Soil Security 11 (2023) 100088

Contents lists available at ScienceDirect

Soil Security

journal homepage: www.sciencedirect.com/journal/soil-security

Estimating soil aggregate stability with infrared spectroscopy and pedotransfer functions

Thomas Chalaux Clergue [a,b], Nicolas P.A. Saby [a,*], Alexandre M.J.-C. Wadoux [c], Bernard G. Barthès [d], Marine Lacoste [a,*]

[a] INRAE, Info&Sols, 45075, Orléans, France
[b] Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL), Unité Mixte de Recherche 8212 (CEA/CNRS/UVSQ), Orme des Merisiers, Gif-Sur-Yvette 91191, France
[c] Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia
[d] Eco&Sols, Université de Montpellier, IRD, CIRAD, INRAE, Institut Agro, 34060 Montpellier, France

ARTICLE INFO

Keywords:
Clay
Mid-infrared
Swelling
Slaking
Cubist
Multiple linear regression
Permutation importance

ABSTRACT

Soil aggregate stability is an important indicator of soil condition and is directly related to soil degradation processes such as erosion and crusting. Aggregate stability is conventionally measured by testing the aggregate resistance to water disturbance mechanisms. Such measurements, however, are costly and time-consuming, which make them difficult to implement at a regional or country scale. In this study, we explore two different approaches to estimate soil aggregate stability by means of commonly-measured soil properties or mid-infrared spectroscopy measurements. The first approach relies on land use and soil properties. In the second approach aggregate stability is estimated by a model fitted with mid-infrared spectroscopic data. We tested the two approaches with a dataset composed of 202 soil samples from mainland France, in which aggregate stability was measured with a fast wetting test. We found that simple linear models based on common soil properties and models based on mid-infrared spectral data yielded similar results. Interpretation of the models revealed well-known relationships: land use had a major role in predicting aggregate stability, followed by organic carbon and clay content. Overall, we conclude that both approaches offer a reliable, cheap and time-efficient alternative to estimating soil aggregate stability. These approaches offer a tool to estimate aggregate stability over large geographical areas, which can support the development of erosive risk management plans and the implementation of adaptive management strategies to mitigate threats to soil and improve the overall soil condition.

1. Introduction

Soil aggregate stability is the ability for the soil aggregates to maintain their structure under different stresses due to the action of water. Wet aggregate stability depends on soil aggregation, which is the process by which particles of different sizes join and are held together by organic and inorganic materials (Amezketa, 1999; Six et al., 2004). Aggregation processes depend on various factors, some are biological (quantity and quality of organic matters, organic stabilizing agents, biological activity) (Gupta and Germida, 2015; Lehmann et al., 2017), whereas others are inorganic (clay flocculation, ions, metalic oxide and hydroxides, carbonates and gypsum). Aggregation also depends on external factors such as climate, land use and soil management strategies. The latter indirectly impacts aggregate stability through modification of the aggregates size. Aggregate stability can also be viewed as the ability to maintain intra-aggregate porosity, compared to inter-aggregate porosity which is the predominant place of preferential water flow. In turn, aggregate stability influences water infiltration and storage, biological activity and crop growth, as well as soil sensitivity to erosion and crusting (Barthès and Roose, 2002; Le Bissonnais, 1996a). For example, a soil with a higher aggregation is expected to be less prone to erosion, crusting, deep compaction, to have smaller risk of pollution transfer, better water infiltration, and to be favourable for plant growth (Amezketa, 1999). As a consequence, soil aggregate stability is a conventional soil indicator commonly used to estimate water erosion risk, for example in European agro-ecosystems (Algayer et al., 2014; Barthès and Roose, 2002).

Aggregate stability is conventionally measured by testing the

* Corresponding authors.
E-mail addresses: nicolas.saby@inrae.fr (N.P.A. Saby), marine.lacoste@inrae.fr (M. Lacoste).

https://doi.org/10.1016/j.soisec.2023.100088
Received 15 May 2022; Received in revised form 12 February 2023; Accepted 26 February 2023
Available online 28 February 2023
2667-0062/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

aggregate resistance to one or several destabilizing mechanisms linked to the action of water, such as slaking, differential swelling of clays, mechanical breakdown by raindrop impact and clay physico-chemical dispersion (Algayer et al., 2014; Amezketa, 1999; Le Bissonnais, 1996a). Different methods can be used for this purpose, the most common of which is that of Le Bissonnais (1996a), currently recognized with an international ISO standard (ISO 10930:2012, 2012). In this method, the stability of aggregates is estimated using laboratory measurements, through application of three destabilizing mechanisms: fast wetting, slow wetting and mechanical breakdown. Other methods involve laboratory rainfall simulations (e.g. Le Bissonnais and Singer, 1993), clay dispersion (e.g. Le Bissonnais, 1996a), and ultrasonic vibrations (e.g. Emerson, 1967). These existing methods for measuring aggregate stability are usually labor-intensive, time-consuming and require specialized equipment (e.g. sieves, solvents, laboratory apparatus). For example, the three tests of fast and slow wetting and mechanical breakdown described in Le Bissonnais (1996a) need six different sieves, a tension table and a laboratory technician for about 2 hours to obtain a single measurement of aggregate stability. Clearly, the resources and costs required to perform the laboratory analysis are high, so measurement of aggregate stability are rarely part of routine soil surveys.

Several studies, however, have shown that soil aggregate stability has considerable spatial and temporal variation (Algayer et al., 2014). Aggregate stability is affected by land cover (Emadodin et al., 2009), varies spatially, and changes with time in response to soil management practices (Novelli et al., 2013; Paul et al., 2013; Six et al., 2000). Aggregate stability is also modified by soil organic carbon (SOC) content through alteration of clay wettability (Chenu et al., 2000). Lower SOC content usually leads to lower soil aggregate stability. Similarly, changes in land use impact aggregate stability and SOC, e.g. the stability of aggregates and SOC contents are usually different in woodlands and croplands (Guo et al., 2020). Also, within a land use, soil management can impact aggregate stability, like tillage which reduce it by breaking up aggregates (Paul et al., 2013). As an alternative to direct estimation of aggregate stability methods, the use of pedotransfert functions relying either on laboratory analysis (hereafter referred to as the lab-based approach) or on infrared spectroscopy have recently been developed (Afriyie et al., 2020; 2022; Annabi et al., 2017; Erktan et al., 2016b; Gomez et al., 2013; Shi et al., 2020). These approaches show that estimation and prediction of aggregate stability is possible using pedotransfert functions for small geographical areas, at the size of the parcel (Afriyie et al., 2020; 2022) to that of landscape (Erktan et al., 2016b) or that of the small region (Annabi et al., 2017; Gomez et al., 2013; Shi et al., 2020). Using pedotransfert functions allow one to obtain a sampling density that adequately captures the spatial variation of aggregate stability as they are less expensive than laboratory measurement of aggregate stability and thus enable more measurements to be acquired.

Pedotransfer function (PTF, Bouma, 1989) are mathematical functions that relate basic soil properties or mid-infrared spectral data (Landre et al., 2018), to other soil properties that are more difficult to estimate. Different mathematical models were used to develop PTF in soil science, such as multiple linear regression, decision trees, random forest or nearest neighbour (Van Looy et al., 2017). In the literature PTF have commonly been developed for soil hydraulic properties such as the field capacity or wilting point. Recent attempts to estimate aggregate stability with basic soil properties were made in Gomez et al. (2013) and Shi et al. (2020). Using a simple set of soil properties and a linear model, they showed that estimating aggregate stability was possible, and found moderate results in a case study within the Cap Bon region in Northern Tunisia and the central part of Belgium, respectively. As an alternative, the literature has reported promising results on the estimation of aggregate stability using infrared spectroscopic data. Infrared spectroscopy is routinely used as a cost-effective tool to estimate soil properties with a direct spectral response. Using the visible near-infrared (vis-NIR) and mid-infrared (MIR) ranges, a large number of studies have reported excellent results in the estimation of primary soil properties such as SOC, carbonate content, texture and soil water content (Soriano-Disla et al., 2014; Stenberg et al., 2010). Among these basic properties, SOC and clay are known to influence greatly aggregate stability (Amezketa, 1999; Le Bissonnais, 1996a). Several studies have thus proposed to estimate aggregate stability directly from vis-NIR (Afriyie et al., 2022; Gomez et al., 2013; Shi et al., 2020) and/or MIR (Afriyie et al., 2020; Cañasveras et al., 2010; Erktan et al., 2016b) spectra because spectroscopy is a relatively faster and cheaper method compared to the conventional laboratory methods outlined previously.

These studies reported promising but variable results with regards to the accuracy of the estimates, implying that infrared spectroscopy could not always replace conventional methods of analysis. This is because aggregate stability has no direct absorption in the infrared range of the electromagnetic spectrum, and prediction is based on correlation to one or more properties with a direct spectral response. Gomez et al. (2013), for example, found moderate accuracy in the estimation of aggregate stability with vis-NIR data, which they attributed to the very low influence that organic matter has on aggregate stability in their study region. Admittedly, estimating aggregate stability with infrared data depends critically on the quality of the surrogate relationship and on the data used to fit the spectroscopic model. There is a need to further explore the estimation of aggregate stability in large-scale environments showing a large range of variation in the primary soil properties.

