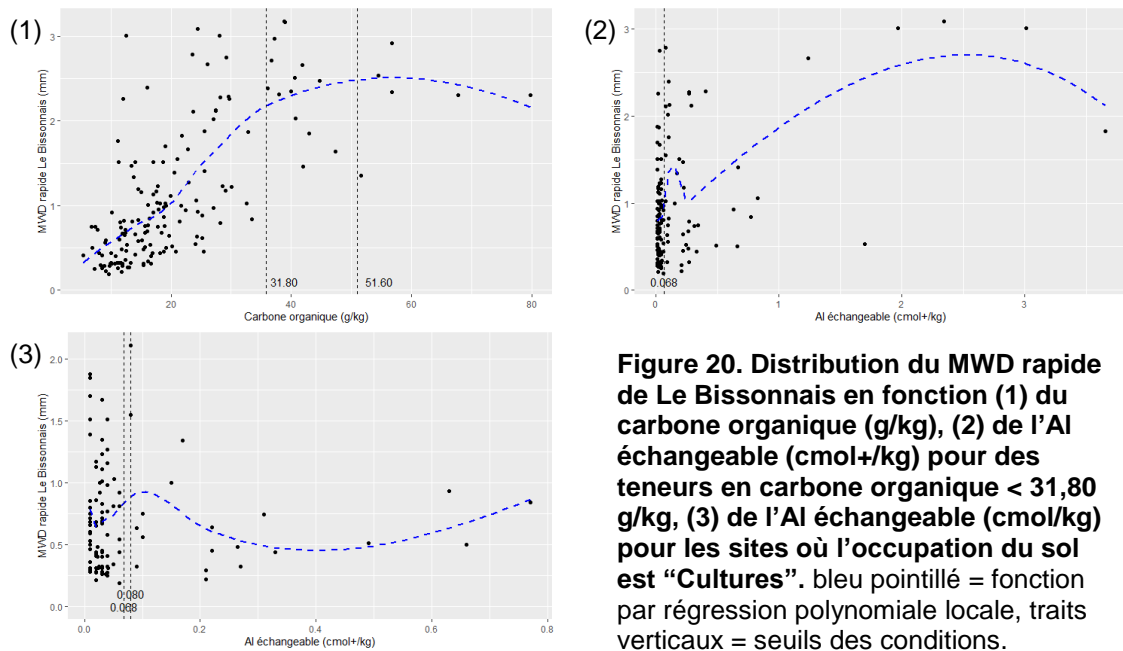


Figure 20. Distribution du MWD rapide de Le Bissonnais en fonction (1) du carbone organique (g/kg), (2) de l'Al échangeable (cmol+/kg) pour des teneurs en carbone organique < 31,80 g/kg, (3) de l'Al échangeable (cmol/kg) pour les sites où l'occupation du sol est "Cultures". bleu pointillé = fonction par régression polynomiale locale, traits verticaux = seuils des conditions.

g. Modèles prédictifs

Les modèles *cubist* à 15 *committees* sont entraînés sur les jeux de données prédictifs réduits (B). Les jeux de textures qui permettent de prédire au mieux les indicateurs de stabilités structurale sont ceux utilisant le LS4 (Tableau XX). Les statistiques sont meilleures que pour les modèles descriptifs équivalents (Tableau XVII).

Tableau XX. Statistiques de la validation croisée (10 blocs) des modèles *cubist* prédictif pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais.

(a)	Texture	Effectif	EM	EQM	r ²	R ²
	LS2	100	-0.01	1.08	0.39	0.39
	LS4	100	-0.08	1.05	0.43	0.42
	ratio	100	-0.05	1.19	0.29	0.26
	LS2ratio	100	0.01	1.09	0.37	0.37
	ALR	100	-0.02	1.05	0.42	0.42

(b)	Texture	Effectif	EM	EQM	r ²	R ²
	LS2	151	-0.01	0.58	0.52	0.48
	LS4	151	-0.01	0.56	0.54	0.51
	ratio	151	-0.03	0.57	0.54	0.50
	LS2ratio	151	-0.02	0.57	0.53	0.49
	ALR	151	-0.01	0.57	0.53	0.50

Les jeux de données sont adaptés selon l'utilisation des variables aux données manquantes. Parmi les nouveaux jeux de texture, le plus adapté pour le SI600 SLAKES semble être celui utilisant le LS2 et celui utilisant le LS2ratio pour le MWD_{rapide} de Le

Bissonnais (Tableau XXI). L'optimisation de jeux de données permet dans les deux cas d'améliorer les statistiques de validation des modèles. Les modèles possèdent encore une fois de meilleurs statistiques que les modèles descriptifs équivalents (Tableau XVIII). Les variables utilisées par les modèles sont précisées dans l'Annexe 4.

Tableau XXI. Statistiques de la validation croisée (10 blocs) des modèles *cbist* pour chaque jeu de texture (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais. Filtres : bm = biomasse microbienne ; bmf = bm + Fe libre ; daf = densité apparente + Fe libre ; dafdb = daf + diversité bactérienne.

(a) Filtre	Effectifs	Set	EM	EQM	r ²	R ²	(b) Filtre	Effectifs	Set	EM	EQM	r ²	R ²
bm	121	LS2	-0.05	1.05	0.45	0.45	daf	180	LS2	-0.04	0.52	0.57	0.56
bmf	111	LS4	-0.06	1.10	0.42	0.42	daf	180	LS4	-0.03	0.53	0.55	0.54
bm	121	ratio	-0.04	1.20	0.29	0.29	dafdb	160	ratio	-0.04	0.53	0.55	0.54
bm	121	LS2ratio	-0.08	1.06	0.44	0.44	daf	180	LS2ratio	-0.06	0.51	0.59	0.58
bmf	111	ALR	-0.03	1.11	0.41	0.41	dafdb	160	ALR	-0.04	0.52	0.58	0.57

C. Prédications sur le RMQS

Tout d'abord, le jeu de données du RMQS (2145 sites) est filtré pour retirer les sites dont les catégories des variables qualitatives ne sont pas présentes dans les jeux de données d'entraînement des modèles. Ainsi, l'occupation du sol est limitée à "cultures", "prairies" et "bois et forêts" et le domaine du matériau parental à "cristallin" et "sédiments". Le jeu de données obtenu contient 1864 sites. A partir de ce jeu de données sont prédites les valeurs des indicateurs de stabilité structurale SI600 et MWD_{rapide} grâce aux modèles prédictifs définis précédemment. Le total de sites observées est de 1874 sites pour le SI600 SLAKES et 1875 sites pour le MWD_{rapide} Le Bissonnais. La distribution des valeurs du SI600 est asymétrique à droite avec une plus grande densité pour les valeurs "moyennement stable" (Fig. 21.a). Tandis que la distribution du MWD_{rapide} est plus symétrique au tour de la classe "moyennement stable" (Fig 21.b). On remarque des effets de seuils plus marqués pour le MWD_{rapide} que pour SI600. Notamment à environ 0,6, 1,5 et 2,2. Ces effets de seuils doivent résulter des conditions mises en place par le modèle prédictif.

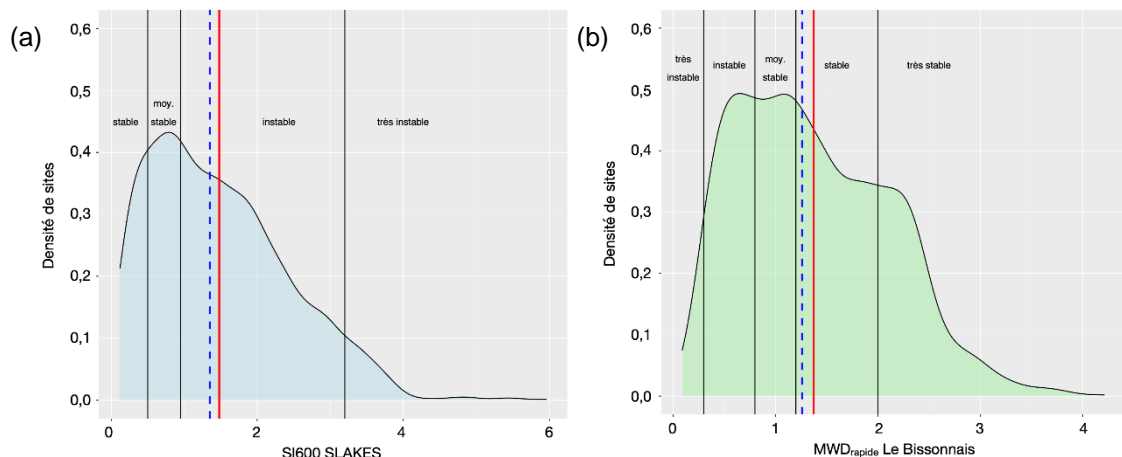


Figure 21. Distribution des valeurs de stabilité structurale (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais observées (resp. 132 ; 183) et prédites (resp. 1874 ; 1875) du le RMQS. trait rouge = moyenne ; trait bleu pointillé = médiane.

Les sites sont ensuite classés selon les référentiels d'interprétations adaptés. Les distributions des classes de stabilité sont représentées par les figures 24 et 25. Le nombre de sites dans les classes "moyennement stable" et "très instable" est du même ordre de grandeur pour les deux indicateurs (Fig. 22), respectivement 25% et 15% de différence. Cependant, dans les prédictions du SI600 SLAKES, la catégories "instables", est 2,9 fois plus importante que pour le MWD_{rapide} . Inversement pour les classifications "stable" - "très stable" du MWD_{rapide} qui sont 3 fois importantes que la classe « stable » du SI600 SLAKES. Les prédictions du SI600 montre une majorité de sites « instables » quand le MWD_{rapide} donne une majorité de sites « stables » - « très stable ».

