



**HAL**  
open science

## Mapping Topsoil Behavior to Compaction at National Scale from an Analysis of Field Observations

Anne C Richer-De-Forges, Dominique Arrouays, Songchao Chen, Zamir Libohova, Dylan E Beaudette, Hocine Bourenane

► **To cite this version:**

Anne C Richer-De-Forges, Dominique Arrouays, Songchao Chen, Zamir Libohova, Dylan E Beaudette, et al.. Mapping Topsoil Behavior to Compaction at National Scale from an Analysis of Field Observations. *Land*, 2024, 13 (7), pp.1014. 10.3390/land13071014 . hal-04638517

**HAL Id: hal-04638517**

**<https://hal.inrae.fr/hal-04638517>**

Submitted on 8 Jul 2024

**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.



Distributed under a Creative Commons Attribution 4.0 International License

## Article

# Mapping Topsoil Behavior to Compaction at National Scale from an Analysis of Field Observations

Anne C. Richer-de-Forges <sup>1</sup>, Dominique Arrouays <sup>1,\*</sup>, Songchao Chen <sup>2,3</sup>, Zamir Libohova <sup>4</sup>,  
Dylan E. Beaudette <sup>5</sup> and Hocine Bourennane <sup>1</sup>

<sup>1</sup> INRAE, Info&Sols, 45075 Orléans, France; anne.richer-de-forges@inrae.fr (A.C.R.-d.-F); hocine.bourennane@inrae.fr (H.B.)

<sup>2</sup> ZJU-Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 311215, China; chensongchao@zju.edu.cn

<sup>3</sup> College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

<sup>4</sup> USDA-ARS Dale Bumpers Small Farms Research Center, 6883 S. State Hwy. 23, Booneville, AR 72927, USA; zamir.libohova@usda.gov

<sup>5</sup> USDA-NRCS, Soil and Plant Science Division, 19777 Greenley Rd, Sonora, CA 95370, USA; dylan.beaudette@usda.gov

\* Correspondence: dominique.arrouays@inrae.fr

**Abstract:** Soil compaction is one of the most important and readily mitigated threats to soil health. Digital Soil Mapping (DSM) has emerged as an efficient method to provide broad-scale maps by combining soil information with environmental covariates. Until now, soil information input to DSM has been mainly composed of point-based quantitative measurements of soil properties and/or of soil type/horizon classes derived from laboratory analysis, point observations, or soil maps. In this study, we used field estimates of soil compaction to map soil behavior to compaction at a national scale. The results from a previous study enabled clustering of six different behaviors using the in situ field observations. Mapping potential responses to soil compaction is an effective land management tool for preventing future compaction. Random forest was used to make spatial predictions of soil behavior to compaction over cultivated soils of mainland France (about 210,000 km<sup>2</sup>). Modeling was performed at 90 m resolution. The map enabled us to spatially identify clusters of possible responses to compaction. Most clusters were consistent with known geographic distributions of some soil types and properties. This consistency was checked by comparing maps with both national and local-scale external sources of soil information. The best spatial predictors were available digital maps of soil properties (clay, silt, sand, organic carbon (SOC) content, and pH), some indicators of soil structural quality using SOC and clay content, and environmental covariates (T °C and relief-related covariates). Predicted maps were interpretable to support management recommendations to mitigate soil compactness at the soil–scape scale. Simple observational field data that are usually collected by soil surveyors, then stored and available in soil databases, provide valuable input data for digital mapping of soil behavior to compaction and assessment of inherent soil sensitivity to compaction.

**Keywords:** soil; compactness; on-field observations; indicators; sensitivity to compaction; clustering; digital soil mapping; random forest



**Citation:** Richer-de-Forges, A.C.; Arrouays, D.; Chen, S.; Libohova, Z.; Beaudette, D.E.; Bourennane, H. Mapping Topsoil Behavior to Compaction at National Scale from an Analysis of Field Observations. *Land* **2024**, *13*, 1014. <https://doi.org/10.3390/land13071014>

Received: 9 May 2024

Revised: 27 June 2024

Accepted: 4 July 2024

Published: 8 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Digital soil mapping (DSM) started to emerge as a mature soil mapping method at the beginning of the 2000s [1,2]. The initial concept of DSM relied on mapping soil properties and/or soil classes using information on soils and spatially exhaustive covariates. In their seminal article, McBratney et al. [2] proposed a model called SCORPAN (Soil, Climate, Organisms, Relief, Parent material, Age, and N for location). The basic principle of SCORPAN is to model statistical relationships between information from soil sample measurements, soil descriptions, or pre-existing soil maps, and spatialized environmental

covariates. The SCORPAN model is based on the work and concepts initially developed by Jenny [3]. Minasny and McBratney [4] provided a brief history of the DSM from the beginning. They identified the development of the first approaches towards the end of the 1970s, followed by significant development between the 1980s and the beginning of 2000, and then by exponential growth after 2000. Initially developed in the academic realm, DSM is now operational at numerous scales and on many continents, delivering many products from landscape to world scale (see reviews and syntheses [5–10]). One of the recent trends is the move from DSM of soil attributes to digital soil assessment (DSA) of soil function, behaviors, and services [11].

Most of the DSM soil input data come largely from georeferenced soil databases (e.g., [9,11–22]). Consequently, soil input data to DSM are most frequently quantitative data, e.g., point-based laboratory soil analyses at various soil depths or proximal- or remote sensing-derived soil data (e.g., [23–29]), though some studies use conventional soil maps (e.g., [30–35]).

However, qualitative data from field expertise and soil pit descriptions have been much less used as soil input data for DSM, though some inferences were derived from soil types or classes (e.g., [36–40]) or from specific soil features and diagnostic horizons (e.g., [41–47]).

Soil scientists and soil surveyors conduct many soil descriptions during their field work. These descriptions use national instructions and standard protocols described in detail in field books (e.g., [48–50]). However, the qualitative information gathered by soil surveyors and stored in soil databases was rarely used especially for estimating soil physical properties and behaviors and relating them to quantitative soil properties. Additionally, to our present knowledge, field observations have never been used as input data for predictive mapping of soil physical properties. Among these physical properties, we focus on topsoil compactness. Rather than a static prediction of soil compactness, we aim to map various soil behaviors to compaction, which is an important component of changes in soil condition and health.