The work presented here builds on the study of Gomez et al. (2013), who investigated the strategies for estimating aggregate stability using lab-based analyses and vis-NIR spectroscopy. In this study we explore two strategies to estimate aggregate stability by means of pedotransfer functions based on commonly-measured soil properties and mid-infrared spectroscopy, at the scale of a country with soils showing substantial variation in properties and land use types. The first approach relies on land use and commonly measured soil properties. The second approach estimate aggregate stability with mid-infrared spectroscopic data. The quality of the predictions of the two estimation strategies are compared and the best modelling strategies are interpreted and analysed.

2. Materials and methods

2.1. Case study and data

The dataset is composed of 202 soil samples from the French national soil monitoring network (*Réseau de Mesures de la Qualité des Sols*, RMQS) (Arrouays et al., 2003) (Fig. 1). The RMQS is based on a regular grid of 16 km × 16 km which leads to good spatial coverage, i.e. the sites are uniformly spread over the study area. This is profitable for PTF modelling, mapping and estimation of spatial means (Brus and Saby, 2016). The selection of the sites of this study was purposely done to built as much as possible a representative dataset of the RMQS, and therefore of the diversity of mainland French soils. The distribution of the three land use for the 202 sites is: 129 cropland sites, 52 grassland sites and 21 woodland sites.

For all 202 soil samples, we measured in the laboratory the aggregate stability and six other soil properties: texture (i.e. clay, silt and sand contents), organic carbon, pH_{water} and $CaCO_3$.

Aggregate stability (Chenu et al., 2011) was measured according to method of Le Bissonnais (1996a) (ISO 10930:2012, 2012). Among the three tests described in this protocol, we conducted the fast wetting (FW) test. The norm ISO 10930:2012 (2012) enables to only use this test, as it is faster than making the three tests together without substantial loss of accuracy in the estimation of aggregate stability. Hereafter, values of aggregate stability obtained by this test are referred to as MWD_{FW} (MWD stands for mean weight diameter).

In this FW test, undisturbed soil samples were spaded in the field over a thickness of 20 cm. Samples were air-dried and gently crumbled by hand, and then 5 mm sieved by hand. Aggregates of the 3 to 5 mm fraction are recovered and dried at 40 °C for 40 hours. For Le Bissonnais'

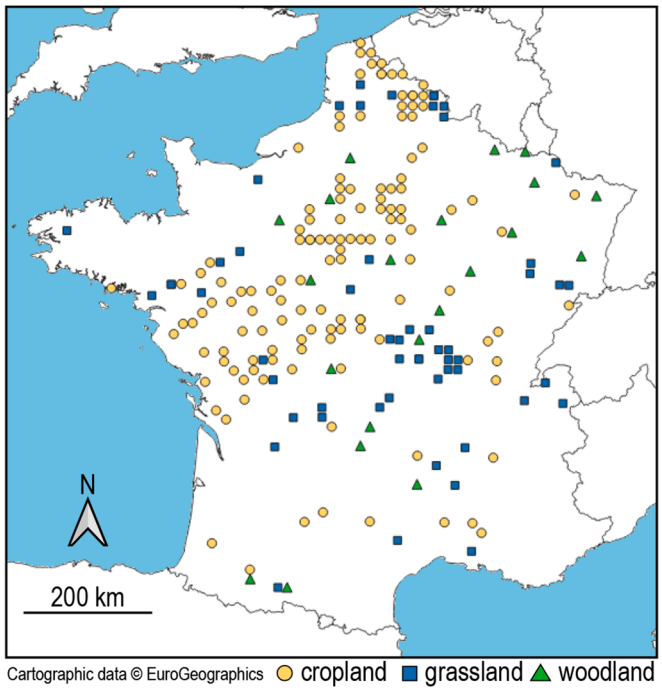


Fig. 1. Location of the soil samples used in this study with their land use: yellow circle = cropland, blue square = grassland and green triangle = woodland.

FW test, 5 g of aggregates are gently immersed in deionzed water for 10 minutes. The water is then sucked off with a pipette. The soil material is transferred to a 50 µm sieve previously immersed in ethanol, and then gently moved by hand. Both fractions above and below 50 µm are re-immersed in a small amount of ethanol. The fraction above 50 µm is sieved by hand on a column of six sieves: 2000, 1000, 500, 200, 100 and 50 µm (Le Bissonnais, 1996a). Each resulting fraction is oven-dried at 40 °C for 48 hours and then weighted. The fraction < 50 µm, which can not be weighed, is obtained indirectly by the difference between the initial mass and the sum of the 6 sieves. The final indicator, MWD_{FW}, is the sum of the mass fractions multiplied by the mean aggregates size for each sieve (i.e. the mean between two consecutive sieves). Measured values of MWD_{FW} have a direct relationship with aggregate stability and crustability (Table 1).

For texture (i.e. clay (<2 µm), silt (2-50 µm) and sand (50-2,000 µm) contents), organic carbon, pH_{water} and $CaCO_3$, 25 individual soil samples were taken from the topsoil (0–30 cm) at each of the 202 RMQS sites we selected, using an unaligned sampling design within a 20 m × 20 m square area. Texture was determined using the pipette method, sedimentation and sieving under water, without decarbonation according to ISO 11277:2009 (2009). The soil organic carbon (SOC) was determined with elemental analysis and dry combustion (ISO 10694:1995, 1995). Finally, pH_{water} was determined by 1/5 dilution (ISO 10390:2005, 2005) and $CaCO_3$ content was determined according to the volumetric method (ISO 10693:1995, 2014). In order to compare the variables importance for the prediction of the aggregate stability, these elementary soil properties were standardized, with their mean reduced to zero and their standard deviation to one.

Table 1
Classes of stability and crustability according to MWD values measured with the three treatments of Le Bissonnais (1996a) (ISO 10930:2012, 2012).

Class	MWD value	Stability	Crustability
1	< 0.4	Very unstable	Systematic crust formation
2	0.4-0.8	Unstable	Crusting frequent
3	0.8-1.3	Medium	Crusting moderate
4	1.3-2.0	Stable	Crusting rare
5	> 2.0	Very stable	No crusting

Mid-infrared spectra were measured on soil samples during the study of Grinand et al. (2012). The samples were grounded to < 0.2 mm and aliquots of about 0.5 g were scanned 32 times from 4000 to 400 cm^{-1} (2500-25000 nm) at 3.86 cm^{-1} resolution with a Nicolet 6700 Diffusive Reflectance Fourier Transform Spectrophotometer (Thermo Fisher Scientific Instruments, Madison, WI, USA). The 32 MIR spectra obtained for an individual sample were then averaged and the process repeated to have one spectrum for each of the 202 samples. Spectra were recorded as reflectance and then transformed by a log-transform of the inverse to absorbance, which was used for modelling. For this work, only the wavenumbers between 4000 to 500 cm^{-1} were analyzed, since bands between 400 and 500 cm^{-1} had a low signal to noise ratio. The spectra were pre-treated for noise reduction with a Savitzky Golay filter using a window size of 11 and the R package prospectr (Stevens and Ramirez-Lopez, 2021).

2.2. Prediction of soil aggregate stability

Two approaches have been studied in order to predict the MWD_{FW} (Fig. 2). In the first approach, we use commonly-measured soil variables (the lab-based approach, LB). In the second approach, we use a spectroscopic approach using MIR spectra. Processing and modelling were carried out under the R programming environment (R Core Team, 2022) following the workflow described in Wadoux et al. (2021).

2.2.1. Lab-based modelling

Two types of statistical models were used to build LB pedotransfer functions to model aggregate stability based on commonly-measured soil properties and land use: multiple linear regression (MLR noted LB_{MLR}) and cubist (LB_{cubist}).

In multiple linear regression, a least-square linear model is fitted between a response variable (i.e. the MWD_{FW}) and the set of explanatory variables. Models were built using all the explanatory variables (i.e. land use, organic carbon, texture, pH_{water} and $CaCO_3$). The models were built using the standard lm function from the R package stats (R Core Team, 2022). When a model was selected, the significance of each explanatory variable contribution was assessed with the associated T-test *p*-value obtained with the lm function. The T-test indicates whether or not the predictor is meaningful for the model. The α threshold was set to 0.10, which means that *p*-value must be bellow α.

Cubist (Quinlan et al., 1992) is a form of piecewise linear regression model that relies on decision trees. The model is fitted in two steps. In the first step, the explanatory variables are partitioned rowwise into smaller homogeneous datasets using the standard error of the

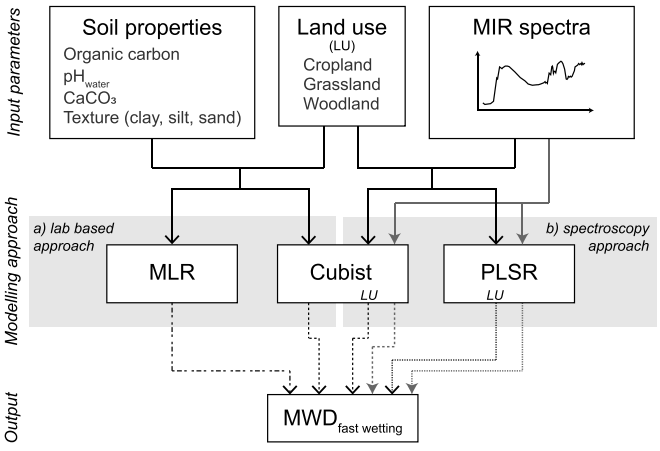


Fig. 2. Modelling approaches followed in this study for the prediction of the aggregate stability.

partitioned dataset as splitting criterion. The splitting within the partitioned dataset continues until a stopping criterion is met: either a minimum number of explanatory variables or if there is no further reduction is the error. The result is a decision tree whose final nodes are subsets of samples where the explanatory variables which have similar characteristics. In the second step, prediction is made by fitting a multiple linear least-square regression model on the partition of the final node between the aggregate stability and the explanatory variables of the partition. The node MLR models are fitted on a subset of explanatory variables. The final cubist model is a set of rules based on the explanatory variables, which are then applied for prediction of the aggregate stability. The model was built using the cubist function from the R package Cubist (Kuhn and Quinlan, 2021) using a optimized number of committees. The number of committees is the number of cubist models that are fitted and then combined for prediction in an ensemble of cubist models. It improves stability of the models and reduces the variance of the prediction. We explored a range of committees between 0 (i.e. a single cubist model) to 30. The optimal number of committees was determined by inspection of the RMSE values.