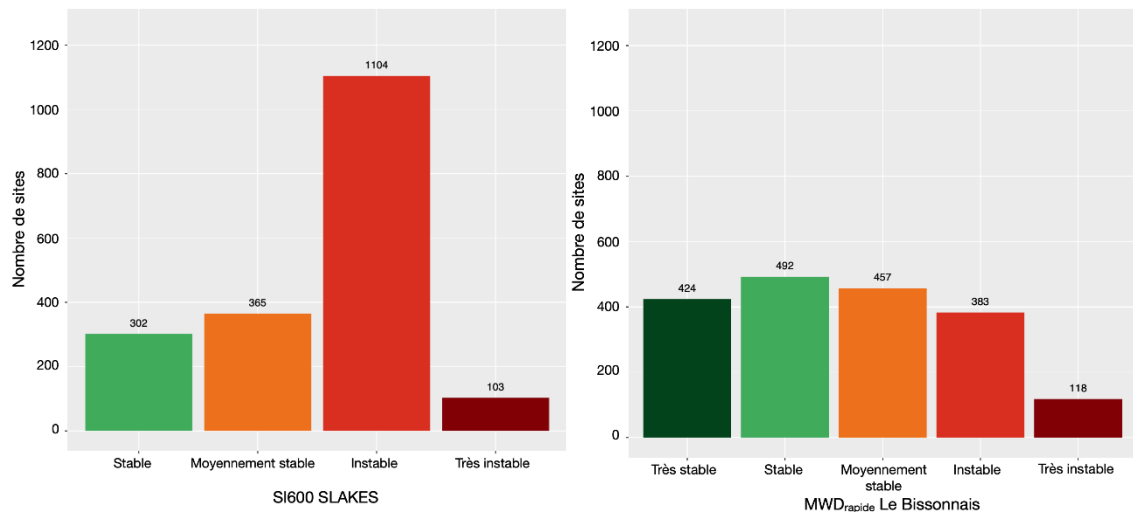


Figure 22. Distribution des sites par classe de stabilité structurale (a) SI600 SLAKES et (b) MWD_{rapide} Le Bissonnais, sites observées 132 et 183 respectivement, et sites prédits 1742 et 1692 respectivement, du RMQS selon le référentiel adapté.

Les cartes des indicateurs de stabilité (Fig. 23) sont semblables dans leurs répartitions globales. Les zones « stables » et « très instables » sont très semblables.

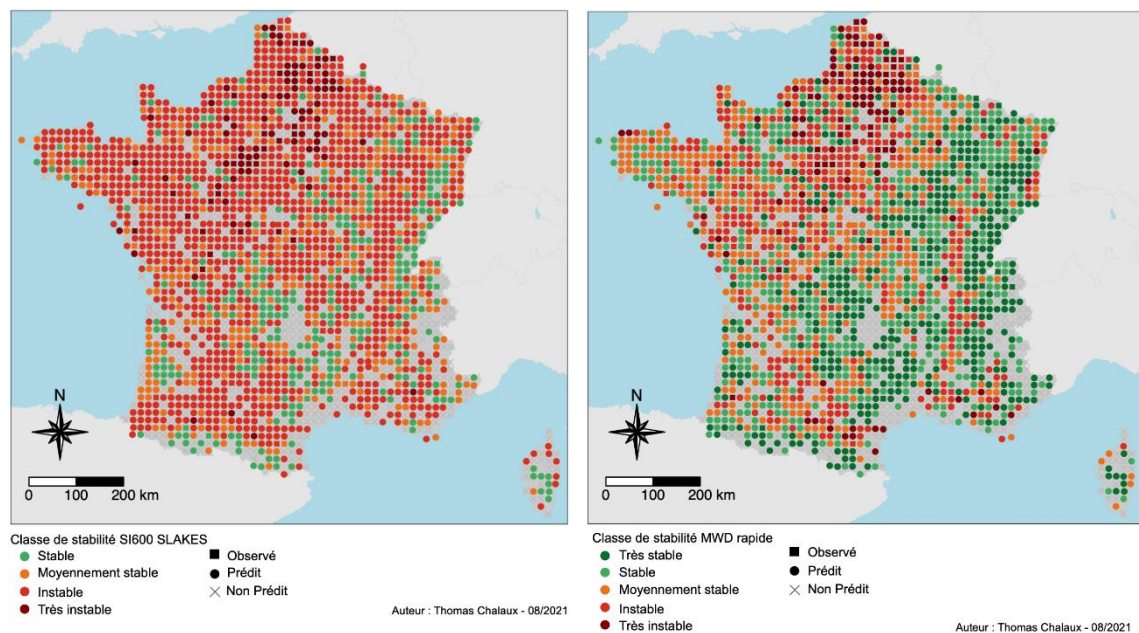


Figure 23. Carte de France de la stabilité structurale des agrégats du sol selon (a) l'indicateur SI600 SLAKES et (b) l'indicateur MWD_{rapide} Le Bissonnais, nombre de sites observés 132 et 182 respectivement et sites prédits 1742 et 1692 respectivement.

Cependant, le SI600 présente de très grandes zones « instable » homogènes quand le MWD_{rapide} y apporte des nuances. Pour les deux indicateurs, les grands ensembles pédologiques français métropolitain sont visibles. Le Bassin Parisien, le Massif Armoricain, le Bassin Aquitain, le Bassin du Rhône et de la Saône, la région méditerranéenne sont des zones de faible stabilité, classées “instables” à “très instables”. Le bassin parisien est le lieu majoritaire des sites “très instables”. Tandis que le Massif Central, les Pyrénées, les Alpes et les Vosges sont des zones plus stables, de “moyennement stable” à “très stable”.

IV. Discussion

A. Jeux de données

Le jeu de données issue du RMQS présente une grande complexité. En effet, la sélection des sites n'a pas pu se faire *a priori* mais en fonction des données et échantillons disponibles ainsi que de l'avancement de la 2nd campagne du RMQS. Néanmoins, les échantillons de la campagne 2 proviennent d'un travail de sélection de sites permettant de bien caractériser les distributions des propriétés physiques du RMQS. De plus, certains sites de la 2nd campagne du RMQS possédaient une mesure SLAKES mais certaines de leurs variables physico-chimiques n'étaient pas encore acquises. De manière à éviter ces données physico-chimiques manquantes, les données de ces mêmes sites issues de la 1^{ère} campagne du RMQS ont été utilisées. Ainsi, certains sites possèdent des données physico-chimiques « composite » des deux campagnes du RMQS. Le biais introduit est difficile à quantifier. Cependant, la constitution progressive de la base de données de la 2nd campagne du RMQS devrait faire disparaître ces manques à l'avenir.

Les jeux de données SLAKES et Le Bissonnais mis en place dans le cadre du projet ECLAT sont, pour la majorité des variables physico-chimiques et descriptives, statistiquement représentatifs de la distribution à l'échelle du RMQS (Tableau II.a et II.b). Néanmoins, quelques variables d'importances ne le sont pas, comme le carbone organique, l'azote total et les carbonates. Ces variables sont fortement corrélées aux indicateurs SLAKES et Le Bissonnais (Fig. 18) et sont utilisées par les modèles descriptifs (Fig. 20 et 21). Il faudrait donc améliorer le jeu de données de manière à atteindre une meilleure représentativité de la distribution du RMQS pour toutes les variables physico-chimiques. A plan d'expérience permettrait de rechercher les sites du RMQS à rajouter pour atteindre cette représentativité, par exemple grâce à un hypercube latin (Minasny & McBratney, 2006 ; Louis et al., 2014 ; Clifford et al., 2015 ; Landré et al., 2018). Dans les jeux de données l'occupation du sol « cultures » et « prairies » sont sur-représentés. Cependant, ce biais est constitutif du projet ECLAT qui s'oriente vers l'évaluation de cette méthode dans le contexte agricole. En effet, l'évaluation des risques érosif est majoritairement réalisée dans le contexte agricole de grandes cultures. Le manque de sites en « vergers et vignes », plus particulièrement de « vignes », représente un problème car ce sont des lieux connus pour leur sensibilité à l'érosion (Paroissien et al., 2010 ; Proscocimi et al., 2016). Sachant que, l'étude de Le Bissonnais et al. (2007) a montré la bonne adéquation de l'indicateur Le Bissonnais dans ce contexte.

B. Répétabilité de la mesure et sélection de l'indicateur SLAKES

L'indicateur SLAKES le plus adapté pour faciliter l'acquisition de données sans pour autant renoncer à sa précision, est le SI600 médian. Le nombre minimal de répétitions est de 30 à 70 mesures de 3 agrégats (90 à 210 agrégats) pour un niveau de confiance de 90 à 99% (Tableau X et XI). Le nombre 40 mesures (120 agrégats), permettant d'atteindre un niveau de confiance à 95%, nous semble accessible pour la mise en œuvre de la méthode. Le nombre de répétitions de la mesure déterminé ici dépasse largement la recommandation de 3 répétitions faite par Fajardo & McBratney (2021). Dans les conditions étudiées 3 répétitions (9 agrégats) ne permettent pas de différencier statistiquement les sols. La nécessité d'utiliser un grand nombre d'agrégats peut s'expliquer par la nature très variable de la stabilité structurale, dans l'espace et dans le temps (Amézketa et al., 1999 ; Annabi et al., 2017 ; Algayer, 2012). Ainsi, l'utilisation d'un grand nombre d'agrégats permet de lisser ces variations pour déterminer un comportement moyen du sol. Le nombre d'agrégats proposé ici, 90 à 120, reste jusqu'à deux fois inférieur à celui utilisé par la méthode Le Bissonais, 150 à 200 agrégats (Mousset, 2021).