A previous study [51] used data on numerous in situ topsoil observations conducted in the framework of conventional soil surveys over the French territory. This study involved (i) comparing field observations to usual indicators of soil compactness and sensitivity to compaction, (ii) identifying relationships between field estimates of soil compactness degree and some of the measured properties, and (iii) distinguishing six clusters of topsoil behavior to compaction by using hierarchical clustering. However, this study did not fully explain and interpret all the clusters, as some of them had incomplete data.

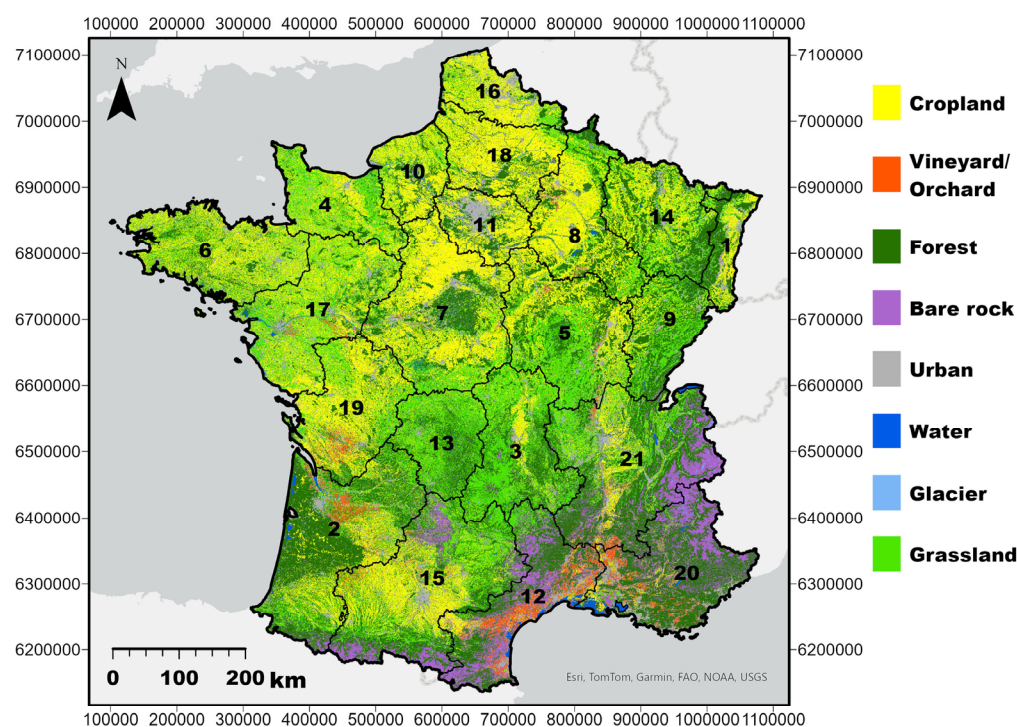
We postulate that using a DSM approach for spatial prediction of these clusters could provide additional information about the controlling factors of soil behavior to compaction. Indeed, searching for relevant covariates, such as particle-size distribution, indicators of sensitivity to compaction, and other potentially relevant spatial data (e.g., climate, relief, soil parent material), could bring new insights to the understanding and interpretation of soil behavior to compaction. In addition, from a practical point of view, mapping these behaviors is an essential step to inform local actors and advisers about inherent soil sensitivity to compaction. Therefore, we aim to use DSM to spatially predict clusters of cultivated topsoil behavior to compaction over the French territory. This approach clearly moves from DSM prediction of rather static soil attributes to DSA of soil behavior to compaction.

Our main objectives are to (i) map soil behavior to compaction at a national scale, (ii) check if this map makes sense from a soil geography point of view, (iii) provide additional insights into the controlling factors of soil behavior to compaction, and (iv) check if these spatial clusters can inform decision-making and at which scale.

## 2. Material and Methods

### 2.1. Study Area

Located in western Europe, mainland France is in the temperate climatic zone, characterized by warm summers and moderately cold winters. Depending on latitude, longitude, and elevation, mainland France exhibits some contrasts from the climates of oceanic, continental, and Mediterranean settings. Mainland France covers about 550,000 km<sup>2</sup>, including about 210,000 km<sup>2</sup> of arable lands. Agricultural systems are very diverse, including large farms of intensively managed field crops, mainly located on large plains of deep silty soils and irrigated calcareous soils. Livestock farming is dominant in certain regions such as Brittany, part of Normandy, and Massif Central. Perennial crops include vineyards and orchards (Figure 1).



**Figure 1.** Main land uses in mainland France after the Theia OSO Land Cover Map [52]. Areas: cropland: 160,000 km<sup>2</sup>; vineyard/orchard: 10,000 km<sup>2</sup>; forest: 160,000 km<sup>2</sup>; bare-rock: 34,000 km<sup>2</sup>; urban: 56,000 km<sup>2</sup>; water: 6000 km<sup>2</sup>; glacier: 235 km<sup>2</sup>; grassland: 80,000 km<sup>2</sup> (permanent); 40,000 km<sup>2</sup> (non-permanent). Regions: 1: Alsace; 2: Aquitaine; 3: Auvergne; 4: Basse-Normandie; 5: Bourgogne; 6: Bretagne; 7: Centre-Val-de-Loire; 8: Champagne-Ardennes; 9: Franche-Comté; 10: Haute-Normandie; 11: Ile-de-France; 12: Languedoc-Roussillon; 13: Limousin; 14: Lorraine; 15: Midi-Pyrénées; 16: Nord-Pas-de-Calais; 17: Pays de la Loire; 18: Picardie; 19: Poitou-Charentes; 20: Provence-Alpes-Côte-d’Azur; 21: Rhône-Alpes.

The present study is limited to arable soils (croplands, vineyards/orchards, and non-permanent grasslands). French soils are highly diverse due to a combination of complex geology, various climates, and a large range of elevations and geomorphology [53].

Mainland France has a diverse range of cropping systems influenced by its varied climate, soil types, and agricultural traditions. One of the strong trends of the last 50 years is the specialization of agricultural systems—some producing crops without livestock, others breeding animals almost without crops—which has led to the specialization of some regions that were traditionally dedicated to mixed crop agriculture–livestock systems. Some regions with suitable soil and climate conditions have become major crop regions. Other regions benefiting from favorable commercial and industrial situations have concentrated livestock activities or vegetable production. The modernization of agriculture has, therefore,



resulted in the specialization of some territories to a limited number of productions. Here, we present a brief overview of the main cropping systems found across different regions.