The two models were fitted to explain the aggregate stability index, MWD_{FW}, using elementary soil properties (i.e. OC, pH_{water}, $CaCO_3$ and texture) and the land use as explanatory variables. A predictor set was composed of land use, pH_{water}, $CaCO_3$ and one set of texture. In order to avoid multicollinearity between texture fractions, leading to the situation where two or more explanatory variables are strongly correlated, we created six disjoints texture sets for which distinct models were created. The subsets are: clay, silt, sand, clay & silt, clay & sand, and silt & sand. Before building a model, the linear correlation between properties were studied with the Pearson's *r* correlation coefficient. To avoid multicollinearity among explanatory variables, we calculate the variance inflation factor (VIF). The VIF is the inverse of the difference between 1 and the unadjusted coefficient of determination for regressing the ith explanatory variable on the remaining ones. If the VIF value is between 5 and 10, multicollinearity will inflate prediction accuracy. Therefore only variables with a VIF values lower than 5 were be kept (Akinwande et al., 2015). If a variable was removed, it was removed for both LB models (i.e. MLR and cubist).

Considering the usually costly laboratory analysis of $CaCO_3$ and by acknowledging that prediction formula should be made as simple as possible, we considered two supplementary approaches. The properties $CaCO_3$ and pH_{water} provide an equivalent information on acidic soil conditions, it may not be necessary to use both them simultaneously. We fitted LB_{MLR} and LB_{cubist} models using three sets: one using all variables, one excluding pH_{water} and one excluding $CaCO_3$. Hereafter, the results were shown for each of the three types of set and the two models.

2.2.2. Spectroscopic modelling

Two types of statistical models were used to relate the MIR spectra to the aggregate stability MWD_{FW}: partial least squares regression (MIR_{PLSR}) and cubist (MIR_{cubist}). The theory for cubist was described in Section 2.2.1.

Partial least squares regression (PLSR) is an alternative to MLR model when the data are highly collinear. In PLSR, a set of orthogonal factors (called PLS components) are obtained by a projection to a new space of the dependent (i.e. MWD_{FW}) and independent (i.e. the spectral data) variables. The orthogonal factors are determined to maximize the covariance between the factor and the response variable. A least square is used to fit a linear regression model using PLS components as predictors. Prior to modelling all explanatory and response variables are standardized to zero mean and unit variance. The optimal number of PLS components is determined by a *K*-fold cross-validation strategy, using the root mean square error as criterion. The PLSR model was implemented using the plsr function from the R package pls (Liland et al., 2021).

Cubist and PLSR for spectroscopic modelling were fitted with either MIR (MIR_{cubist} and MIR_{PLSR}) or MIR and land use (hereafter LU) ($MIR_{cubist+LU}$ and $MIR_{PLSR+LU}$). Land use was added to the spectroscopic modelling because it is an easy-to-access information in field surveys. The number of components for the PLSR model was selected by visual inspection of the RMSE values, by testing between 1 to 20 components, whereas the number of committees of the cubist model was tested for values between 1 to 30. We selected a number of components and committees as small as possible, and for which larger number would show only a marginal improvement in the RMSE value.

2.2.3. Accuracy assessment

For every model, the prediction performance was assessed based on the results of a *K*-fold cross-validation. The dataset was randomly split into $K = 10$ approximately same size folds. At each step, the soil samples of one fold were kept apart for the validation and the remaining soil samples ($K - 1$ folds) were used as calibration dataset. The results of the *K*-fold predictions were collected all together and used to calculate the prediction performance statistics. Folds were defined previously to the models construction to ensure that each models use the same folds to evaluate the prediction accuracy of LB and spectroscopic models and make the statistics comparable.

Models were compared using the mean error (ME), the root mean square error (RMSE), the squared Pearson's *r* correlation coefficient (r^2), and the modelling efficiency coefficient (MEC), defined with the following equations:

Mean error:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n} z_i - \widehat{z}_i, \tag{1}$$

where z_i and $\widehat{z}_i$ denote the measured and predicted values of soil aggregate stability for site *i*, respectively, and *n* is the total number of measured values.

Root mean square error:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \widehat{z}_i)^2}. \tag{2}$$

Squared Pearson's *r* correlation coefficient:

$$r^2 = \left(\frac{\sum_{i=1}^{n}(z_i - \overline{\mathbf{z}})(\widehat{z}_i - \overline{\widehat{\mathbf{z}}})}{\sqrt{\sum_{i=1}^{n}(z_i - \overline{\mathbf{z}})^2}\sqrt{\sum_{i=1}^{n}(\widehat{z}_i - \overline{\widehat{\mathbf{z}}})^2}}\right)^2, \tag{3}$$

where $\overline{\mathbf{z}}$ is the mean of the measured values and $\overline{\widehat{\mathbf{z}}}$ is the mean of the predicted values.

Modelling efficiency coefficient:

$$\mathrm{MEC} = 1 - \frac{\sum_{i=1}^{n}(z_i - \widehat{z}_i)^2}{\sum_{i=1}^{n}(z_i - \overline{\mathbf{z}})^2}. \tag{4}$$

The ME represents the bias and the RMSE the magnitude of the error. Both have an optimal value of 0, but the ME can be negative. The r^2 describes the linear correlation between measured and predicted values and ranges between 0 (no linear correlation) to 1 (perfect linear correlation). The MEC optimal value is 1 but it can be negative if the mean of the measured values is a better predictor than the model. Positive MEC values can be interpreted as an amount of variance explained by the model.

2.2.4. Model interpretation

For each type of models, LB or MIR, only the best model is kept. The selection of this model is based on the comparison of the prediction performance among models. For the LB approach the choice is between MLR and cubist, and for the MIR approach it is between cubist and PLSR.

To study which variables (i.e. soil properties or wavenumbers) are useful for the models, a feature importance by permutation approach was followed. The feature importance measure how important is a

feature to the model prediction quality. To measure the importance, the model calculates the increase of the model prediction error after permuting a variable while keeping all other variables fixed. If the model prediction error rises the variable is considered as important. The variable is not important when the prediction error is not affected. The feature importance measurement was done with the package iml (Molnar et al., 2018). The feature importance was measured by the difference of RMSE with 500 permutations.

In order to simplify the reading of the wavenumber importance the wavenumber resolution was downscaled from 3.6 cm^{-1} to 11 cm^{-1} with a moving window, using the package prospectr.

3. Results

3.1. Soil characteristics

Descriptive statistics of the conventional laboratory soil analyses and MWD_{FW} measured in laboratory are summarized in Table 2. The physico-chemical properties of the soils collected in this study showed great variations. In particular, organic carbon ranged from 5.37 to 79.70 $g{\cdot}kg^{-1}$ with an average value of about 21 $g{\cdot}kg^{-1}$. This is similar to the values reported for the whole RMQS dataset in France (Institut National De La Recherche Agronomique, 2021, see in the Supplementary Materials, Table 1). However, our dataset showed a lower organic matter content than RMQS (median is 17.00<19.60 $g{\cdot}kg^{-1}$, mean: 20.79<25.59 $g{\cdot}kg^{-1}$) (Table 2 and Table 1 in the Supplementary Materials). This can be explained by the over-representation of cropland in our dataset, as cropland have usually a lower organic matter content than other land uses, i.e. grassland and woodland, median are 14.63, 25.60 and 26.80 $g{\cdot}kg^{-1}$, respectively (Table 1 and 2 in the Supplementary Materials). About half of samples in our dataset is poor in $CaCO_3$ (median is 0 $g{\cdot}kg^{-1}$) while the remaining samples show a large range of calcareous content from 0 to 866 $g{\cdot}kg^{-1}$. The texture is variable and a wide variety of texture classes are represented. A wide range of pH_{water} is explored, from extremely acidic (4.0) to moderately alkaline (8.3), with an average neutral values of pH_{water} (i.e. 6.7).

MWD_{FW} values vary greatly over the studied area (Table 2, Fig. 3). MWD_{FW} ranged from 0.19 to 3.32, with a mean value of 0.88. The five stability classes defined by Le Bissonnais were represented (Le Bissonnais, 1996b). Most samples were from soil that were categorized under *unstable* to *very unstable* soil aggregate stability classes (46%). There was also large disparity in MWD_{FW} values among land-uses. Cropland soils had on average lower MWD_{FW} values, which correspond to *unstable* and *very unstable* classes. Grassland soils had large variation in MWD_{FW} values and range from *unstable* to *very stable*. Finally, soil samples from woodlands had overall high MWD_{FW} values and in consequence high stability (categories *stable* and *very stable*).