L'algorithme de reconnaissance des agrégats utilisé par l'application pourrait être adapté au contexte français. Pour reconnaître les agrégats l'algorithme procède à un seuillage sur la bande bleu à un seuil fixé (Fajardo et al., 2016). Or, l'algorithme reconnaît difficilement les sols clairs (Mousset, 2021) ce qui complique l'acquisition de données et leur fiabilité. Ce qui peut s'expliquer par l'origine Australienne de la méthode, où les sols sont plus richement colorés que les sols français. De plus, il est possible de faire évoluer l'application pour adapter le seuil afin de s'adapter à la teinte du sol, par une calibration en début de mesure par exemple. L'utilisation du SI600 permettrait aussi de simplifier la méthode d'acquisition. En effet, le SI600 ne nécessite qu'une mesure de la surface initiale des agrégats puis d'une mesure de la surface à 10 minutes, soit 2 photographies. Une application pour ordinateur pourrait aussi être développée, pour permettre l'utilisation d'un véritable appareil photographique. Sachant que, ces appareils permettent de sélectionner le format de sauvegarde RAW où l'image est brute. De plus, ces appareils permettent de sélectionner et fixer la sensibilité du capteur à la luminosité (ISO) ainsi que la vitesse de prise d'image (focale). De plus, le choix d'un objectif adapté permet de réduire les risques de distorsion ou aberrations chromatiques qui peuvent apparaître sur les téléphones portables. L'augmentation de la dimension de la profondeur de champs -zone où l'image est nette- peut aussi permettre d'augmenter la surface captée et donc d'utiliser de plus grands récipients. Il serait alors possible de mesure plus d'agrégats en même temps, et ainsi, réduire le temps nécessaire pour mesure de la totalité des agrégats. De plus, l'utilisation d'un anneau lumineux se fixant à l'objectif permettrait d'améliorer la gestion de la lumière pour un faible investissement. L'éclairage serait alors plus homogène, plus constant et l'appareil ne pourrait plus créer d'ombre portée. Cela permettrait aussi de minimiser les fluctuations de luminosité créées par les déplacements de l'expérimentateur.

Chaque mesure SLAKES dure 10 minutes brutes, et environ 12 minutes de manipulation (Mousset, 2021 ; Jabely, 2021). Réaliser les 40 mesures, représente un total de 400 à 480 minutes. L'utilisation du SI600, qui est un indicateur peu sensible aux différences entre téléphones, permet de paralléliser la mesure avec plusieurs téléphones. Il est donc possible de réduire le temps de manipulation. Ainsi, si l'on se place dans les conditions de l'INRAe d'Orléans avec 7 smartphones, on réduit le temps de manipulation à environ 1 h 10. En comparaison, pour réaliser l'ensemble de la mesure du MWD_{rapide} soit les tests avec un temps de repos (10 min), les tamisages et les pesées

il faut considérer environ 1 h 30 à 2 h auxquelles s'ajoutent 48 de séchage entre les tests et les pesées. Sachant qu'il est possible de paralléliser la phase de tests et repos sur plusieurs sites. Dans les conditions présentées, la méthode SLAKES présente un avantage temporel et donc en coup de main d'œuvre par rapport à la méthode Le Bissonnais. De plus, elle est plus simple à réaliser. En effet, elle nécessite très peu de manipulations, donc plus rapide à maîtriser. Mis à part l'acquisition des téléphones, elle ne nécessite pas de matériel spécifique. Contrairement à la méthode Le Bissonnais qui exige des tamis particuliers, de l'éthanol et potentiellement un appareil pour l'agitation mécanique.

L'utilisation du SI600 permet aussi d'éviter l'étape d'ajustement d'une équation de Gompertz sur la dynamique d'étalement des agrégats. Une étape où beaucoup d'agrégats peuvent être écartés du jeu final à cause des erreurs d'ajustement (Fajardo et al., 2016 ; Sakande, 2020 ; Peluchon, 2021 ; Mousset, 2021). Le rejet de certains agrégats sur la qualité d'ajustement de la Gompertz n'est plus nécessaire. On peut donc récupérer la totalité des agrégats mesurés. Cependant, il faut rester attentif aux agrégats non valides car ils possèdent un gravier en leur cœur. Ces agrégats ne se désagrègent pas et apparaissent très stables. Afin de vérifier cela, Peluchon (2021) proposa d'écraser d'un doigt tous les agrégats en fin d'expérience et de noter si l'on observe la présence d'un gravier. Ce qui permet de retirer cet agrégat *a posteriori*. L'évolution d'application pourrait aussi intégrer la possibilité de retirer des agrégats à la fin de la mesure. Selon la norme ISO de la méthode Le Bissonnais (Afnor, 2012), les sols dont la teneur en gravier est supérieure à 40 % sont considérés comme : non significatifs. Il n'est alors pas cohérent d'y réaliser une mesure SLAKES. Toujours selon la norme ISO une teneur en gravier dans la gamme de 10 à 40 % est une source d'erreurs importantes. Ces limites pourraient être affichées par l'application pour éviter une utilisation en condition non valable.

Un aspect plus complexe concerne les agrégats dont la désagrégation peut ne pas être terminée au bout de 10 minutes. Il faudrait alors explorer d'autres durées d'expériences pour s'assurer qu'il n'y a pas meilleur indicateur physique que le SI600. Dans leur étude, Flynn et al. (2020) ont comparé le SI600 au SI480 et au SI380. Le SI600 ressort comme le meilleur indicateur. On peut envisager d'explorer des temps supérieurs à 10 minutes.

D'autre part, il est aussi possible d'explorer l'ajustement d'autres courbes théoriques sur les données de dynamique d'étalement. Néanmoins, l'utilisation du SI600 semble légitime. De plus, son utilisation apporte une simplicité pour l'implémentation de la méthode en laboratoire permettant l'utilisation d'un appareil photographie de bonne qualité, car la mesure se résume en la prise d'une photographie au début de l'expérience puis à 10 mins.

C. Référentiel d'interprétation

Le référentiel d'interprétation mis en place pour le SI600 médian SLAKES s'appuie sur la transposition statistique sur celui définis pour la méthode Le Bissonnais. L'approche vise à avoir un taux de bonne répartition par rapport au référentiel d'interprétation du MWD rapide de Le Bissonnais, et plus particulièrement pour les classes « très instable » et « instable », qui sont très peu sous-estimées (Tableau XV). Néanmoins, cette classification surestime l'instabilité des sols en comparaison de la méthode Le Bissonnais. L'acquisition d'un plus grand nombre de données permettrait d'améliorer le référentiel. De plus, les mesures réalisées jusqu'alors ne correspondent pas au nombre de répétitions défini précédemment. Mesurer de nouveau ces sites en respectant ces

recommandations permettrait d'améliorer la qualité des données utilisées. De plus, le jeu de données possède un nombre de sites « stable » et « très sable » insuffisant, ce qui n'a pas permis leur bonne caractérisation. Pour le moment, le référentiel proposé surestime ces sites.

Le référentiel d'interprétation de Le Bissonnais s'appuie sur une évaluation du lien entre la mesure et le comportement réel sur le terrain (Le Bissonnais & Le Souder, 1995 ; Le Bissonnais et al., 1995 ; Amezketa et al., 1996). Une approche de terrain similaire à celle réalisée par Le Bissonnais et Arrouays (1997) est à envisager pour affiner, confirmer ou infirmer l'approche numérique proposée dans ce travail.

D. Formules de pédotransfert

Le SI600 est corrélé à des propriétés physico-chimiques du sol reconnues pour influencer la stabilité structurale (Fig. 17) (Yoder, 1936 ; Hénin 1938 ; Kemper & Koch 1966 ; Tisdall & Oades, 1982 ; Le Bissonnais, 1988 ; Le Bissonnais, 1996 ; Amézketa, 1999 ; Chenu et al., 2000). Il est positivement corrélé à la teneur en limons, qui est défavorable à la stabilité structurale. Inversement, il est négativement corrélé à des propriétés favorables à la stabilité structurale, le carbone organique, la biomasse microbienne (Chan & Heenan, 1999 ; Bossuyt et al., 2001 ; Cosentino et al., 2006), les argiles et les ions polyvalents (Mg, Fe, Al) qui participent à la floculation des argiles (Emerson, 1967 ; Le Bissonnais, 1996 ; Amézketa, 1999). Le MWD rapide présente le même profil avec des relations inversées. Cependant, certaines corrélations vont à l'encontre des connaissances. Le MWD rapide est négativement corrélé au Ca échangeable, un ion bivalent, et est positivement corrélés au K échangeable qui est un ion monovalent. De même pour le SI600, qui est positivement corrélé à la teneur en sable. Ces anomalies peuvent être le résultat de confusions d'effet ou d'associations fortuites car le jeu de données est trop réduit. D'autres variables physico-chimiques ayant une influence reconnue n'ont ici pas de corrélation significative ou importante avec les indicateurs, comme les carbonates.

Certaines associations fortuites ou confusions d'effets dans les modèles peuvent être liées à la taille limitée des jeux de données. Les modèles *cubists* sont le plus souvent utilisés pour des jeux de données de l'ordre de 1000 sites. Malgré tout, les statistiques de validation des modèles prédictifs *cubists* présentés se situent dans la normale des formules de pédotransfert publiées en pédologie (Kirkby et al., 2000 ; Flanagan, 2004 ; Bagarello et al., 2009 ; Fu et al., 2011 ; Wang et al. 2012 ; Lei et al., 2016). Le modèle *cubist* apporte la prise en compte de la non linéarité par rapport à un modèle linéaire. Les modèles linéaires ont été explorés, mais les résultats sont moins bons que pour les modèles présentés. L'augmentation du nombre de *committees* a permis de largement améliorer les modèles, entre la version descriptive et prédictive. Dans les deux cas, le MWD_{rapide} est mieux modélisé que le SI600.