- Cereal cropping systems mainly include wheat, barley, and maize crops. Wheat is mainly grown in northern and central France, particularly in regions such as Ile-de-France, Picardie, and Centre-Val-de-Loire. Winter wheat is the most common variety. Barley is grown alongside wheat in Lorraine, Nord-Pas-de-Calais, Picardie, and areas with more marginal soils. Maize is widely cultivated in Aquitaine, Midi-Pyrénées, and in Alsace.
- Oilseed and protein crops systems are based on rapeseed (Canola), commonly grown in Nord-Pas-de-Calais and Picardie and in eastern France, including Bourgogne, Champagne-Ardenne, and Alsace, whereas sunflower and soybean are mainly grown in Poitou-Charentes, Aquitaine, and Midi-Pyrénées.
- In these two cropping systems, conventional tillage by plowing is the most frequent practice. However, many practices between direct sowing and conventional tillage are currently developing. An example of such practices is occasional plowing, which is carried out to resolve situations that have become problematic after a few years of stopping plowing.
- Root and tuber cropping systems include sugar beet production and potatoes in Nord-Pas-de-Calais, Picardie, and Bretagne. Conventional tillage includes deep plowing to prepare the soil for planting, followed by harrowing and sometimes ridging to ensure proper root and tuber development.
- Viticulture systems are located in small terroirs characterized by specific combinations of soil and climate conditions. However, the Languedoc-Roussillon region is characterized by a large area of vineyards. Most vines are plowed deeply before planting, then either grassed or tilled superficially depending on the climate and water availability. While many vineyards are shifting towards reduced tillage or no-till practices, tillage is still used in some areas for weed control and soil management. This may include plowing between vine rows or harrowing to manage the soil surface.
- Vegetable production systems are mainly concentrated in some key regions, including Bretagne (e.g., artichokes, cauliflower, and carrots), Provence-Alpes-Côte-d'Azur (e.g., tomatoes, zucchini, and eggplants), Nord-Pas-de-Calais and Picardie (potatoes, peas, beans), and along the large valleys of the main rivers. Conventional tillage is common, with frequent plowing and soil preparation to manage diverse vegetable crops. This system often involves multiple cropping cycles per year.
- Fruit systems include apple orchards, prominent in Bretagne and Basse-Normandie, while stone fruits (apricots, peaches, and cherries) are common in the Rhône-Alpes and Provence-Alpes-Côte-d'Azur. Orchards are either grassed or tilled superficially.
- Forage and pasture systems include permanent extensive grasslands found in regions like Auvergne, Basse-Normandie, Bretagne, Franche-Comté, and Limousin, supporting dairy and beef cattle. These cropping systems were excluded from this study. Non-permanent forage crops are included in some rotations to support livestock feed needs, especially in areas where mixed farming systems still exist. In the latter case, the soils are often plowed before seedling. Non-permanent forage crops are included in the present study.

While conventional tillage remains common, there is a growing trend towards conservation tillage and no-till practices in France, driven by energy costs, environmental concerns, and the need for sustainable farming practices.

## 2.2. Soil Data

Soil data extraction and filtering are detailed in [51]. Briefly, topsoil data from cultivated French lands were extracted from available records of the soil profiles from the French national database. We applied filtering processes to retain only cultivated topsoil and remove a couple of O horizons that were observed in fields recently converted from forest to cultivation. Topsoils were collected from a large variety of crop rotations and

management practices. Soil compactness was described in the field as moisture slightly lower than field capacity and far from the permanent wilting point. The compactness is an ordered qualitative variable defined in 4 soil compactness classes (SCC). We chose this on-field test because of its simplicity and the large amount of available information, i.e., SCC observations collected on about 14,500 topsoil horizons (Figure S1). For these horizons, we also extracted quantitative variables corresponding to laboratory analyses, e.g., particle-size fractions, soil organic carbon content (SOC), bulk density (BD), pH, cation exchange capacity (CEC), C/N ratio, and CaCO<sub>3</sub> content (see Table 2 in [51]).

The SCCs were compared to nine commonly used indicators, either belonging to general indicators of structural quality and various soil physical properties such as soil structural stability and soil porosity or specifically focused on soil compactness. These nine indicators are described in [51]. In an orthonormal space of silt% on the abscissa and clay% on the ordinate, we divided the space into squares of 2% × 2%, called soil texture cells (STcells) hereafter. Within each STcell, we calculated the frequency of each SCC. From this frequency, we derived four quantitative variables (SCC-1, SCC-2, SCC-3, and SCC-4), each equal to the frequency of a given SCC within an STcell.

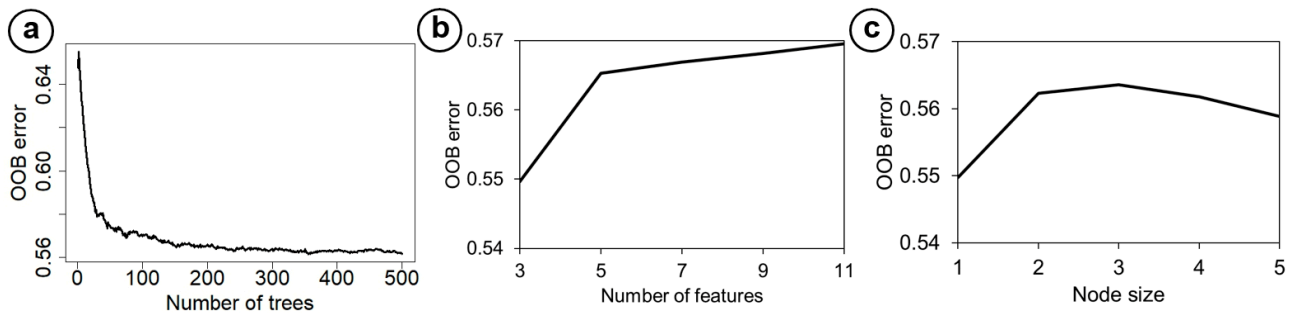
We then performed a clustering of STcells using the four above-mentioned quantitative variables to characterize each STcell. Hierarchical clustering methods and results are described in [51]. In order to choose the number of clusters, we compared the following methods: (i) changes in the sum of squared errors with an increasing number of clusters; (ii) goodness of clustering measure, i.e., the “gap” statistic [54]; (iii) the number of STcells by cluster with the increasing number of clusters. Full results are detailed in [51]. The optimal number of clusters was six. Clusters were also assigned to each observation point, which made it possible to analyze their relationships with soil structure, compaction indicators, and some soil properties [51].