Table 3 shows that the soil aggregate stability index MWD_{FW} was strongly positively correlated with soil organic carbon ($r = 0.67$), to a lesser extent with clay ($r = 0.34$) but weakly and negatively correlated with silt (r = -0.30) and pH_{water} (r = -0.28). Soil properties were not strongly correlated one to another, with the exception of SOC and clay ($r = 0.44$), pH_{water} and clay ($r = 0.38$) and pH_{water} and $CaCO_3$ ($r = 0.40$) which were moderately correlated.

Table 2
Summary statistics of measured values of soil properties and MWD_{FW} for the 202 soil samples. sd: standard deviation.

Property	Minimum	Maximum	Median	Mean	sd
MWD_{FW} (mm)	0.19	3.32	0.88	1.16	0.83
Organic carbon ($g{\cdot}kg^{-1}$)	5.37	79.70	17.00	20.79	12.46
pH_{water}	4.0	8.3	6.7	6.7	1.2
$CaCO_3$ ($g{\cdot}kg^{-1}$)	0.0	866.0	0.0	58.0	171.1
Clay ($g{\cdot}kg^{-1}$)	81	648	236	260	119
Silt ($g{\cdot}kg^{-1}$)	101	789	452	463	172
Sand ($g{\cdot}kg^{-1}$)	13	816	213	275	197

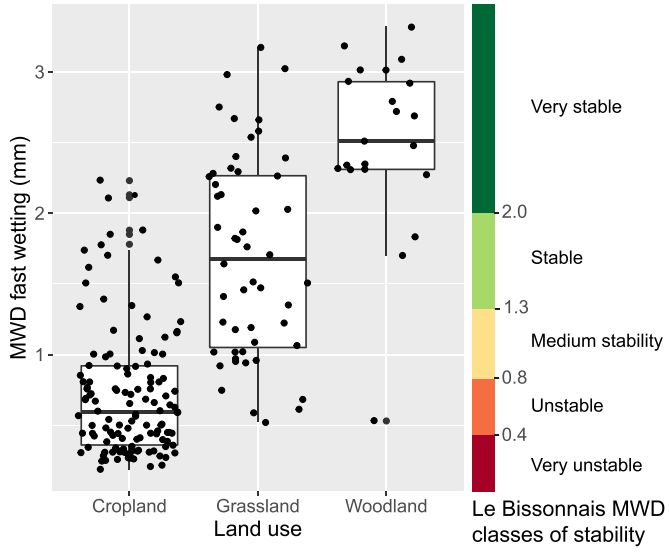


Fig. 3. Distribution of the measured values of MWD_{FW} (mm) by land use type (center lines of each box are median values, while upper and lower lines are the first and third quartiles).

Table 3
Person's *r* correlation coefficient between the measured soil properties and the MWD_{FW}.

	MWD_{FW}	SOC	$CaCO_3$	pH_{water}	Clay	Silt
SOC	0.67					
$CaCO_3$	0.16	0.13				
pH_{water}	-0.30	-0.16	0.40			
Clay	0.34	0.44	0.17	0.38		
Silt	-0.28	-0.17	-0.06	0.11	-0.12	
Sand	0.04	-0.12	-0.05	-0.32	-0.50	-0.80

Table 4
Summary 10 folds cross-validation statistics of multiple linear regression (MLR) models using basics soil properties (i.e. organic carbon, pH_{water} and/or $CaCO_3$) and different texture sets (i.e. clay, silt, sand, clay & silt, clay & sand or silt & sand) to predict MWD_{FW}. ME = mean error; RMSE = root mean square error; MEC = modelling efficiency coefficient.

Properties	Texture	ME (mm)	RMSE (mm)	r^2	MEC
Land use, SOC, pH_{water}, $CaCO_3$	Clay	0.00	0.48	0.66	0.66
	Silt	0.00	0.51	0.63	0.63
	Sand	0.00	0.52	0.61	0.61
	Clay & silt	0.00	0.48	0.66	0.66
	Clay & sand	0.00	0.48	0.66	0.66
	Silt & sand	0.00	0.48	0.66	0.66
Land use, SOC, pH_{water}	Clay	0.00	0.49	0.65	0.65
	Silt	0.00	0.51	0.62	0.62
	Sand	0.00	0.52	0.61	0.61
	Clay & silt	0.00	0.49	0.65	0.65
	Clay & sand	0.00	0.49	0.65	0.65
	Silt & sand	0.00	0.49	0.65	0.65
Land use, SOC, $CaCO_3$	Clay	0.00	0.49	0.64	0.64
	Silt	0.00	0.51	0.63	0.63
	Sand	0.00	0.52	0.61	0.61
	Clay & silt	0.00	0.49	0.64	0.64
	Clay & sand	0.00	0.49	0.64	0.64
	Silt & sand	0.00	0.49	0.64	0.64

Table 5
Summary 10 folds cross-validation statistics of cubist models using basics soil properties (i.e. organic carbon, texture, pH_{water} and/or $CaCO_3$) and different texture sets (i.e. clay, silt, sand, clay & silt, clay & sand or silt & sand) to predict MWD_{FW}. ME = mean error; RMSE = root mean square error; MEC = modelling efficiency coefficient.

Properties	Textures	ME (mm)	RMSE (mm)	r^2	MEC
Land use, SOC, pH_{water}, $CaCO_3$	Clay	-0.05	0.51	0.62	0.62
	Silt	-0.03	0.53	0.60	0.59
	Sand	-0.04	0.54	0.58	0.57
	Clay & silt	-0.04	0.51	0.62	0.62
	Clay & sand	-0.05	0.50	0.64	0.64
	Silt & sand	-0.04	0.50	0.64	0.64
Land use, SOC, pH_{water}	Clay	-0.05	0.50	0.65	0.64
	Silt	-0.05	0.51	0.63	0.62
	Sand	-0.06	0.51	0.63	0.62
	Clay & silt	-0.06	0.49	0.65	0.64
	Clay & sand	-0.06	0.50	0.65	0.64
	Silt & sand	-0.06	0.50	0.65	0.64
Land use, SOC, $CaCO_3$	Clay	-0.07	0.53	0.60	0.60
	Silt	-0.07	0.51	0.62	0.62
	Sand	-0.07	0.55	0.58	0.57
	Clay & silt	-0.07	0.51	0.63	0.62
	Clay & sand	-0.06	0.53	0.59	0.58
	Silt & sand	-0.08	0.51	0.62	0.61

Table 6
Standardized regression coefficients for multiple linear regression models fitted with land use, SOC, clay & silt, pH_{water} and $CaCO_3$. Significance of correlation: "****" p-value < 0.001; "***" p-value < 0.01; "**" p-value < 0.05; "." p-value < 0.1; " " p-value < 1.

Property	Value	T-test p-value	T-test Significance
Intercept 'cropland'	0.8340	<2e-16	***
Land use: woodland	1.3050	<2e-16	***
Land use: grassland	0.6906	4.07e-6	***
Clay	0.1971	1.03e-5	***
Soil organic carbon	0.1804	1.99e-4	***
pH_{water}	-0.0841	5.08e-2	.
Silt	-0.0630	6.99e-2	.

Table 7
Summary 10 folds cross-validation statistics of partial least squares regression and cubist models using only either mid-infrared spectra (4,000 to 400 cm^{-1}) or both MIR and land use (denoted LU) to predict soil aggregate stability. ME = mean error; RMSE = root mean square error; MEC = modelling efficiency coefficient.

Models	ME (mm)	RMSE (mm)	r^2	MEC
MIR_{PLSR}	0.00	0.62	0.45	0.44
$MIR_{PLSR+LU}$	0.00	0.50	0.63	0.63
MIR_{cubist}	-0.05	0.60	0.49	0.48
$MIR_{cubist+LU}$	-0.05	0.54	0.59	0.58

3.2. Lab-based approach

Based on the procedure described in Section 2.2.1, we calculated the VIF values on the MLR models. None of the VIF values were above 1.5, which represents a low level of multicollinearity among explanatory variables. All the explanatory variables can thus be used as input variables. For cubist calibration, the optimal number of committees was found to be 10 for all models.

The results of the cross-validation for the MLR and cubist models are shown in Tables 4 and 5 respectively. Overall, the validation statistics were very similar between the two models (i.e. MLR and cubist). However cubist models showed a greater bias (ME about -0.05) compared to MLR (ME = 0). Both models had a similar range of error (RMSE close to 0.50). Also, both the Person's r correlation coefficient and the MEC were close to 0.6, which means that about 60% of the variation of the MWD_{FW} data was explained by the models. Despite large similarities in validation statistics between the models, the MLR models had slightly better validation statistics, in particular for the variable set composed of: land use, SOC, pH_{water}, $CaCO_3$, and with very slight differences between the four textures sets: clay, clay & silt, clay & sand, and silt & sand. Considering a texture set with clay only or the addition of either sand or silt to clay led to similar validation statistics for the MLR models, we choose to use clay and silt which better describe texture than clay only.

Tables 4 and 5 also show that including $CaCO_3$ and pH_{water} or including only one of them leads to similar prediction quality statistics. For example, in the MLR model, the use of $CaCO_3$ and pH_{water} compared to the single use of pH_{water} decreased the RMSE of 1.2%, increased both r^2 and MEC of 0.01, which were small changes. In the same way, using $CaCO_3$ decreased the RMSE of 1.2%, increased both r^2 and MEC of 0.02, which were also small changes but these were slightly higher than the single use of pH_{water}. It can be explained with the higher linear correlation between MWD_{FW} and pH_{water} ($r = -0.30$) which was about twice higher than MWD_{FW} and $CaCO_3$ ($r = 0.16$) (Table 3). Theses two parameters act as redundant information, so the use of only one of these two parameters was possible. We chose to keep pH_{water} which is a more accessible analysis.