E. Cartes de stabilité structurale

La distribution de la stabilité structurale est très différente entre les deux indicateurs de stabilité, même si certaines classes sont cohérentes, comme « très instable » et « moyennement stable ». Le nombre de sites classés « instables » par SI600 est plus important par rapport aux prédictions sur le MWD_{rapide}.

La formule de pédotransfert mise en place pour le SI600 semble illustrer la stabilité structurale inhérente au sol. En effet, seules des variables stables à plus ou moins long terme sont utilisées. Tandis que le MWD, en prenant en compte l'occupation du sol,

illustre plus facilement l'état actuel du sol. L'absence de l'utilisation l'occupation du sol par les modèles du SI600 peut-être liée au nombre de sites insuffisant. En effet, on a montré précédemment la capacité du SI600 à différencier les modes de travail du sol, ce qui avait été mis en avant par Flynn et al. (2020) et confirmé par Jones et al. (2021).

Les cartes des prédictions sont cohérentes entre les deux indicateurs : les zones de faible et forte stabilité sont situées aux mêmes endroits. Les nuances dans les zones intermédiaires sont totalement gommées pour le SI600. Le MWD rapide apporte de la nuance notamment grâce à l'utilisation de l'occupation du sol. L'Est de la France, le Massif Central et la Basse Normandie qui sont des régions d'élevage bovin avec des prairies, passent d'un SI600 « instable » à un MWD « stable » à « très stable ». Inversement, pour le Bassin Parisien, zone de grandes cultures où la stabilité structurale est réduite avec un plus grand nombre de sites « très instables ». De plus, la prédictibilité de certaines zones est à remettre en cause. En effet, pour les sols très riches en sables ou en graviers l'évaluation d'une stabilité structurale n'a peut-être pas de sens, comme dans les Landes. La mise en place de filtre sur la valeur pédologique de la prédiction serait à mettre en place.

V. Conclusion

Ce travail de stage avait pour objectif d'évaluer une nouvelle méthode *SLAKES* d'évaluation de la stabilité structurale à partir d'analyse d'image du suivi de la désagrégation d'agrégats plongés dans l'eau. Le travail a été effectué dans le contexte des sols de France métropolitaine. Tout d'abord, il s'agissait d'affiner le calcul de l'indicateur produit par cette nouvelle méthode et d'évaluer sa robustesse et précision sur un grand nombre de sols. Ensuite, d'évaluer l'indicateur retenu par rapport à la mesure MWD de la méthode Le Bissonnais afin de proposer un référentiel d'interprétation adapté au contexte français métropolitain. Ensuite, nous avons cherché à déterminer les propriétés du sol et les variables environnementales (occupation notamment) reliée statistiquement à la distribution des indicateurs *SLAKES* et Le Bissonnais dans un objectif de caractérisation de la distribution spatiale à l'échelle du territoire pour définir les zones à risque.

Tout d'abord, afin de faciliter la manipulation des mesures issues des mesures *SLAKES* sur smartphones, une chaîne de traitement a été développée et est diffusable sous la forme d'un package R : « *slakestable* ». Ce travail a permis également de proposer un nouvel indicateur *SLAKES* calculé à partir des analyses d'image : le SI600, plus précis et plus sensible à la texture et aux modes de travail du sol que l'indicateur proposé dans la publication initiale. L'analyse statistique d'un grand nombre de répétitions de la mesure en laboratoire a permis d'estimer que 30 à 70 répétitions sont nécessaires pour d'espérer estimer le SI600 selon un intervalle de confiance de 90 à 99%. L'indicateur *SLAKES* Asym, initialement proposé par Fajardo et al. (2016), ne présente pas la même précision et demande un nombre de répétitions de la mesure plus important pour être mis en œuvre (> 160).

L'analyse des corrélations entre les indicateurs et les propriétés physico-chimiques et descriptives ainsi que la modélisation par modèle *cubist* mettent en avant des liens attendus avec les propriétés pédologiques participant à la floculation des argiles. Le lien entre le mode d'occupation (culture, prairies, bois et forêts) n'a été pas pu être mis en avant uniquement pour l'indicateur Le Bissonnais. Le jeu de données étant réduit et contraint par des aspects opérationnels du RMQS, la répartition statistique n'est pas optimale pour toutes les variables considérées dans ce travail. De plus, le trop faible nombre de sites en « vergers et vignes » n'a pas permis d'y évaluer le comportement des indicateurs, alors que ce sont des parcelles connues pour leur sensibilité à l'érosion. Néanmoins, les modèles mis en place présentent de bonnes statistiques de validation, et sont de bonne qualité.

La comparaison entre les deux méthodes est complexe. Notamment, leurs protocoles de mesure comportent des différences importantes sur l'action mécanique appliquée aux agrégats après ré-humectation. La méthode Le Bissonnais impose ainsi une étape de tamisage des agrégats désagrégés qui est absente pour le protocole *SLAKES*. Nous avons considéré que l'agrégat réhumecté en fin d'expérience dans le protocole de la méthode *SLAKES* peut avoir atteint un équilibre limitant l'étalement des particules. Ce protocole semble donc proposer une action mécanique plus douce sur les agrégats.

De manière à améliorer la qualité des mesures réalisées sur les sites du RMQS, mon travail recommande d'augmenter le nombre de répétitions, ce qui permettrait d'atteindre une meilleure précision pour chaque site. De plus, le protocole de la mesure *SLAKES* pourrait être modifié de manière à améliorer la détection des agrégats

contenant un gravier. Il semble important de vérifier de leur présence en fin d'expérience avec le doigt. Enfin, il apparaît que l'algorithme d'analyse d'images par l'application pourrait être adapté pour mieux détecter les sols de couleurs claires.

Le méthode SLAKES pourrait être comparé aux autres tests Le Bissonnais comme l'humectation lente car mécaniquement plus douce.

Le référentiel d'interprétation proposé ne se base que sur un lien statistique en les indicateurs SLAKES et Le Bissonnais. Une approche terrain du référentiel à l'image de l'étude de Le Bissonnais et Arrouays (1997) permettrait de préciser, de valider ou d'invalider le référentiel proposé.

L'objectif initial d'une mesure de la stabilité structurale réalisable par tous s'éloigne. En effet, au vu des résultats, le nombre de mesures demandé est important et demande des précautions d'usage (lumière). Cependant, il est imaginable d'intégrer cette méthode dans un laboratoire de mesure, car une fois équipé la mesure est rapide et simple à mettre en place.

VI. Bibliographie

- AFNOR. (2012). ISO 10930:2012—Qualité du sol—Mesure de la stabilité d'agrégats de sols soumis à l'action de l'eau. <https://www.iso.org/fr/standard/46433.html>
- Algayer, B. (2012). Short term variability of soil aggregate stability and associated physico-chemical processes. [Life Science [Thèse], Université d'Orléans]. tel-00844812.
- Almajmaie, A., Hardie, M., Acuna, T., & Birch, C. (2017). Evaluation of methods for determining soil aggregate stability. *Soil and Tillage Research*, 167, 39-45. <https://doi.org/10.1016/j.still.2016.11.003>
- Amézketa, E. (1999). Soil Aggregate Stability: A Review. *Journal of Sustainable Agriculture*, 14(2-3), 83-151. https://doi.org/10.1300/J064v14n02_08
- Amézketa, E., Singer, M. J., & Le Bissonnais, Y. (1996). Testing a New Procedure for Measuring Water-Stable Aggregation. *Soil Science Society of America Journal*, 60(3), 888-894. <https://doi.org/10.2136/sssaj1996.03615995006000030030x>
- Annabi, M., Raclot, D., Bahri, H., Bailly, J. S., Gomez, C., & Le Bissonnais, Y. (2017). Spatial variability of soil aggregate stability at the scale of an agricultural region in Tunisia. *CATENA*, 153, 157-167. <https://doi.org/10.1016/j.catena.2017.02.010>
- Arrouays, D., Jolivet, C., Boulonne, L., Bodineau, G., Ratié, C., Saby, N., & Grolleau, E. (2003). Le Réseau de Mesures de la Qualité des Sols de France (RMQS). *Etude et Gestion des Sols*, 10(4), 241.
- Arrouays, D., Jolivet, C., Boulonne, L., Bodineau, G., Saby, N., & Grolleau, E. (2002). Une initiative nouvelle en France : La mise en place d'un réseau institutionnel de mesure de la qualité des sols (RMQS). *Comptes Rendus de l'Académie d'Agriculture de France*, 88(5), 93.
- Bagarello, V., Stefano, C. D., Ferro, V., Giuseppe, G., & Iovino, M. (2009). A PEDOTRANSFER FUNCTION FOR ESTIMATING THE SOIL ERODIBILITY FACTOR IN SICILY. *Journal of Agricultural Engineering*, 40(3), 7-13. <https://doi.org/10.4081/jae.2009.3.7>
- Baize, D. (2016). Petit lexique de pédologie : Nouvelle édition augmentée. Quae.
- Baize, D., & Girard, M. C. (2008). *Référentiel pédologique 2008*. Association française pour l'étude du sol (Afes). Quae. http://www.afes.fr/wp-content/uploads/2017/11/Referentiel_Pedologique_2008.pdf
- Baize, D., Girard, M. C., Arrouays, D., & Aubert, G. (1995). *Référentiel pédologique 1995*. Quae.
- Ben-Hur, M., & Lado, M. (2008). Effect of soil wetting conditions on seal formation, runoff, and soil loss in arid and semiarid soils—A review. *Soil Research*, 46(3), 191. <https://doi.org/10.1071/SR07168>
- Bossuyt, H., Deneff, K., Six, J., Frey, S. D., Merckx, R., & Paustian, K. (2001). Influence of microbial populations and residue quality on aggregate stability. *Applied Soil Ecology*, 16(3), 195-208. [https://doi.org/10.1016/S0929-1393\(00\)00116-5](https://doi.org/10.1016/S0929-1393(00)00116-5)
- Delaunois et Maucorps, 1980 - Présentation de la carte des sols de l'Aisne. Service de conseil agronomique - SUAD - Chambre d'Agriculture de l'Aisne. 12 p.