### 2.3. Digital Soil Mapping (DSM)

#### 2.3.1. DSM Model, Parameters, and Evaluation

Random forest (RF) is a machine learning algorithm that produces multiple decision trees, randomly choosing features to make decisions when splitting nodes to create each tree. It then takes these randomized observations from each tree and averages them out to build a final model [55]. RF is one of the most used machine learning for DSM at broad-scales [8], and more details about its theory and applications can be found in [56]. A prediction model was generated from the SCC observation points using the *randomForest* package (version 4.7-1.1 [57]) in R. For model learning or training, we assigned its corresponding cluster class to each SCC observation point. Random forest algorithms have three main hyper-parameters that must be set before training. As relying on default parameters can introduce uncertainty, we performed a sensitivity analysis by testing the effect of increasing the hyper-parameters values on the out-of-bag error to determine the optimal settings for these parameters. Figure 2 shows that the default values offered by the *randomForest* function, namely 500, 3, and 1, for the number of trees, number of features (mtry), and node size, respectively, were suitable.

We used 10-fold cross-validation. Note that this 10-fold cross-validation is not comparable to the usual process of cross-validation, as the left-out folds of data were not real point “truth” observations but the STcell cluster classes the points belong to. A confusion matrix was built in order to compare point predictions with the original clusters to which they belong. The overall agreement (OA) corresponded to the sum of the points located in the diagonal of the matrix divided by the total number of points. The Kappa and Tau agreement indices were calculated as follows.



**Figure 2.** The estimation of the out-of-bag error (OOB) according to the variation in the hyper-parameters of the random forest model: (a) number of trees; (b) number of features (mtry); (c) node size.

The Kappa index is calculated as:

$$K = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

where  $P_o$  is the proportion of correctly classified samples, and  $P_e$  is the probability of random agreement. Kappa results can range from  $-1$  to  $+1$ .

The Tau was calculated to better account for unbalanced sample class distribution. The Tau index was calculated using the tauW function in the R *aqp* package (version 2.0.2. [58]). The Tau indicator measures the improvement of the classification over a random chance [59–61]. When the value of the Tau index is close to 1, it indicates an almost perfect agreement.

$$Tau = \frac{\theta_1 - \theta'_2}{1 - \theta'_2} \quad (2)$$

where

$$\theta_1 = \sum_{i=1}^r P_{ii} \quad (3)$$

$$\theta'_2 = \sum_{i=1}^r P_i \times P_{+i} \quad (4)$$

The importance of covariates was assessed using the Gini coefficient [62]. The average decrease in the Gini coefficient is a measure of how each variable contributes to the homogeneity of nodes and leaves in the resulting random forest. The higher the value of the Gini score, the more important the variable is in the model. As we had no clear indication of what the “true values” are, the Gini index coefficient is more relevant than accuracy-based indices.

The frequency of SCCs in each predicted point-based cluster was calculated and compared to the box plots of frequencies of the SCCs in STcell clusters. For each predicted cluster, box plots of selected quantitative data and/or indices classes were plotted and compared to the box plots obtained for STcell clusters. Maps of predicted clusters were displayed. For displaying maps, a mask of non-cropped soil was applied using (OSO [52]).

### 2.3.2. Covariates

The covariates (Table 1) included raster soil properties predicted from GlobalSoilMap-France [63] and the 5–15 cm layer for particle-size fractions (clay, silt, sand),  $pH_{water}$ , and SOC. From these SOC and clay rasters, we calculated a raster of the SOC/clay ratio [64,65] and a raster of the SOC/SOC<sub>exp</sub> ratio [66]. The SOC/SOC<sub>exp</sub> ratio was calculated using a slightly different method than proposed in [66], and this method is described in [51].

**Table 1.** List of covariates: acronyms, types, spatial resolution, unit, and reference.

| Name                   | Covariate  | Resolution        | Unit               | Reference   |
|------------------------|--|-------------------|--------------------|---|
| clay_GSM               | Clay   | 90 m              | g kg <sup>-1</sup> | GlobalSoilMap-France [63]   |
| silt_GSM               | Silt   | 90 m              | g kg <sup>-1</sup> | GlobalSoilMap-France [63]   |
| sand_GSM               | Sand   | 90 m              | g kg <sup>-1</sup> | GlobalSoilMap-France [63]   |
| SOC_GSM                | SOC  | 90 m              | g kg <sup>-1</sup> | GlobalSoilMap-France [63]   |
| ph_GSM                 | pH <sub>water</sub>                                  | 90 m              | pH                 | GlobalSoilMap-France [63]   |
| SOC/clay               | SOC/clay ratio                                       | 90 m              | -                  | calculated from<br>GlobalSoilMap-France [63]  |
| SOC/SOC <sub>exp</sub> | SOC/SOC <sub>exp</sub> ratio                         | 90 m              | -                  | calculated from<br>GlobalSoilMap-France [63]  |
| tmax                   | maximum annual temperature                           | 1 km <sup>2</sup> | °C                 | WorldClim<br><a href="http://www.worldclim.org/">http://www.worldclim.org/</a><br>(accessed on 3 July 2024) |
| typo                   | climate typology                                     | 90 m              | 8 classes          | [67]  |
| geol                   | Soil parent material                                 | 90 m              | 9 classes          | [68]  |
| IDPR                   | Rate of river network<br>development and persistence | 1:50 K            | -                  | [69]  |
| SRTM                   | Elevation  | 90 m              | m                  | SRTM (Shuttle Radar Topography<br>Mission), [70]  |
| OCS                    | Landscape Sentinel-2 2018                            | 10 m              | 23 classes         | Theia OSO Land Cover Map [52]   |

Climatic data included the maximum annual average temperature and a climate typology in 8 classes [67]. The lithology was derived from the dominant soil parent material stored in the 1:1,000,000 Geographical Soil Database of France (BDGSF; [68,71,72]). The soil parent material was classified into 9 classes. The IDPR index (Network Development and Persistence Index Values) was created by the French Institute for geology and mining research (BRGM) in 2014 [69]. It assesses the ability of the bedrock to induce water surface flow or infiltration. The elevation was given by the 90 m SRTM DEM (Shuttle Radar Topography Mission Digital Elevation Model [70]). Land cover (OCS for “OCcupation du Sol” in French) came from the Theia OSO Land Cover Map [52]. This product was compiled from a series of multi-temporal optical images with high spatial resolution (Sentinel-2, 2018). From this raster, we created a mask in order to keep only the cropped areas. These covariates were re-sampled using the R *terra* package (version 1.7-39) to have the same resolution (Figures S2 and S3). The reference resolution is that of SRTM. Covariates selection was refined using recursive feature elimination (RFE; *rfeControl* function of the *caret* package version 6.0-94 [73]). From the thirteen types of covariates described in Table 1, ten were retained by RFE: clay\_GSM, silt\_GSM, sand\_GSM, pH\_GSM, SOC/clay, SRTM, IDPR, tmax, SOC\_GSM, and SOC/SOC<sub>exp</sub>. Strictly speaking, as some qualitative covariates were split into many presence–absence covariates, RFE dropped the total number of covariates from 50 to 10. The maps of the covariates are available in Figures S2 and S3.