Hereafter, we focus the results on the model from Tables 4 and 5 which shows the best validation statistics, which is the lab-based model with the variable set composed of: land use, SOC, clay & silt, pH_{water}.

Table 6 shows the standardized regression coefficients of the MLR model (i.e. a model fitted with variables having zero mean and unit variance) as the resulting model is readily interpretable, with their corresponding T-test p-values. The T-test p-values indicate whether or not the predictor is meaningful for the model. The α threshold was set to 0.1. Note that the intercept corresponds to land use: "cropland". Land use is used as an indicator variable. The resulting coefficient for land use is the sum of the intercept (cropland) with the other land use coefficients (woodland or grassland). The standardized coefficients of the different land uses can be ranked in descending order as follows: woodland (coefficient is 2.14), grassland (coefficient is 1.52), cropland (coefficient is 0.83). The grassland coefficient was about twice higher than that cropland (1.83), or about 2.56 times higher than that of woodland. The global variable order of importance was: land use, clay, SOC, pH_{water} and silt. Clay and SOC coefficients were very close to one another (8%). Silt and pH_{water} had a minor contribution to the MWD_{FW} prediction (respectively 10 and 7% of the intercept) and the negative value of their coefficient shows that they had a negative contribution to the prediction of the MWD_{FW} value.

Using unstandardized variables for the MLR fitting with land use, SOC, clay & silt and pH_{water} we obtained Eq. 5, which can be reused with raw data in further works.

$$MWD_{FW} = \begin{cases} 0.7630 + 0.0017 \times clay + 0.0145 \times SOC - 0.0719 \times pH_{water} - 0.0004 \times silt, & \text{if croplands} \\ 2.068 + 0.0017 \times clay + 0.0145 \times SOC - 0.0719 \times pH_{water} - 0.0004 \times silt, & \text{if woodlands} \\ 1.4536 + 0.0017 \times clay + 0.0145 \times SOC - 0.0719 \times pH_{water} - 0.0004 \times silt, & \text{if grasslands} \end{cases} \quad (5)$$

with MWD_{FW} in mm and SOC, clay, and silt in g.·kg^{-1}.

3.3. Mid-infrared spectroscopy approach

Recall that for the mid-infrared approach we fitted PLSR and cubist models. For PLSR, the optimal number of components was found to be 5 whereas for the cubist model the optimal number of committees was 10.

Table 7 shows the validation statistics obtained by cross-validation for the spectroscopic models based on PLSR and cubist, built with either mid-infrared data only or both mid-infrared data and land use. Overall, models that included land use had better validation statistics. For example, the MEC of $MIR_{PLSR+LU}$ was 0.63 whereas it was 0.44 for the MIR_{PLSR} model fitted without land use, which meant a gain of 0.19 in model efficiency when using land use. Models were all without substantial bias, but the cubist models had a very small negative bias (ME = -0.05). The absence of bias explains why the r^2 and the MEC were close for all modelling approaches. Hereafter we interpret the best model for spectroscopic modelling, that is, the PLSR model which was built with both mid-infrared spectra and land use (i.e. $MIR_{PLSR+LU}$).

Fig. 4 shows the variable importance of the $MIR_{PLSR+LU}$ model obtained by permutation. The PLSR importance of predictors shows again that land use played an important role for the prediction quality of MDW_{FW} with a difference of RMSE about 4000 times higher than MIR wavenumbers. There are great variations in the importance of the different wavenumbers, with some prominent features (Fig. 4). Four important wavenumbers (i.e. with RMSE differences between values larger than 2.0×10^{-4}) are detected. There are in decreasing order of importance at: 1880 cm^{-1}, 600 cm^{-1}, 1790 cm^{-1} and 1145 cm^{-1}. In addition, they are some wavenumbers with a slightly lower importance: 2930 cm^{-1}, 1670 cm^{-1}, 3100 cm^{-1}, 2720 cm^{-1}, 2820 cm^{-1} and 500 cm^{-1}. Furthermore, lowest importance wavenumbers are the following: 1530 cm^{-1}, 3570 cm^{-1}, 3600 cm^{-1}, 720 cm^{-1}, 3870 cm^{-1}, 4000 cm^{-1}, 2520 cm^{-1}, 3960 cm^{-1}, 880 cm^{-1}, 950 cm^{-1}, 1380 to 1390 cm^{-1} and 770 cm^{-1}. Finally, very few regions had a minor negative impact (i.e. they do not contribute to the prediction, or are worsening the prediction) from 3620 to 3800 cm^{-1}, 1920 to 2450 cm^{-1} and 2560 cm^{-1} and 1430 cm^{-1}.

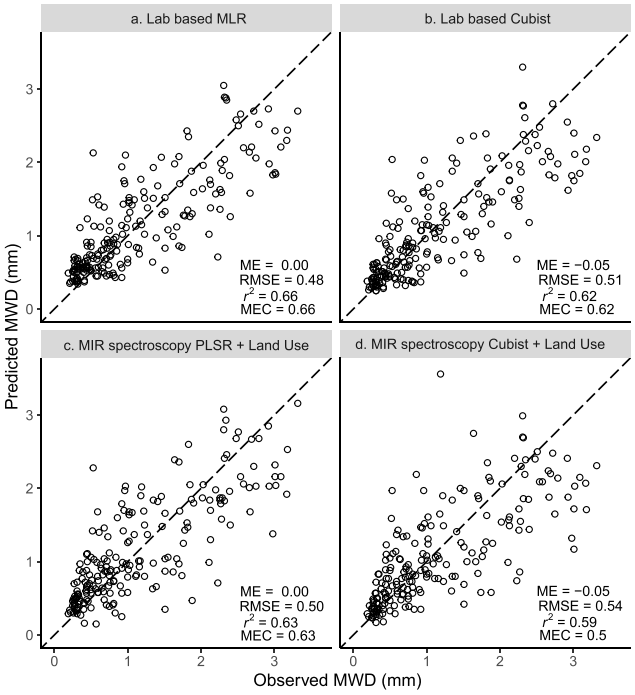


Fig. 5. Observed versus predicted values of MWD_{FW} obtained using a 10-fold cross-validation via models with the best prediction statistics for LB functions (a. MLR; b. cubist, Variables: land use, SOC, clay, silt, and pH_{water}) or MIR spectroscopic modelling (c. PLSR ; d. PLSR + land use). The dotted line represents y = x.

3.4. Comparison of the two approaches

The absence of bias observed previously for all modelling approaches is confirmed by the scatterplots of observed versus predicted values of MWD_{FW} (Fig. 5). The scatterplots show the predicted values from the best LB functions (with variables: land use, SOC, clay, silt, pH_{water}) (Table 4 and Table 5) and spectroscopic models (with variables: MIR spectra and land use) (Table 7). While there was no systematic over- or

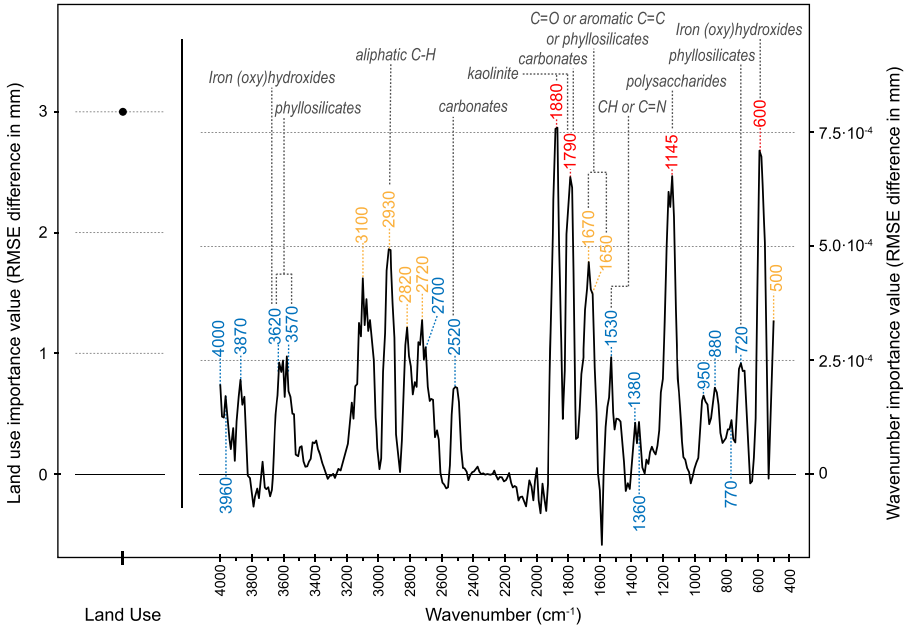


Fig. 4. Importance of predictors for MWD_{FW} prediction for the $PLSR_{LU}$ model (i.e. fitted with mid-infrared spectra and land use), numbers above the peaks give the wavenumber, colour indicate the importance of the peaks (red, orange, and blue, for respectively very important, moderately important, or minor importance).

under-estimation, there was more dispersion for MWD_{FW} values that are above 1.5.