- Chan, K. Y., & Heenan, D. P. (1999). Microbial-induced soil aggregate stability under different crop rotations. *Biology and Fertility of Soils*, 30(1), 29-32. <https://doi.org/10.1007/s003740050583>
- Chenu, C., Bissonnais, Y. L., & Arrouays, D. (2000). Organic matter influence on clay wettability and soil aggregate stability. *Soil Science Society of America Journal*, 64(4), 1479.
- Clifford, D., Payne, J. E., Pringle, M. J., Searle, R., & Butler, N. (2014). Pragmatic soil survey design using flexible Latin hypercube sampling. *Computers & Geosciences*, 67, 62-68. <https://doi.org/10.1016/j.cageo.2014.03.005>
- Cosentino, D., Chenu, C., & Bissonnais, Y. L. (2006). Aggregate stability and microbial community dynamics under drying-wetting cycles in a silt loam soil. *Soil Biology and Biochemistry*, 68(8), 2053. <https://doi.org/10.1016/j.soilbio.2005.12.022>
- Edwards, A. P., & Bremner, J. M. (1967). Dispersion of soil particles by sonic vibration. *Journal of Soil Science*, 18(1), 47-63. <https://doi.org/10.1111/j.1365-2389.1967.tb01487.x>
- Emerson, W. (1967). A classification of soil aggregates based on their coherence in water. *Soil Research*, 5(1), 47. <https://doi.org/10.1071/SR9670047>
- Erktan, A., Legout, C., De Danieli, S., Daumergue, N., & Cécillon, L. (2016). Comparison of infrared spectroscopy and laser granulometry as alternative methods to estimate soil aggregate stability in Mediterranean badlands. *Geoderma*, 271, 225-233. <https://doi.org/10.1016/j.geoderma.2016.02.025>
- Fajardo, M., & McBratney, A. (2019). *Slakes : A soil aggregate stability smart-phone app* (2.1) [Mobile application software]. https://play.google.com/store/apps/details?id=slaker.sydneyuni.au.com.slaker&hl=en_AU
- Fajardo, M., McBratney, Alex. B., Field, D. J., & Minasny, B. (2016). Soil slaking assessment using image recognition. *Soil and Tillage Research*, 163, 119-129. <https://doi.org/10.1016/j.still.2016.05.018>
- FAO. (2019). Soil erosion : The greatest challenge to sustainable soil management. (p. 100). Licence: CC BY-NC-SA 3.0 IGO.
- Field, D. J., McKenzie, D. C., & Koppi, A. J. (1997). Development of an improved Vertisol stability test for SOILpak. *Soil Research*, 35(4), 843-852. <https://doi.org/10.1071/s96118>
- Flanagan, D. (2004). Pedotransfer functions for soil erosion models. In *Developments in Soil Science* (Vol. 30, p. 177-193). Elsevier. [https://doi.org/10.1016/S0166-2481\(04\)30011-5](https://doi.org/10.1016/S0166-2481(04)30011-5)
- Flynn, K. D., Bagnall, D. K., & Morgan, C. L. S. (2020). Evaluation of SLAKES, a smartphone application for quantifying aggregate stability, in high-clay soils. *Soil Science Society of America Journal*, 84(2), 345-353. <https://doi.org/10.1002/saj2.20012>
- Fu, S., Liu, B., Liu, H., & Xu, L. (2011). The effect of slope on interrill erosion at short slopes. *CATENA*, 84(1), 29-34. <https://doi.org/10.1016/j.catena.2010.08.013>

- Gis Sol. (s. d.). *Erosion des sols*. Gis Sol. Consulté 25 août 2021, à l'adresse <https://www.gissol.fr/thematiques/erosion-des-sols-48>
- Gomez, C., Le Bissonais, Y., Annabi, M., Bahri, H., & Raclot, D. (2013). Laboratory Vis-NIR spectroscopy as an alternative method for estimating the soil aggregate stability indexes of Mediterranean soils. *Geoderma*, 209-210, 86-97. <https://doi.org/10.1016/j.geoderma.2013.06.002>
- Hénin, S. (1938). Etude physico-chimique de la stabilité structurale des terres [Science des sols, Université de Paris]. tel-00752081.
- Hénin, S., Monnier, G., &ambeau, A. (1958). Méthode pour l'étude de la stabilité structurale des sols. *Annales Agronomiques*, 9, 73-92.
- Jabely, B. (2021). Etude de la variabilité de la stabilité structurale du sol au sein d'une parcelle instrumentée à l'aide de la technologie SLAKES [Rapport de stage L3]. Université de Tours.
- Jones, E. J., Filippi, P., Wittig, R., Fajardo, M., Pino, V., & McBratney, A. B. (2021). Mapping soil slaking index and assessing the impact of management in a mixed agricultural landscape. *SOIL*, 7(1), 33-46. <https://doi.org/10.5194/soil-7-33-2021>
- Kemper, W. D. (1965). Aggregate Stability. In C. A. Black (Éd.), *Agronomy Monographs* (p. 511-519). American Society of Agronomy, Soil Science Society of America. <https://doi.org/10.2134/agronmonogr9.1.c40>
- Kemper, W. D., & Koch, E. J. (1966). Aggregate stability of soils from Western United States and Canada : Measurement procedure, correlations with soil constituents: Vol. No. 1355 (Agricultural Research Service, US Department of Agriculture.).
- Kirkby, M. J., Le Bissonais, Y., Coulthard, T. J., Daroussin, J., & McMahon, M. D. (2000). The development of land quality indicators for soil degradation by water erosion. *Agriculture, Ecosystems & Environment*, 81(2), 125-135.
- Kuhn, M., & Quinlan, R. (2021). Cubist : Rule- and Instance-based regression modeling. R package version 0.3.0. <https://CRAN.R-project.org/package=Cubist>
- Lark, R. M., & Bishop, T. F. A. (2007). Cokriging particle size fractions of the soil. *European Journal of Soil Science*, 58(3), 763-774. <https://doi.org/10.1111/j.1365-2389.2006.00866.x>
- Landre, A., Saby, N. P. A., Barthès, B. G., Ratié, C., Guerin, A., Etayo, A., ... & Cornu, S. (2018). Prediction of total silicon concentrations in *French* soils using pedotransfer functions from mid-infrared spectrum and pedological attributes. *Geoderma*, 331, 70-80.
- Le Bissonais, Y. (1988). Analyse des mécanismes de désagrégation et de la mobilisation des particules de terre sous l'action des pluies [Sciences du Vivant [q-bio], Université d'Orléans]. tel-02857690.
- Le Bissonais, Y. (1996). Aggregate stability and assessment of soil crustability and erodibility : I. Theory and methodology. *European Journal of Soil Science*, 47(4), 425-437. <https://doi.org/10.1111/j.1365-2389.1996.tb01843.x>
- Le Bissonais, Y. (1996). Aggregate stability and assessment of crustability and erodibility : 1. Theory and methodology. *European Journal of Soil Science*, 47, 425-437.: *Reflections by Y. Le Bissonais*. *European Journal of Soil Science*, 67(1), 2-4. https://doi.org/10.1111/ejss.2_12311
- Le Bissonais, Y., & Arrouays, D. (1997). Aggregate stability and assessment of soil crustability and erodibility : II. Application to humic loamy soils with various organic

- carbon contents. *European Journal of Soil Science*, 48(1), 39-48.
<https://doi.org/10.1111/j.1365-2389.1997.tb00183.x>
- Le Bissonnais, Y., Blavet, D., De Noni, G., Laurent, J.-Y., Asseline, J., & Chenu, C. (2007). Erodibility of Mediterranean vineyard soils: Relevant aggregate stability methods and significant soil variables. *European Journal of Soil Science*, 58(1), 188-195. <https://doi.org/10.1111/j.1365-2389.2006.00823.x>
- Le Bissonnais, Y., & Le Souder, C. (1995). Mesurer la stabilité structurale des sols pour évaluer leur sensibilité à la battance et à l'érosion. *Etude et Gestion des sols*, 2(1), 43-56.
- Le Bissonnais, Y., Thorette, J., Bardet, C., & Daroussin, J. (2002). *L'érosion hydrique des sols en France*.
- Lei, W.-J., Tang, X.-Y., Reid, B. J., & Zhou, X.-Y. (2016). Spatial distribution of soil hydraulic parameters estimated by pedotransfer functions for the Jialing River Catchment, Southwestern China. *Journal of Mountain Science*, 13(1), 29-45.
<https://doi.org/10.1007/s11629-014-3390-9>
- Levoday, J., & Pyle, J. C. (1973). *The Emerson dispersion test and its relationship to hydraulic conductivity* (p. 7).
<https://publications.csiro.au/publications/publication/Plprocite:1af6079d-0bb1-4471-9948-f6f9a6eae5d1>
- Liu, J., Hu, F., Xu, C., Wang, Z., Ma, R., Zhao, S., & Liu, G. (2021). Comparison of different methods for assessing effects of soil interparticle forces on aggregate stability. *Geoderma*, 385, 114834. <https://doi.org/10.1016/j.geoderma.2020.114834>
- Louis, B. P., Saby, N. P. A., Orton, T. G., Lacarce, E., Boulonne, L., Jolivet, C., Ratié, C., & Arrouays, D. (2014). Statistical sampling design impact on predictive quality of harmonization functions between soil monitoring networks. *Geoderma*, 213, 133-143.
<https://doi.org/10.1016/j.geoderma.2013.07.018>
- Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, 32(9), 1378-1388. <https://doi.org/10.1016/j.cageo.2005.12.009>
- Mousset, P. (2021). Etude de la stabilité structurale des sols par l'application « Slakes » [Rapport de stage IUT]. IUT de la Roche-Sur-Yon.
- Paroissien, J.-B., Lagacherie, P., & Bissonnais, Y. L. (2010). A regional-scale study of multi-decennial erosion of vineyard fields using vine-stock unearthing–burying measurements. *CATENA*, 82(3), 159-168.
<https://doi.org/10.1016/j.catena.2010.06.002>
- Peluchon, M. (2021). Evaluation de la stabilité structurale du sol par une application mobile : SLAKES [Rapport de stage]. Agro Campus Ouest.
- Prosdocimi, M., Cerdà, A., & Tarolli, P. (2016). Soil water erosion on Mediterranean vineyards: A review. *Catena*, 141, 1-21.
- Quinlan (1992). Learning with continuous classes. Proceedings of the 5th Australian Joint Conference On Artificial Intelligence. pp. 343-348
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (4.0.4) [Computer software].