### 3. Results

#### 3.1. Random Forest Modeling Confusion Matrix

Table 2 shows the confusion matrix derived from 10-fold cross-validation. The lines are the STcell clusters from which point data were extracted. The columns correspond to the RF predictions of each individual point belonging to a cluster. Therefore, lines and columns do not have the same meaning and support. This consideration is important to keep in mind when interpreting the confusion matrix. Another important point to consider is the unbalanced number of points between clusters.

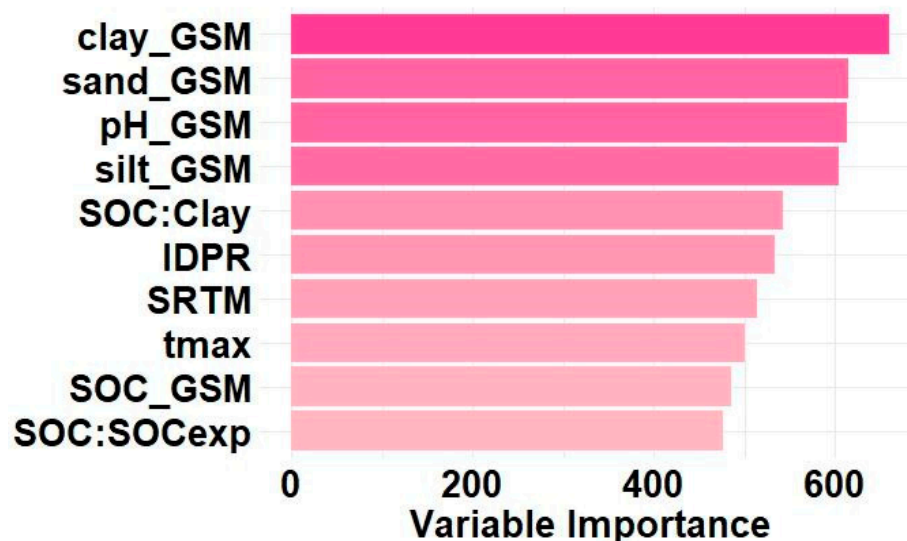
**Table 2.** Confusion matrix (OA: overall agreement). For each column and line, recall% is the ratio between the number of points belonging to the same cluster and the total number of points in the corresponding STcell and RF predicted clusters, respectively.

| STcell Cluster | Random Forest Cluster |      |      |      |      |      | Total | Recall%         |
|----------------|-----------------------|------|------|------|------|------|-------|-----------------|
|                | 1                     | 2    | 3    | 4    | 5    | 6    |       |                 |
| 1              | 309                   | 101  | 764  | 93   | 29   | 8    | 1304  | 23.7            |
| 2              | 92                    | 380  | 601  | 190  | 102  | 67   | 1432  | 26.5            |
| 3              | 338                   | 306  | 3091 | 294  | 120  | 35   | 4184  | 73.9            |
| 4              | 86                    | 228  | 775  | 264  | 78   | 37   | 1468  | 18.0            |
| 5              | 35                    | 138  | 280  | 85   | 154  | 32   | 724   | 21.3            |
| 6              | 11                    | 116  | 105  | 49   | 41   | 51   | 373   | 13.7            |
| <b>Total</b>   | 871                   | 1269 | 5616 | 975  | 524  | 230  | 9485  |                 |
| <b>Recall%</b> | 35.5                  | 29.9 | 55.0 | 27.1 | 29.4 | 22.2 |       | <b>OA = 45%</b> |

The overall agreement (OA) between the STcell clusters from which point data were extracted and the point predicted clusters using RF was 45%. One striking result is that most of the points (73.9%) belonging to STcells cluster 3 were allocated to cluster 3 by RF. Conversely, only 55.0% of the points predicted as cluster 3 came from the STcell cluster 3. Overall, RF predictions tended to smooth the differences between recall% values when compared to STcell clustering. The Kappa coefficient was 0.20, which was substantially lower than the OA, indicating that a large portion of the apparent classification agreement could be due to chance. The Tau index was 0.34, i.e., closest to the OA. Overall, OA, Kappa, and Tau were rather low if we compare them with the range of values usually observed in soil map assessments (e.g., [74,75]).

### 3.2. Covariates Importance

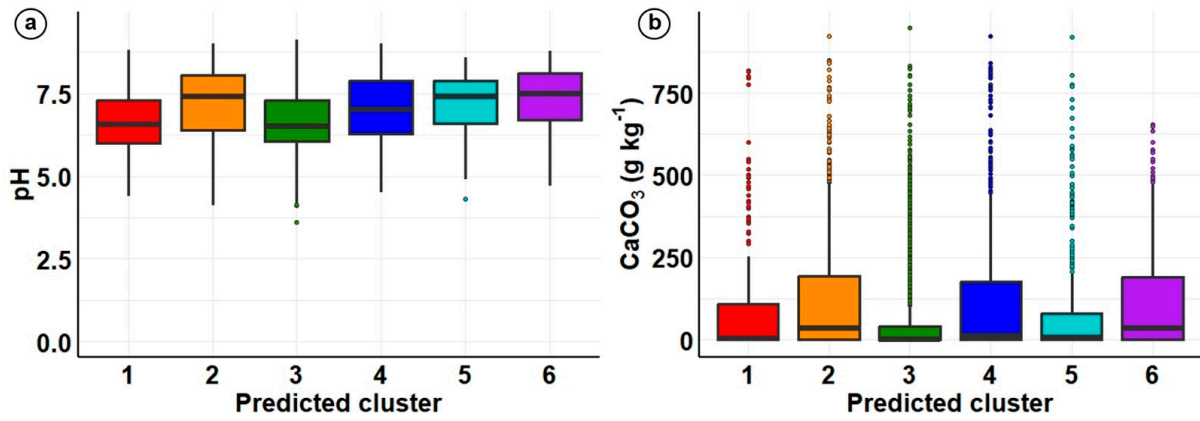
Figure 3 shows, for each variable, its importance in the classification of the data. The variables are presented in order of decreasing importance according to the Gini coefficient. Clay and sand were the most important covariates, followed by pH. As expected from previous results [51], the particle-size fractions were among the most important covariates. The SOC/clay index was in the fifth position, followed by non-soil covariates (IDPR, SRTM, tmax), SOC, and SOC/SOC<sub>exp</sub>.