4. Discussion

4.1. Estimating soil aggregate stability

Estimating the aggregate stability by the LB approach gave better result with MLR model using land use, SOC, clay, silt and pH_{water} and for MIR spectra approach with the PLSR model using MIR spectra and land use. The two strategies tested in this study for estimating aggregate stability with either common soil properties or mid-infrared spectra yielded similar validation statistics (±4% for all statistics, Tables 4 and 7). However, the strategy based on lab measurements had slightly better results (i.e. similar bias but slightly higher MEC and r^2) than those based on MIR spectra data. This is consistent with other comparable studies, in which it was found that LB gave slightly better predictions statistics than spectroscopy. For example Gomez et al. (2013), with a RMSE of +6% and a r^2 of −0.04 for the approach based on MIR (for a study site of about 800 km^2). However, it was found in others studies that the two approached yielded very different results, (e.g. Shi et al., 2020) in which the RMSE was +42% and r^2 was −0.13% (for a study site of 600 km^2). Overall, in our study both approaches showed good prediction of aggregate stability. Our LB PTF approach showed better results than the one developed by Annabi et al. (2017) in which they obtained a MEC of 0.29 vs. 0.66 in our study, and Gomez et al. (2013) in which they found a r^2 of 0.44 vs. 0.66 in our study. We found worse results than Shi et al. (2020) where the r^2 was 0.85 vs. 0.66 in our case. The MIR spectra approach showed better results than Gomez et al. (2013) ($r^2 = 0.48$ vs. 0.63 in our study) and Erktan et al. (2016b) ($r^2 = 0.45$ and RMSE = 0.62 vs. resp. 0.63 and 0.50 our study) but worse than Afriyie et al. (2020), Afriyie et al. (2022), and Shi et al. (2020) which had r^2 +0.15 in comparison to our study. The lower results in our study could be attributed to the use of a *K*-fold cross-validation which was not the case for Afriyie et al. (2020) (validation set with 30% of samples), Afriyie et al. (2022) (validation set with 30% of samples), and Shi et al. (2020) (validation set with 25% of samples), and also to the large variability of MWD_{FW} values, ranging from 0.19 to 3.32 mm (sd = 0.83 mm, Table 2), which makes modelling more challenging. This range of MWD_{FW} values is wider than the range observed in other studies (Afriyie et al., 2020; 2022; Annabi et al., 2017; Erktan et al., 2016a; 2016b; Gomez et al., 2013; Shi et al., 2020), whose widest MWD_{FW} range was 0.22-2.77 mm and which had MWD_{FW} values with less diversity (sd ranging between 0.30 and 0.55 mm). Also, our study shows a large variability in clay and SOC values. Other studies (e.g. Annabi et al., 2017; Erktan et al., 2016b; Gomez et al., 2013; Shi et al., 2020) had less clay variability, whose widest clay range was 35–179 $g{\cdot}kg^{-1}$ which is narrower than our study with 81–648 $g{\cdot}kg^{-1}$ (Table 2). However, the studies of Gomez et al. (2013) and Annabi et al. (2017), which had the same study site, had a wider clay range (46 to 747 $g{\cdot}kg^{-1}$) than our study site (81 to 648 $g{\cdot}kg^{-1}$). In addition, the range of SOC in other studies (e.g. Afriyie et al., 2020; 2022; Annabi et al., 2017; Gomez et al., 2013; Shi et al., 2020) was lower to ours (13 against 74 $g{\cdot}kg^{-1}$), but their minimums are lower than our (on average 2.2±0.9 $g{\cdot}kg^{-1}$ against 5.37 $g{\cdot}kg^{-1}$).

4.2. Interpretation of the models

Both approaches of this study confirmed that soil aggregate stability had a strong positive correlation with SOC and clay content, as it is well-documented (Baveye et al., 2020; Le Bissonnais et al., 2007). However, our results showed a smaller correlation between aggregate stability and SOC than some previous studies made in a similar context (e.g. Le Bissonnais et al., 2007; Shi et al., 2020). Shi et al. (2020), for example, recently found in a study on the region of the Belgian Loam Belt that correlation was $r = 0.84$, while it was 0.67 in our study (Table 3). This can be related to the differences between sampling material, as SOC was measured on the first 30 cm of soil and from 25 composite samples while aggregate stability was measured on undisturbed aggregate from the very topsoil layer. The LB PTF developed fitted with the MLR model (i.e. LB_{MLR}) highlighted known relationships between soil properties and aggregate stability. Clay and SOC had an important and positive effect, as shown by the regression coefficients in the LB_{MLR} model (Table 6). This is consistent with some of the previous studies: for example Annabi et al. (2017) and Gomez et al. (2013) showed the same order of magnitude for clay and SOC coefficients, but Gomez et al. (2013) found a negative coefficient for SOC in the northern region of Tunisia which is under a Mediterranean climate, which contrasts the findings of Le Bissonnais et al. (2007) who showed the predominant positive role of SOC under a Mediterranean climate in the South of France. Shi et al. (2020) found similar coefficients for SOC and pH_{water} in the Belgian loam belt under a temperate climate. However, the coefficient value for clay was much higher than in our study, which can be explained with the lower range of clay (7 to 14 $g{\cdot}kg^{-1}$ against 81 to 648 $g{\cdot}kg^{-1}$). In addition, clay content is not the only parameter by which clays can influence stability. Clays nature plays an ambivalent role in aggregation, for example clays with a higher cation exchange capacity (CEC) (e.g. smectitic clays) are more efficient on aggregation due to their higher physiochemical interaction capacity but they are sensitive to soil water electrolyte concentration which can destabilize these interactions (Amezketa, 1999; Boivin et al., 2004). The way in which SOC and clay interacts also influences wet aggregate stability, for examples by modifying its wettability (Chenu et al., 2000). These factors can not be studied here due to the few number of parameters selected for the LB approach, however, it should be noted that the use of clay content remains a robust way of assessing wet aggregate stability, especially when combined with other measures such as organic carbon (Barzegar et al., 1997; Le Bissonnais, 1996a). Further, the LB PTF in our study had a negative coefficient associated to pH_{water}, which implies a negative impact of pH_{water} on soil aggregate stability (i.e. the greater the pH, the lower the aggregate stability). In the literature, pH_{water} is usually negatively correlated to aggregate stability (Amezketa, 1999; Le Bissonnais and Arrouays, 1997). The accuracy statistics and preliminary results did not emphasize the need to use $CaCO_3$, and we thus did not proceed any further. The negligible difference in results when adding $CaCO_3$ to the set of common soil property is also found the literature (Erktan et al., 2016b; Gomez et al., 2013; Le Bissonnais et al., 2007). Theses results somehow contrasts with the literature on aggregate stability mechanisms (Amezketa, 1999; Le Bissonnais, 1988), but can be attributed to the low $CaCO_3$ content is our dataset (median of $CaCO_3$ is 1.0 $g{\cdot}kg^{-1}$). Further, the positive correlation between $CaCO_3$ and pH_{water} ($r = 0.40$, Table 3) can explain the need to use only one of the two variables. However, the LB PTF model did not allows to assess relation between land use and soil properties (i.e. SOC and texture), as properties coefficients are the same among land use formula. The LB cubist model was investigate with this purpose. However, its prediction quality was slightly lower than the LB PTF model, which means that even with the possibility of taking into account these relations it was not significant to improve the prediction quality. In this way, the size of the current data might be a limitation to study these interactions as a high variety of soil texture were explored and maybe with too few samples per land use to draw a tendency.

It was documented elsewhere (e.g. Stenberg et al., 2010) that soil aggregate stability does not link to a very specific spectral response, but instead a combination of overlapping regions. Most previous studies used vis-NIR spectral data for aggregated stability, whereas in our case we use MIR. The MIR spectra contains fundamental vibrations linked to soil chemical compounds, some of which are visible in Fig. 4 and can be interpreted. We associate the different wavenumbers to different soil component according to Stenberg et al. (2010), Soriano-Disla et al. (2014) and Afriyie et al. (2020) as follows: the peaks at about 1880 cm^{-1} and 1790 cm^{-1} are within the double-bond regions and are usually difficult to interpret but can be associated to the broad vibrational bond of kaolinite, quartz minerals or even organic matter functional groups

(Nguyen et al., 1991). The 1790 cm^{-1} peak can also be associated to carbonates (Gomez et al., 2022). The peak at 1170 cm^{-1} is related to polysaccharides (Stenberg et al., 2010). The organic matter is correlated with different bands: 2920 cm^{-1} (aliphatic C-H stretching), 1720 cm^{-1} (C═ O strecthing in carboxyls), 1660-1650 cm^{-1} (C═ O of amide groud), 1650-1600 cm^{-1} (aromatic C═ C and olefinic groups, conjugated C═ O groups), 1517–1590 cm^{-1} (N-H bending and C═ N stretching) and 1120-1000 (C-O stretching in alcohols, polysaccharides-type structures) (Fernández-Getino et al., 2010). Phyllosilicates occur at various wavenumbers (e.g. 3622, 3790, 1635 and 795 cm^{-1}) (Pärnpuu et al., 2022). Iron (oxy)hydroxydes occur around 3622, 795, and 400–700 cm^{-1} whereas carbonates occur near 2515, 1800, 1400, 880 and 715 cm^{-1} (Gomez et al., 2022; McDevitt and Baun, 1964; Stenberg et al., 2010). Overall, the important wavenumbers for prediction of the aggregate stability show that functional groups related to organic matter and clay minerals were essential. Some wavenumbers used by our model were also used by other studies models (Afriyie et al., 2020; 2022; Erktan et al., 2016b; Shi et al., 2020), for example the peaks 1790, 1530, 1380, 1145, 950, 880, 720 and 600^{-1}, mainly associated to organic matter, phyllosilicates and (oxy)hydroxydes. However, some wavenumbers found to be important in other studies were not important in our case (i.e. 2851, 1800, 1610, 1275 and 1200 cm^{-1} in Afriyie et al. (2020) or the bands near 1900 cm^{-1} in Shi et al. (2020)). The opposite was also observed for the peaks at 1145 cm^{-1}, 600 cm^{-1} which were important in our case, but for which we did not find similar results in the literature. These wavenumbers usually refer to SOC (Fernández-Getino et al., 2010) and iron oxydes (McDevitt and Baun, 1964). The differences between our models and the literature may be related to the large diversity of soils in our dataset. Indeed, the work of Afriyie et al. (2020); Shi et al. (2020) are transposable to a sub-part of our soils, the North Eastern part of France, and the work of Erktan et al. (2016b); Gomez et al. (2013) to the South Eastern part of France. The main differences with other studies might be due to the large geographical scale of our study, covering a wide range of different soils and pedoclimatic zones. For comparison, the area studied by Afriyie et al. (2020) is about 0.2 km^2 while it is 40 km^2 for Erktan et al. (2016b), 300 km^2 for Gomez et al. (2013), about 700 km^2 for Shi et al. (2020) and 800 km^2 for Annabi et al. (2017). Previous studies also considered a specific land use only, such a cropland in Afriyie et al. (2020), Shi et al. (2020), and Afriyie et al. (2022). In Gomez et al. (2013) and Annabi et al. (2017) the main focus was on pedological or lithological classes and land use types were not listed.