- Sakande, C. (2020). *Analyse statistique de données sur la stabilité structurale des sols obtenues à partir d'une application mobile* [Rapport de stage]. Université d'Orléans Master en mathématiques appliquées, statistiques & InfoSol Centre Val De Loire.
- Saygin, S. D., Erpul, G., & Basaran, M. (2017). Comparison of Aggregate Stability Measurement Methods for Clay-rich Soils in Asartepe Catchment of Turkey. *Land Degradation & Development*, 28(1), 199-206. <https://doi.org/10.1002/ldr.2383>
- Shi, P., Castaldi, F., van Wesemael, B., & Van Oost, K. (2020). Large-Scale, High-Resolution Mapping of Soil Aggregate Stability in Croplands Using APEX Hyperspectral Imagery. *Remote Sensing*, 12(4), 666. <https://doi.org/10.3390/rs12040666>
- Swiderski, C., Saby, N., Boulonne, L., Jolivet, C., & Cousin, I. (2017). Méthodologie d'élaboration du plan d'échantillonnage de la deuxième campagne du Réseau de Mesures de la Qualité des Sols de France par intégration multi-critères : Capacité à détecter une évolution temporelle, évaluation du réservoir en eau utilisable et contraintes logistiques. *Etude et Gestion des Sols*, 24(1), 83.
- Tahar, O. (2021). Evaluation et spatialisation de la stabilité structurale du sol mesurée par l'application mobile « SLAKES » à l'échelle du bassin versant de Naizin. [Rapport de Master 2]. Université de Tours.
- Tisdall, J. M., & Oades, J. M. (1982). Organic matter and water-stable aggregates in soils. *Journal of Soil Science*, 33(2), 141-163. <https://doi.org/10.1111/j.1365-2389.1982.tb01755.x>
- Truman, C. C., Bradford, J. M., & Ferris, J. E. (1990). Antecedent Water Content and Rainfall Energy Influence on Soil Aggregate Breakdown. *Soil Science Society of America Journal*, 54(5), 1385-1392. <https://doi.org/10.2136/sssaj1990.03615995005400050030x>
- UMR SAS Institut Agro Agrocampus Ouest. (2021). Géo SAS Plateforme EFELE. GéoSAS. <http://geowww.agrocampus-ouest.fr/hportails/?portail=efele>
- UMR SAS Institut Agro Agrocampus Ouest, UMR CNRS 6553, Ecobio. (2015). Sols de Bretagne. *Sols de Bretagne* <http://www.sols-de-bretagne.fr>
- Wang, Y., Shao, M., & Liu, Z. (2012). Pedotransfer Functions for Predicting Soil Hydraulic Properties of the Chinese Loess Plateau. *Soil Science*, 177(7), 424-432. <https://doi.org/10.1097/SS.0b013e318255a449>
- Wang, Y., & Witten, I. H. (1997). *Inducing Model Trees for Continuous Classes*. In Proc. of the 9th European Conf. on Machine Learning Poster Papers. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.9768>
- Yoder, R. E. (1936). A Direct Method of Aggregate Analysis of Soils and a Study of the Physical Nature of Erosion Losses ¹. *Agronomy Journal*, 28(5), 337-351. <https://doi.org/10.2134/agronj1936.00021962002800050001x>
- Young, R. A. (1984). A Method of Measuring Aggregate Stability Under Waterdrop Impact. *Transactions of the ASAE*, 27(5), 1351-1354. <https://doi.org/10.13031/2013.32970>

VII. Annexes

Annexe 1 : Le package est accessible sur lien suivant :

Chaloux T., Lacoste M., Saby N. (2021) "Slakestable: R package to explore raw data from the Slakes app", <https://doi.org/10.15454/BGSMUE>.

Annexe 2 : Dispersion de la mesure de stabilité SLAKES Asym et SI600 en fonction du nombre de mesures de 3 agrégats tirées

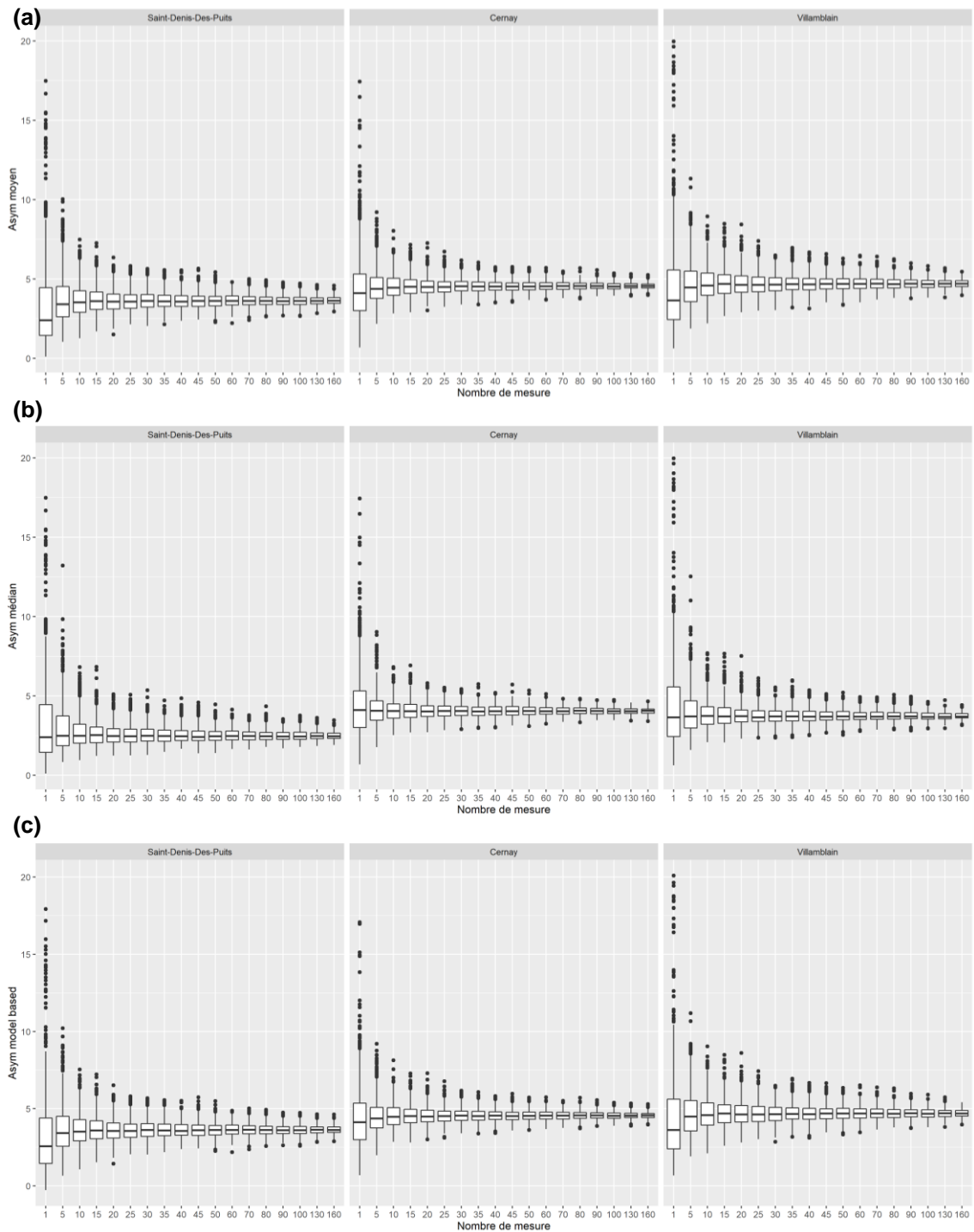


Figure 1. Evolution de la distribution de l'indicateur SLAKES Asym (a) moyen, (b) médian, (c) *model based* en fonction du nombre de répétitions pour chaque site : Saint-Denis-Des-Puits, Cernay et Villamblain