**Figure 3.** Importance of covariates with average decrease in the Gini coefficient.



Considering the relative importance of pH, we wanted to know if the distributions of pH and CaCO<sub>3</sub> content were different among clusters. Figure 4 displays the distribution of these covariates among the clusters. Figure 4a shows that the distribution of pH values was slightly different among clusters, whereas Figure 4b shows that the third quartiles of CaCO<sub>3</sub> were substantially higher for clusters 2, 4, and 6 than for the other ones.

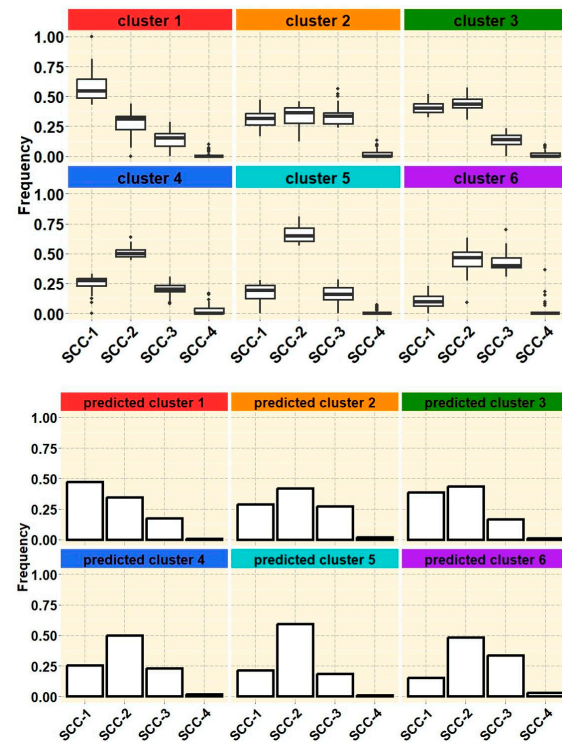


**Figure 4.** Box plots of (a) pH and (b) CaCO<sub>3</sub> content by predicted cluster. Colors correspond to different clusters.

### 3.3. Characteristics of the Predicted Clusters

#### 3.3.1. Frequency of Soil Compactness Classes and Selected Soil Properties per Predicted Cluster

Figure 5 compares the frequency of SCCs in each cluster after the clustering step with the frequency of SCCs in clusters predicted by DSM. The SCC frequencies of the DSM-predicted clusters were very similar to those observed by STcell clusters previously identified by Richer-de-Forges et al. [51].

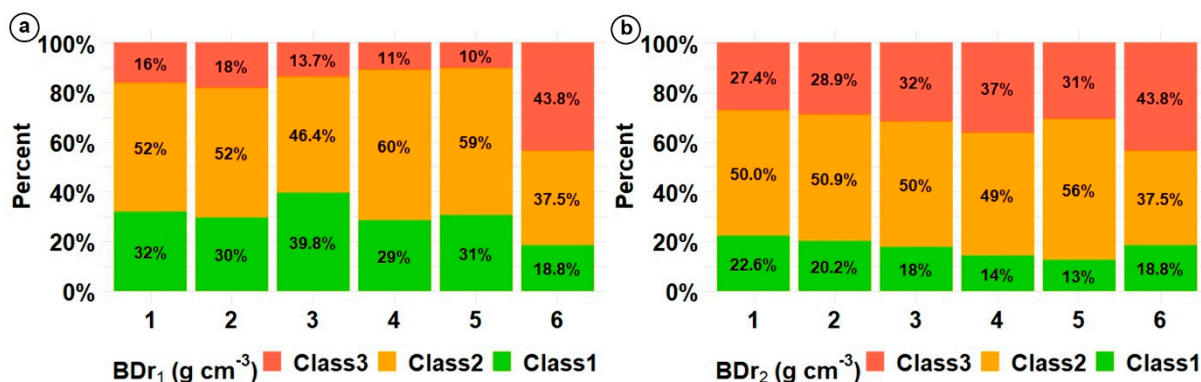


**Figure 5.** Frequency bar graph of soil compactness classes (SCCs). Top: for the six clusters (after Richer-de-Forges et al. [51]), black dots are outliers; bottom: for the six clusters predicted by DSM.

For predicted cluster 1 and cluster 5, the point predictions tended to slightly lower the largest SCC proportions compared to the highest SCC frequencies within STcells clusters. However, these most frequent SCCs remained the same (SCCS-1 for cluster 1 and SCC-2 for cluster 5). Another small difference appeared in cluster 6, where the frequency of SCC-1 was slightly lower than for the corresponding STcell cluster.

### 3.3.2. Relationships between Predicted Clusters and Indicators of Soil Compactness and Soil Properties

Figure 6 displays the relative proportion of two indicators of restrictive bulk density ( $BDr_1$  and  $BDr_2$ ) among the predicted clusters. These two indicators were defined by Jones [76] according to threshold values for limiting root depth, and both were classified into three classes. These threshold values are calculated from linear equations using clay content ( $BDr_1$ ) or (clay + silt) content ( $BDr_2$ ).  $BDr_1$  did not exhibit striking differences between clusters except for cluster 6, where the proportion of class3 was much larger compared to the other clusters.  $BDr_2$  showed a slight decrease in class1 proportion from cluster 1 to 4 with an increase in class3, whereas class2 proportion remained remarkably constant (ranging from 49 to 50.9%). Cluster 6 class proportions were exactly the same for  $BDr_1$  and  $BDr_2$ .



**Figure 6.**  $BDr_1$  (a) and  $BDr_2$  (b) classes by predicted clusters.  $BDr_1$  and  $BDr_2$  classes are calculated using clay ( $BDr_1$ ) or clay + silt content ( $BDr_2$ ). Class1: rooting not restricted. Class2: rooting moderately restricted. Class3: rooting highly restricted.