In our case, both approaches highlighted the importance of land use for the prediction of aggregate stability. Land use is known to influence soil characteristics, such as organic carbon (Chenu et al., 2000; Le Bissonnais et al., 2007), which stress the indirect impact of soil management on aggregate stability (through the addition of carbon to the soil or the changes in SOC dynamics). With the LB_{MLR} approach it was found that woodland had the highest aggregate stability values, whereas cropland had the lowest values. We stress that assessment of interactions between predictor variables was not included in the LB PTF approach using MLR but was indirectly taken into account in the cubist model. By fitting and inspecting the rules produced by a cubist model with 1 committee (not shown), we observed that the nodes of the tree separated the dataset according to land uses into two separate groups (croplands and grassland/woodland). However, as our dataset is unbalanced with a large number of observations for croplands and few data points in woodland areas, these interactions did not produce accurate predictions. Thus, we did not pursue any further with the assessment of interactions between variables but this might be a valuable contribution in future studies if the dataset size is sufficiently large and balanced between land uses.

4.3. Inference strategies to estimate aggregate stability

According to the validation statistics, both approaches had similar level of reliability to estimate aggregate stability. The LB PTF is an easily understandable approach and relies on easy-to-obtain soil properties. Aggregate stability can be estimated using the LB approach based on MLR, using an existing legacy dataset and the estimated regression coefficients reported in this study (Eq. 5). The MIR approach also appears to be reliable. MIR scans of the soil are easier, faster, cheaper to obtain (5-10 € per soil sample), when compared to estimating soil properties (40-50 € per sample for texture, SOC, pH_{water} and $CaCO_3$) or aggregate stability measurement (70-80 € per sample). However the investment for the acquisition of the spectrometer is significant (about 30-80 k€) although it is not unreasonable when compared to other common laboratory equipment used for the analysis of soil texture, SOC or $CaCO_3$. Aggregate stability can exhibit large spatial and temporal variations. Soil scientists and stakeholders (e.g. farmers, water management agencies) can only afford to collect and analyse few samples, leading to poor characterizations of the spatial mean of this property and its variance. Our study shows the importance of land use for the prediction of aggregate stability. Further studies could similarly investigate the possibility of predicting aggregate stability within a specific land use and focus instead on, for example, different cropland managements.

4.4. Limitations

Unlike the LB model, which allows us to see the influence of each property on the outcome of the prediction and then understand its value, the MIR model is an indirect measurement of the soil. As shown in the LB PTF coefficients interpretation (i.e. Section 2.2.1), aggregate stability was strongly correlated to SOC and clay content. For the MIR_{PLSR} model, the link between the spectra used as input variables and the soils properties (e.g. SOC) is less obvious. The LB PTF model, which is based on slowly evolving or stable soil properties while aggregate stability is a very dynamic property, is therefore a first approach to assess the potential aggregate stability of a studied area. However, if some bands of the MIR spectra are linked to soil properties that evolve with the same dynamics as aggregate stability, the MIR measurements could allow a more dynamic prediction of aggregate stability.

Our dataset allowed us to study the possibility of predicting aggregate stability at a country scale with a wide diversity of soil and land uses. However, to study a wide diversity of soil the dataset was built from samples collected from two different sampling campaigns. These two campaigns were separated by about 10 years. Even if we favour measures with no or the least time difference, some sites were composites of these two campaigns. Nevertheless, to build our models, especially for the approach based on pedotransfer functions, we used perennial properties that do not evolve rapidly over time, so we assume the error in measurement due to sampling in two different dates is negligible.

Finally, despite the wide diversity of soils, the $CaCO_3$ content overall was low (median was 1 $g{\cdot}kg^{-1}$). Knowing its importance for aggregate stability (Amezketa, 1999; Le Bissonnais and Arrouays, 1997) the LB model showed that the use $CaCO_3$ or pH_{water} give the same results, so they look to bear the same information. The slightly better modelling results with pH_{water} can be related to its ability to show differences in acidic conditions when it is not possible with $CaCO_3$. However, with the MIR approach the PLSR model show to give importance to bands associated to $CaCO_3$ as it contribute to the prediction with the strong peak at 1800 cm^{-1}, which was also highlighted in similar studies (Erktan et al., 2016b; Gomez et al., 2013).

5. Conclusion

We investigated two approaches to estimate a normalized international method of soil aggregate stability index (denoted MWD_{FW}), using an extensive dataset covering mainland France and by means of i) commonly-measured soil properties (lab-based), and ii) a mid-infrared spectroscopy approaches. From the Results and Discussion we draw the following conclusions:

- We found strong positive correlations between aggregate stability measurements, organic carbon and clay content.
- Land use was important for the prediction of aggregate stability.
- Both approaches to estimate aggregate stability had similar prediction validation statistics, but the lab-based approach had slightly better results compared to the spectroscopic approach.
- The results support the use of MIR measurements as a cost and time effective measure for the prediction of soil aggregate stability, without substantial decrease in prediction accuracy compare to pedotransfer function. When the initial investments for the acquisition of a spectrometer is too high, simple pedotransfer function based on commonly-measured soil properties can be used to estimate aggregate stability.
- It is possible to predict aggregate stability at the scale of a country in the context of a wide diversity of soil and different land uses.

Since structural stability is highly variable in time and space, we envision building a dataset with repeated measurements over time, for example, by measuring it throughout a year or during high erosive periods (summer or autumn). This might help in both reducing the measurement error and in evaluating the aggregate stability change over time and therefore the soil susceptibility to erosion for a region at a country scale. The methodology developed in this study can be used for this purpose. Further, the developed formula can be used to predict wet aggregate stability trends at the country level and thus provide a support to regional management policies. The acquisition of more measurements per land use, the addition of new land use types as well as different managements within a land use will help to improve and clarify models.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available upon request https://doi.org/10.15454/QSXKGA.

Acknowledgements

The soil sampling and the analyses of physicochemical properties of soils were supported by the French Scientific Group of Interest on soils, GIS Sol, involving the French Ministry in charge of the Ecological Transition (MTE), the French Ministry in charge of the Agriculture and Food (MAA), the French Agency for Environment and Energy Management (ADEME), the French Biodiversity Agency (OFB), the National Institute of Geographic and Forest Information (IGN), the French National Research Institute for Sustainable Development, and the National Institute for Agronomic and Environmental Research (INRAE). We thank all the soil surveyors and technical assistants involved in sampling the sites. This project received funding from INRAE department Agro-ecosystem. Spectrum acquisition on RMQS samples was supported by the MIRSol project, which was funded by ADEME (contract 0675C0102). We would like to thank also F Darboux for sharing Agresta data.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.soisec.2023.100088