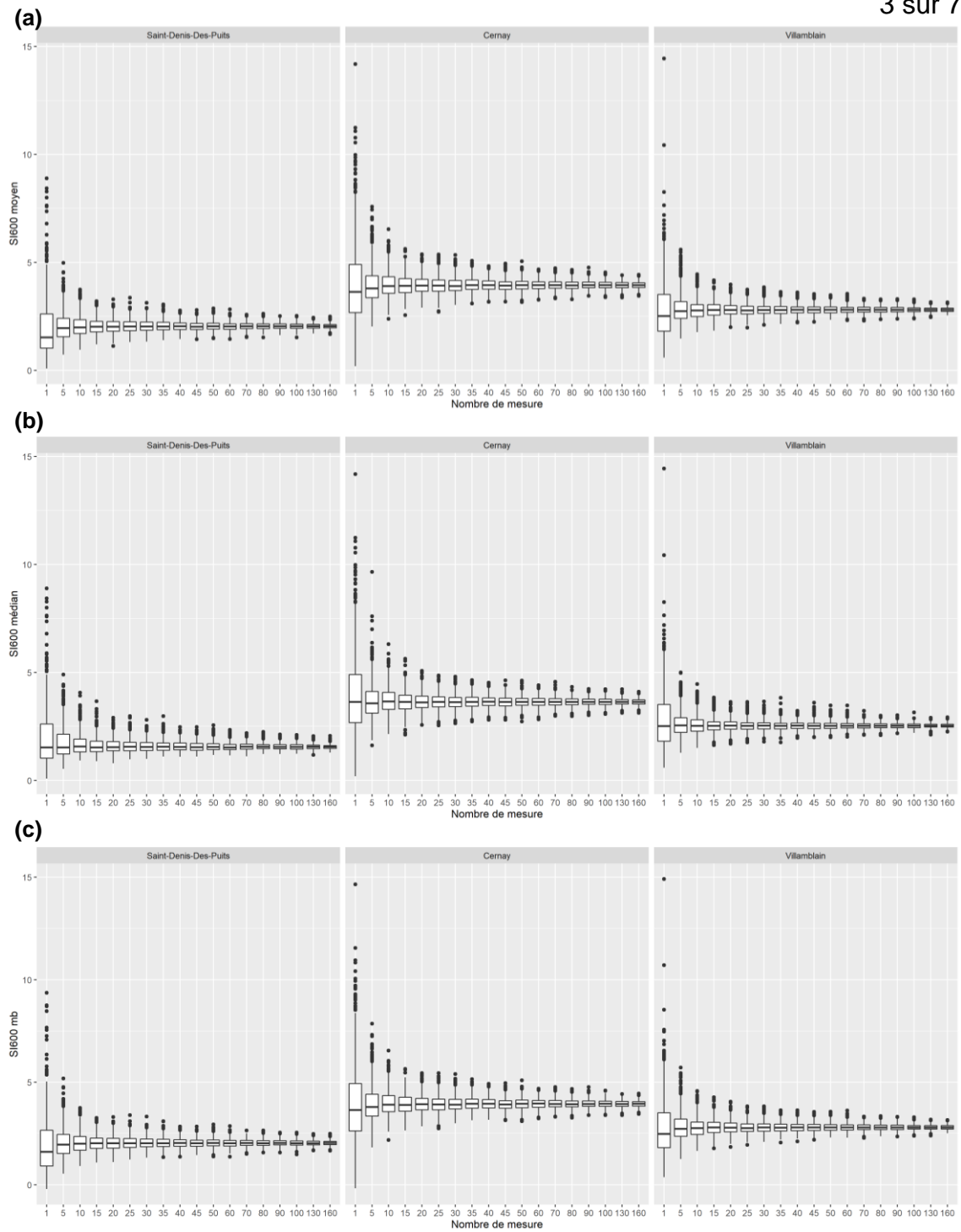


Figure 2. Evolution de la distribution de l'indicateur SLAKES SI600 (a) moyen, (b) médian, (c) *model based* en fonction du nombre de répétitions pour chaque site : Saint-Denis-Des-Puits, Cernay et Villamblain

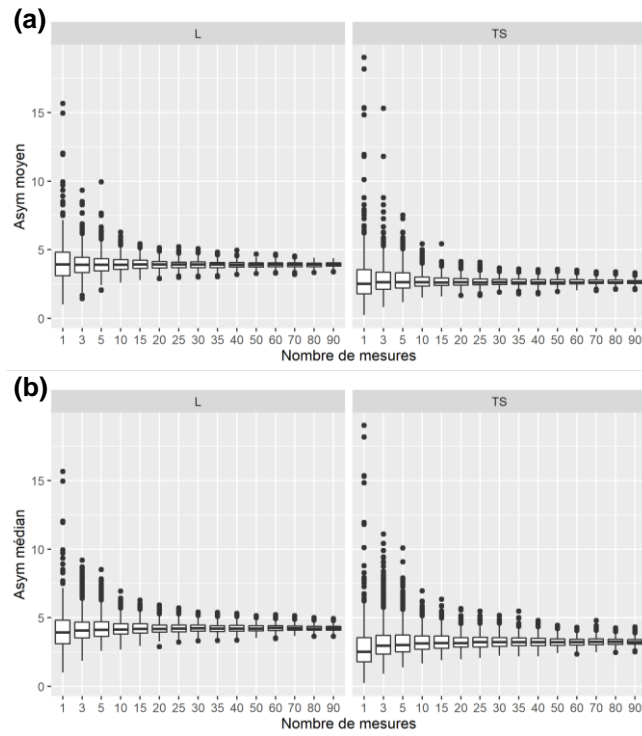


Figure 3. Evolution de la distribution de l'indicateur SLAKES Asym (a) moyen, (b) médian en fonction du nombre de répétitions pour chaque mode de travail du sol : labour (L) et travail simplifié (TS)

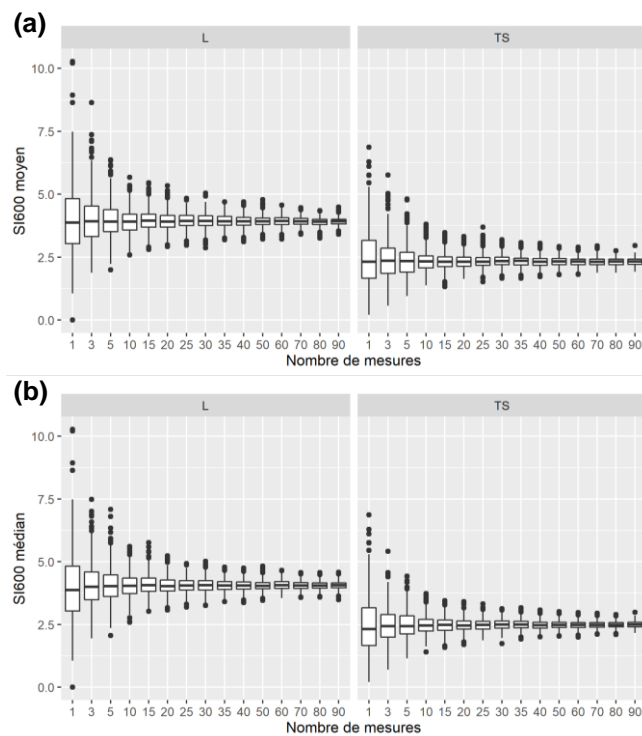


Figure 4. Evolution de la distribution de l'indicateur SLAKES SI600 (a) moyen, (b) médian en fonction du nombre de répétitions pour chaque mode de travail du sol : labour (L) et travail simplifié (TS)

Annexe 3 : Analyse en composante principale entre les indicateurs SLAKES et MWD_{rapide} Le Bissonnais et les propriétés pédologiques.

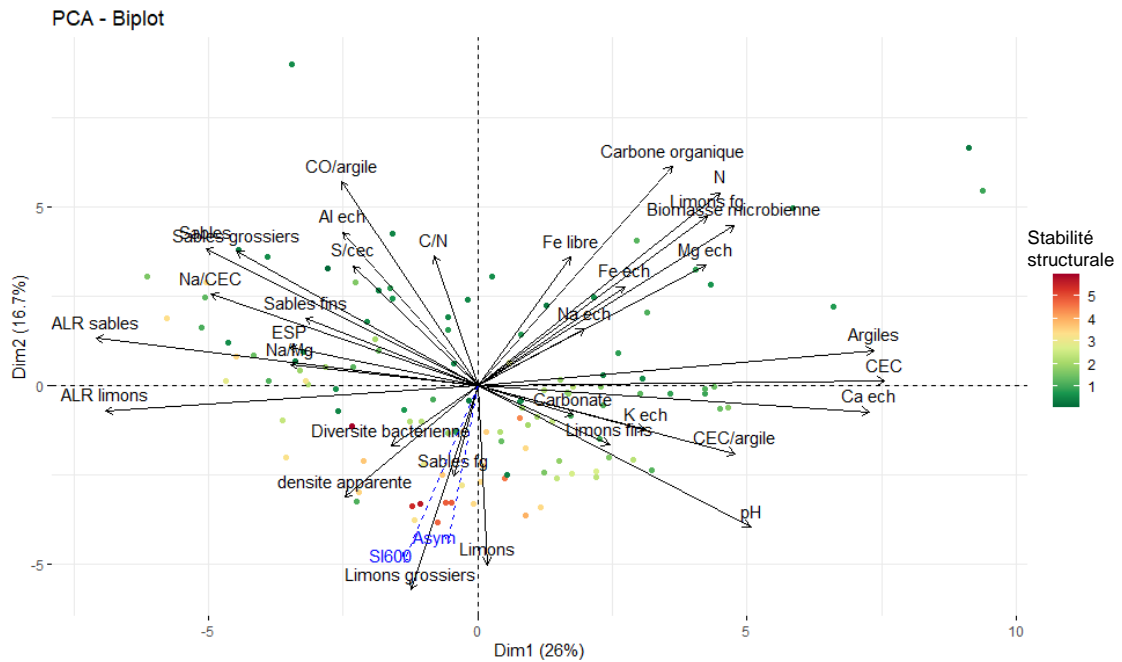


Figure 1. Bitplot de l'ACP entre les propriétés pédologiques et les indicateurs SLAKES Asym et SI600 (n = 117). ech = échangeable ; CO = carbone organique ; fg = ratio de texture fins/grossiers ; ALR = additive log ratio. Coloration des sites selon la stabilité structurale, plus la valeur est faible plus le sol est stable.