Figure 7 displays box plots of selected soil properties and threshold values for  $BDr_2$  for the DSM-predicted clusters.

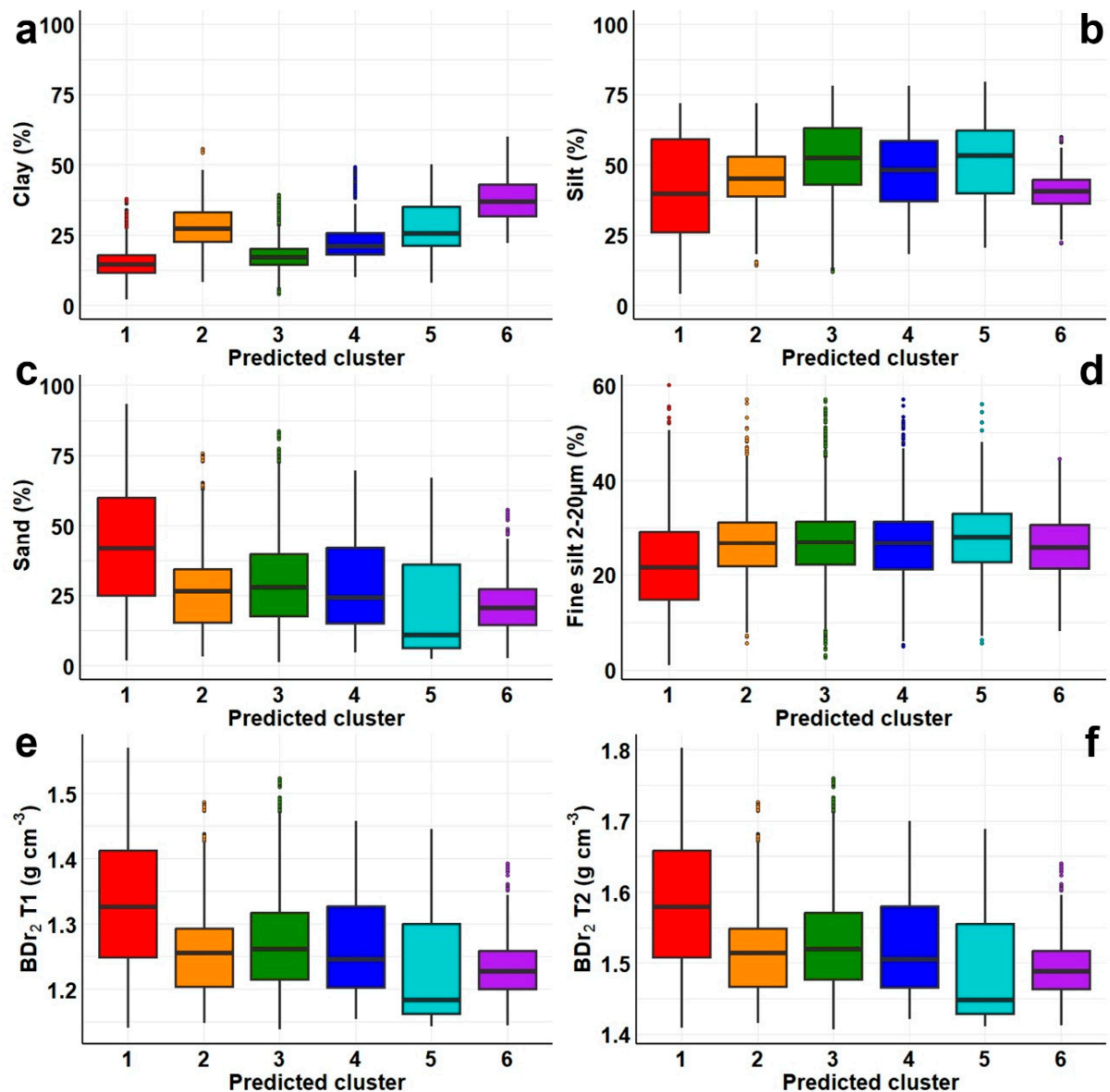
The box plots of particle-size fractions of the predicted clusters were consistent with the distribution of the initial clusters in the ST triangle (see Figure 6 in [51]). Cluster 1 was the sandiest, whereas cluster 6 was the clayest. Cluster 5 was characterized both by high values of silt and fine silt and the lowest median value of clay content.  $BDr_2$  thresholds were distributed in the same way. Note the high values for cluster 1 and the very low median values for cluster 5.

### 3.4. Mapping

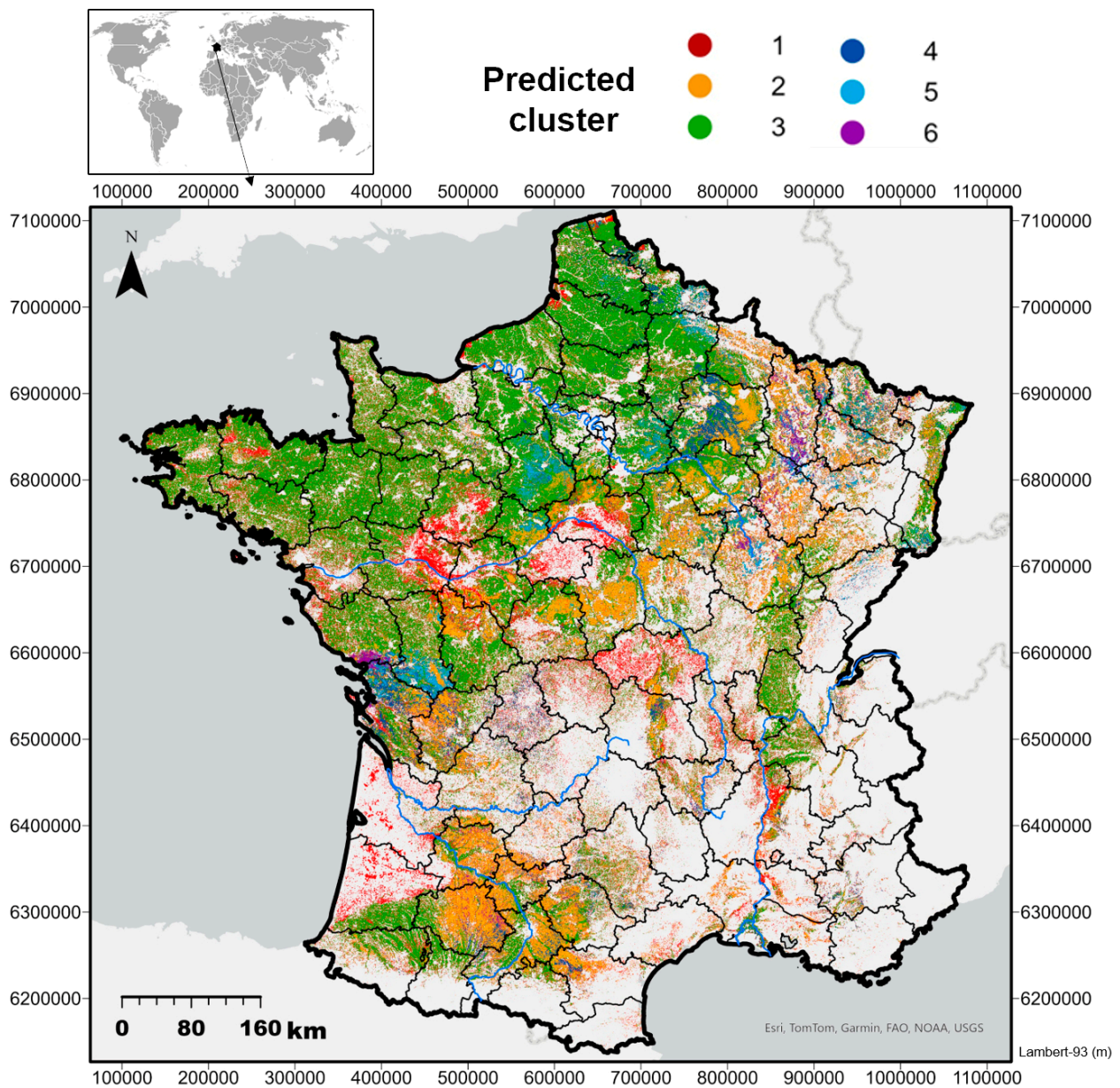
Figure 8 displays the cluster prediction map.