References

Afriyie, E., Verdoodt, A., Mouazen, A.M., 2020. Estimation of aggregate stability of some soils in the loam belt of belgium using mid-infrared spectroscopy. Sci. Total Environ. 744, 140727.
Afriyie, E., Verdoodt, A., Mouazen, A.M., 2022. Potential of visible-near infrared spectroscopy for the determination of three soil aggregate stability indices. Soil Tillage Res. 215, 105218.
Akinwande, M.O., Dikko, H.G., Samson, A., et al., 2015. Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis. Open J. Stat. 5 (07), 754.
Algayer, B., Le Bissonnais, Y., Darboux, F., 2014. Short-term dynamics of soil aggregate stability in the field. Soil Sci. Soc. Am. J. 78 (4), 1168–1176.
Amezketa, E., 1999. Soil aggregate stability: a review. J. Sustain. Agric. 14 (2–3), 83–151.
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.
Arrouays, D., Jolivet, C., Boulonne, L., Bodineau, G., Ratié, C., Saby, N.P.A., Grolleau, E., 2003. Le reseau de mesures de la qualite des sols de france (RMQS). Etude et Gestion des Sols 10 (4), 241–250.
Barthès, B.G., Roose, E., 2002. Aggregate stability as an indicator of soil susceptibility to runoff and erosion; validation at several levels. Catena 47 (2), 133–149.
Barzegar, A.R., Nelson, P.N., Oades, J.M., Rengasamy, P., 1997. Organic matter, sodicity, and clay type: influence on soil aggregation. Soil Sci. Soc. Am. J. 61 (4), 1131–1137.
Baveye, P.C., Schnee, L.S., Boivin, P., Laba, M., Radulovich, R., 2020. Soil organic matter research and climate change: merely re-storing carbon versus restoring soil functions. Front. Environ. Sci. 8, 161.
Boivin, P., Garnier, P., Tessier, D., 2004. Relationship between clay content, clay type, and shrinkage properties of soil samples. Soil Sci. Soc. Am. J. 68 (4), 1145–1153.
Bouma, J., 1989. Using Soil Survey Data for Quantitative Land Evaluation. Advances in Soil Science. Springer, New York, NY, pp. 177–213.
Brus, D.J., Saby, N.P.A., 2016. Approximating the variance of estimated means for systematic random sampling, illustrated with data of the french soil monitoring network. Geoderma 279, 77–86.
Cañasveras, J.C., Barrón, V., Del Campillo, M.C., Torrent, J., Gómez, J.A., 2010. Estimation of aggregate stability indices in mediterranean soils by diffuse reflectance spectroscopy. Geoderma 158 (1–2), 78–84.
Chenu, C.; Abiven, S.; Annabi, M.; Barray, S; Bertrand, M;. Bureau, F.; Cosentino, D.; Darboux, F.; Duval, O.; Fourrié, L.; Francou, C.; Houot, S.; Jolivet, C.; Laval, K.; Le Bissonnais, Y.; Lemée, L.; Menasseri, S.; Pétraud, J.P.; Verbeque, B. (2011). Mise au point d'outils de prévision de l'évolution de la stabilité de la structure de sols sous l'effet de la gestion organique des sols. Etudes et Gestion des Sols, 18(3):137–151.
Chenu, C., Le Bissonnais, Y., Arrouays, D., 2000. Organic matter influence on clay wettability and soil aggregate stability. Soil Sci. Soc. Am. J. 64 (4), 1479–1486.
Emadodin, I., Reiss, S., Bork, H.R., 2009. A study of the relationship between land management and soil aggregate stability (case study near albersdorf, northern-Germany). J. Agric. Biol. Sci. 4 (4), 48–53.
Emerson, W.W., 1967. A classification of soil aggregates based on their coherence in water. Soil Res. 5 (1), 47–57.
Erktan, A., Cécillon, L., Graf, F., Roumet, C., Legout, C., Rey, F., 2016. Increase in soil aggregate stability along a mediterranean successional gradient in severely eroded gully bed ecosystems: combined effects of soil, root traits and plant community characteristics. Plant Soil 398 (1), 121–137.
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.
Fernández-Getino, A.P., Hernández, Z., Buena, A.P., Almendros, G., 2010. Assessment of the effects of environmental factors on humification processes by derivative infrared spectroscopy and discriminant analysis. Geoderma 158 (3–4), 225–232.
Gomez, C., Chevallier, T., Moulin, P., Arrouays, D., Barthès, B.G., 2022. Using carbonate absorbance peak to select the most suitable regression model before predicting soil inorganic carbon concentration by mid-infrared reflectance spectroscopy. Geoderma 405, 115403.
Gomez, C., Le Bissonnais, 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, 86–97.
Grinand, C., Barthès, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C., Caria, G., Bernoux, M., 2012. Prediction of soil organic and inorganic carbon contents at a national scale (france) using mid-infrared reflectance spectroscopy (MIRS). Eur. J. Soil Sci. 63 (2), 141–151.
Guo, L., Shen, J., Li, B., Li, Q., Wang, C., Guan, Y., D'Acqui, L.P., Luo, Y., Tao, Q., Xu, Q., 2020. Impacts of agricultural land use change on soil aggregate stability and physical protection of organic c. Sci. Total Environ. 707, 136049.
Gupta, V.V., Germida, J.J., 2015. Soil aggregation: influence on microbial biomass and implications for biological processes. Soil Biol. Biochem. 80, A3–A9.
Institut National De La Recherche Agronomique, et al. Analyses physico-chimiques des sites du Réseau de Mesures de la Qualité des Sols (RMQS) du territoire métropolitain pour la 1ère campagne (2000–2009), avec coordonnées théoriques. Type: dataset. https://data.inrae.fr/citation?persistentId=doi:10.15454/QSXKGA.
ISO 10390:2005, 2005. Soil quality - Determination of pH. Standard. International Organization for Standardization, Geneva, CH.
ISO 10693:1995, 2014. Soil quality – Determination of carbonate content – Volumetric method. Standard. International Organization for Standardization, Geneva, CH.

ISO 10694:1995, 1995. Soil quality – Determination of organic and total carbon after dry combustion (elementary analysis). Standard. International Organization for Standardization, Geneva, CH.
ISO 10930:2012, 2012. Soil quality – Measurement of the stability of soil aggregates subjected to the action of water. Standard. International Organization for Standardization, Geneva, CH.
ISO 11277:2009, 2009. Soil quality - Determination of particle size distribution in mineral soil material - Method by sieving and sedimentation. Standard. International Organization for Standardization, Geneva, CH.
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.
Landre, A., Saby, N.P.A., Barthès, B.G., Ratié, C., Guerin, A., Etayo, A., Minasny, B., Bardy, M., Meunier, J.-D., 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 Bissonnais, Y., 1988. Analyse des mécanismes de désagrégation et de la mobilisation des particules de terre sous l'action des pluies. Université d'Orléans.
Le Bissonnais, Y., 1996. Aggregate stability and assessment of soil crustability and erodibility: i. theory and methodology. Eur. J. Soil Sci. 47 (4), 425–437.
Le Bissonnais, Y., 1996. Soil Characteristics and Aggregate Stability. In: Agassi, M. (Ed.), Soil Erosion, Conservation and Rehabilitation. Marcel Dekker Inc., New York, pp. 41–60.
Le Bissonnais, Y., Arrouays, D., 1997. Aggregate stability and assessment of soil crustability and erodibility: II. application to humic loamy soils with various organic carbon contents. Eur. J. Soil Sci. 48 (1), 39–48.
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. Eur. J. Soil Sci. 58 (1), 188–195.
Le Bissonnais, Y., Singer, M.J., 1993. Seal formation, runoff, and interrill erosion from seventeen california soils. Soil Sci. Soc. Am. J. 57 (1), 224–229.
Lehmann, A., Zheng, W., Rillig, M.C., 2017. Soil biota contributions to soil aggregation. Nature Ecol. Evol. 1 (12), 1828–1835.
Liland, K. H., Mevik, B.-H., Wehrens, R., 2021. pls: Partial Least Squares and Principal Component Regression. R package version 2.8-0. https://CRAN.R-project.org/package=pls.
McDevitt, N.T., Baun, W.L., 1964. Infrared absorption study of metal oxides in the low frequency region (700-240 cm- 1). Spectrochim. Acta 20 (5), 799–808.
Molnar, C., Bischl, B., Casalicchio, G., 2018. Iml: an r package for interpretable machine learning. JOSS 3 (26), 786. https://doi.org/10.21105/joss.00786.https://joss.theoj.org/papers/10.21105/joss.00786
Nguyen, T.T., Janik, L.J., Raupach, M., 1991. Diffuse reflectance infrared fourier transform (DRIFT) spectroscopy in soil studies. Soil Res. 29 (1), 49–67.
Novelli, L.E., Caviglia, O.P., Wilson, M.G., Sasal, M.C., 2013. Land use intensity and cropping sequence effects on aggregate stability and c storage in a vertisol and a mollisol. Geoderma 195, 260–267.
Pärnpuu, S., Astover, A., Tõnutare, T., Penu, P., Kauer, K., 2022. Soil organic matter qualification with FTIR spectroscopy under different soil types in estonia. Geoderma Region. e00483.
Paul, B.K., Vanlauwe, B., Ayuke, F., Gassner, A., Hoogmoed, M., Hurisso, T.T., Koala, S., Lelei, D., Ndabamenye, T., Six, J., 2013. Medium-term impact of tillage and residue management on soil aggregate stability, soil carbon and crop productivity. Agric. Ecosyst. Environ. 164, 14–22.
Quinlan, J.R., et al., 1992. Learning with continuous classes. 5th Australian joint conference on artificial intelligence, Vol. 92. World Scientific, pp. 343–348.
R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.
Shi, P., Castaldi, F., van Wesemael, B., Van Oost, K., 2020. Vis-NIR spectroscopic assessment of soil aggregate stability and aggregate size distribution in the belgian loam belt. Geoderma 357, 113958.
Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil Tillage Res. 79 (1), 7–31.
Six, J., Paustian, K., Elliott, E.T., Combrink, C., 2000. Soil structure and organic matter i. distribution of aggregate-size classes and aggregate-associated carbon. Soil Sci. Soc. Am. J. 64 (2), 681–689.
Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., Macdonald, L.M., McLaughlin, M. J., 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49 (2), 139–186.
Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J., 2010. Visible and near Infrared Spectroscopy in Soil Science. Advances in Agronomy, Vol. 107. Elsevier, pp. 163–215.
Stevens, A., Ramirez-Lopez, L., 2021. An introduction to the prospectr package. R package version 0.2.2.
Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y.A., Padarian, J., et al., 2017. Pedotransfer functions in earth system science: challenges and perspectives. Rev. Geophys. 55 (4), 1199–1256.
Wadoux, A.M.J.-C., Malone, B., Minasny, B., Fajardo, M., McBratney, A.B., 2021. Soil Spectral Inference With R: Analysing Digital Soil Spectra using the R Programming Environment. Springer Nature, Cham.