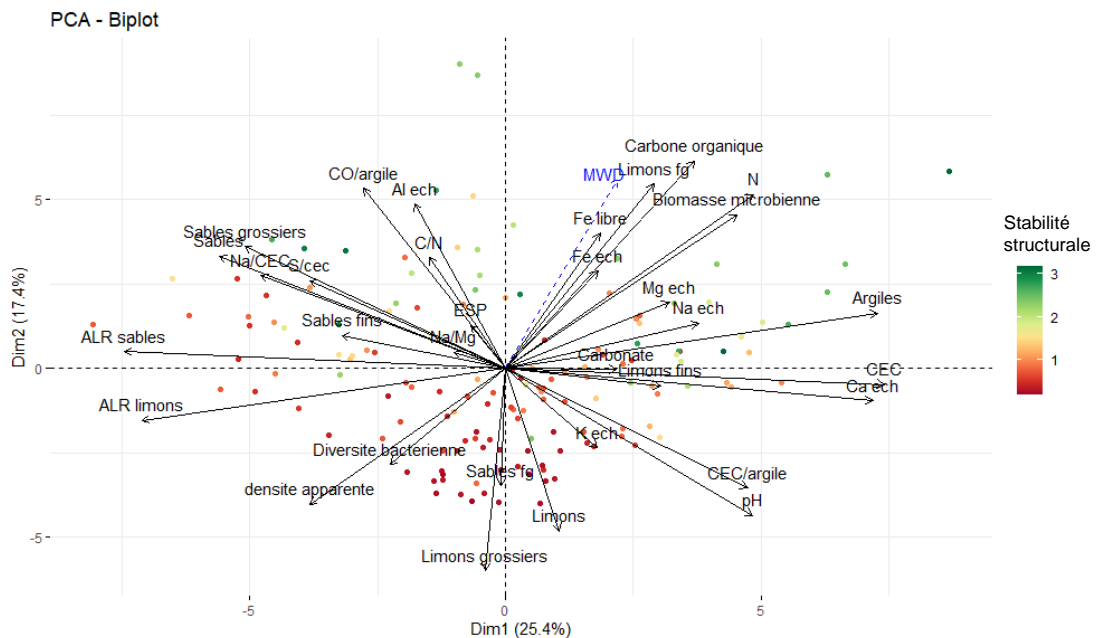


Figure 2. Bitplot de l'ACP entre les propriétés pédologiques et l'indicateur Le Bissonnais MWD_{rapide} (n = 151). ech = échangeable ; CO = carbone organique ; fg = ratio de texture fins/grossiers ; ALR = additive log ratio. Coloration des sites selon la stabilité structurale, plus la valeur est importante plus le sol est stable.

Evaluation de la méthode SLAKES en France métropolitaine

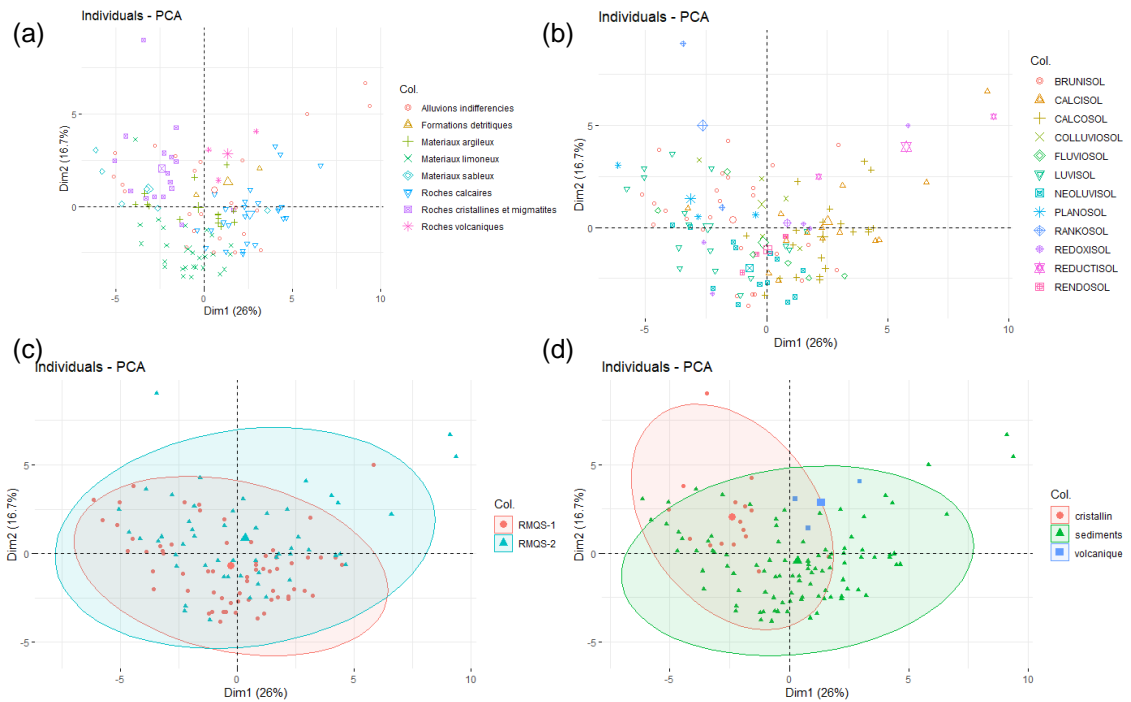


Figure 3. Représentation des individus en analyse en composante principale (ACP) selon (a) le matériau parental, (b) le nom de sol selon le référentiel pédologique (Baize et al., 1995 ; Baize & Girard, 2008), (c) la campagne du RMQS, (d) la nature du matériau parental pour les indicateurs SLAKES Asym et SI600.

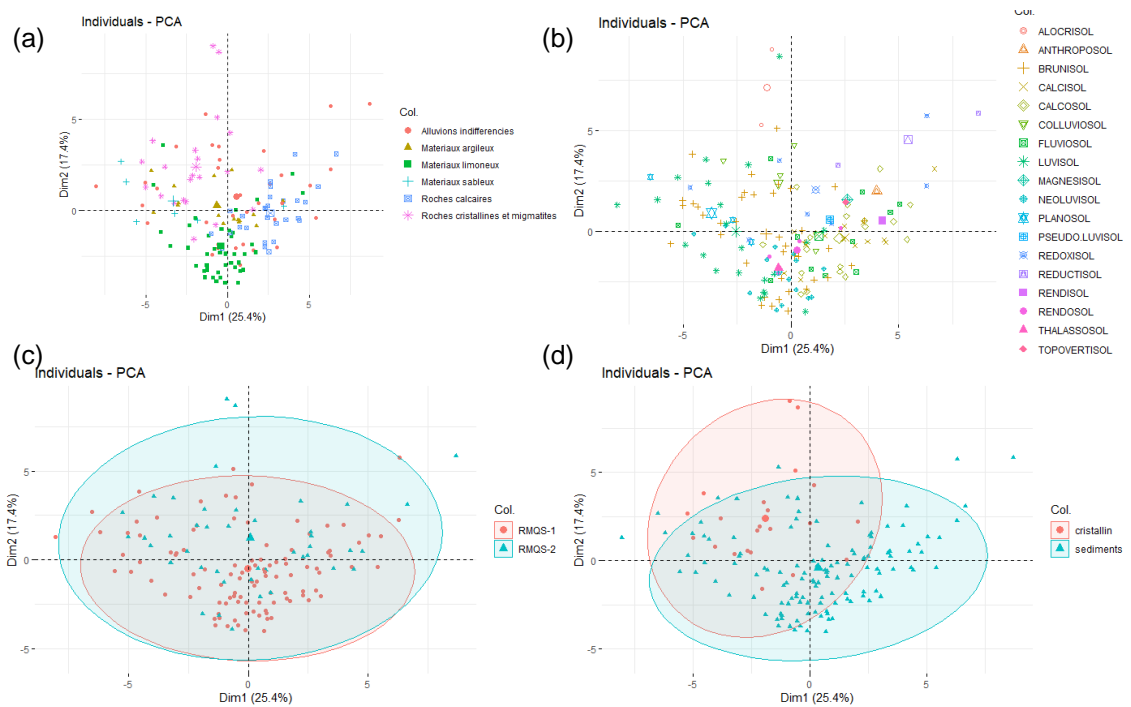


Figure 4. Représentation des individus en analyse en composante principale (ACP) selon (a) le matériau parental, (b) le nom de sol selon le référentiel pédologique (Baize et al., 1995 ; Baize & Girard, 2008), (c) la campagne du RMQS, (d) la nature du matériaux parental pour le MWD_{rapide} Le Bissonnais.

Annexe 4 : Variables utilisées par les modèles prédictifs.

Conditions	Modèle	Variables
33	47	Carbone organique
13	96	Limons
0	52	Mn extractible
0	39	CEC
0	33	Ca échangeable
0	30	Sables
0	28	CO/argile
0	25	K échangeable
0	21	Biomasse microbienne
0	17	N
0	16	S/CEC
0	12	Na échangeable
0	7	Carbonate
0	4	Mg échangeable
0	4	Na/Mg
0	4	pH
0	1	Ca total

Figure 1. Variables utilisées par le modèle *cubist* prédictif de l'indicateur SLAKES SI600 médian.

Conditions	Modèle	Variables
35	0	Occupation du sol
20	43	Fe échangeable
9	52	Carbone organique
8	36	K total
7	34	Mg total
7	0	Matériaux parental
5	27	S/CEC
5	30	Al échangeable
2	7	Na total
2	55	pH
0	57	Sables
0	28	Fe libre
0	69	N total
0	67	Limons
0	38	Densité apparente
0	35	CEC
0	32	C/N
0	31	Al total
0	29	Fe total
0	27	CEC/argile
0	24	CO/argile
0	24	Ca échangeable
0	22	Ca total
0	17	K échangeable
0	13	Biomasse microbienne
0	9	Mn extractible
0	9	Na échangeable
0	4	Mg échangeable
0	3	Na/CEC
0	2	Mn total
0	1	ESP

Figure 2. Variables utilisées par le modèle *cubist* prédictif de l'indicateur Le Bissonnais MWD_{rapide} .