As expected, the predicted map of clusters mainly reflected the national distribution of topsoil texture as mapped by Mulder et al. [63]. All the large patches of sandy soils that were identified by the topsoil texture map are visible in the cluster map and correspond mainly to cluster 1. For example, in the southwest, the “Landes de Gascogne” are covered by podzols developed nearly in pure quartz sands [43,77]. Large patches of cluster 2 mainly correspond to silty clay textures, a large part of which is developed from calcareous parent materials. Cluster 3 dominates in large loamy and clayey silt to silty–clayey areas. Indeed, it dominates all of northwestern France, part of the southwest, and the Alsace plain (Eastern part of France). Cluster 4 is mainly encountered in somewhat poorly drained

conditions, such as Stagnosols in flat plains, depressed areas, former marshlands, and soils under flooding risk, such as some Colluvic cambisols or Fluvisols. As Cluster 4 is mainly dominated by temporal water-logging conditions, it explains why it is poorly correlated with texture. Cluster 5 appears to group very different soils, most of which are characterized by high contents of silt and fine silt (see Figure 7). Cluster 5 is mainly localized either on silt-rich soils such as Luvisols (Eutric) and Haplic Luvisols (Hypereutric) (see the zoom in Figure 9 and silt content in Figure S4) or on soils developed from soft calcareous materials (Cambisols (Calcaric)) with a lot of fine silt calcareous particles (see the zoom in Figure 10).



**Figure 7.** (a–f) Box plots of different topsoil properties per predicted cluster. For BDr<sub>2</sub>, T1 is the lower threshold separating classes 1 and 2, and T2 is the upper threshold separating classes 2 and 3. Colors and numbers on the x axis correspond to different predicted clusters.

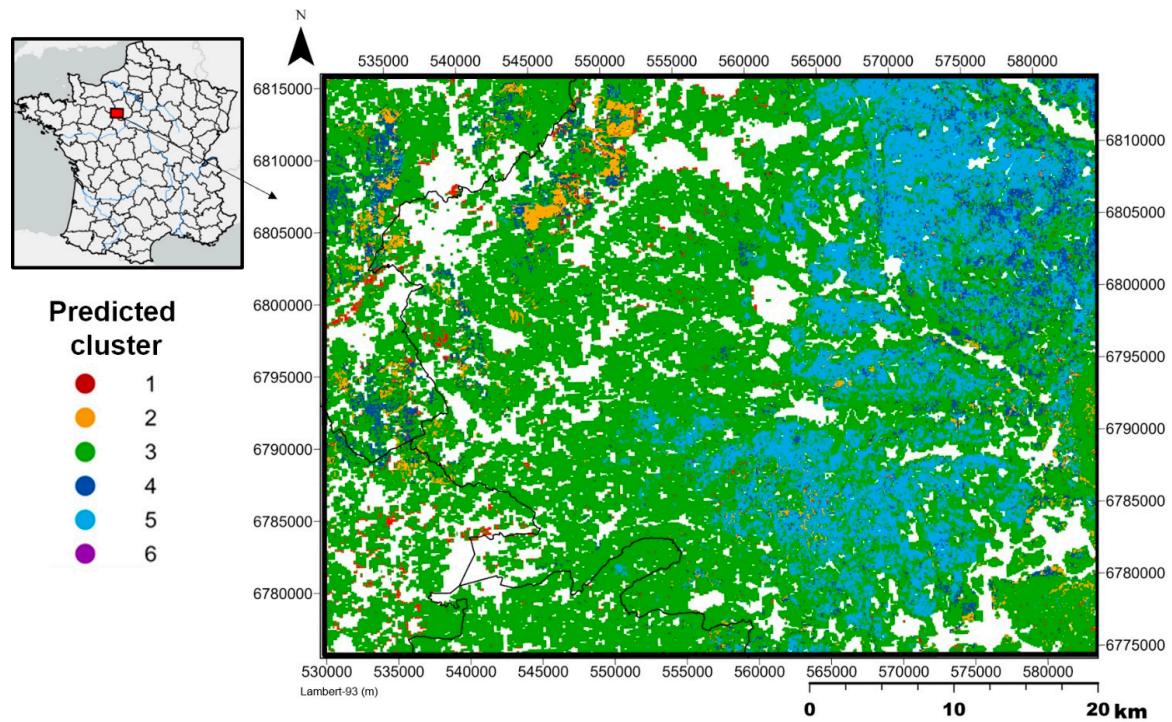


**Figure 8.** Map of predicted clusters. Areas left blank are non-cultivated areas and areas where covariates were not available.

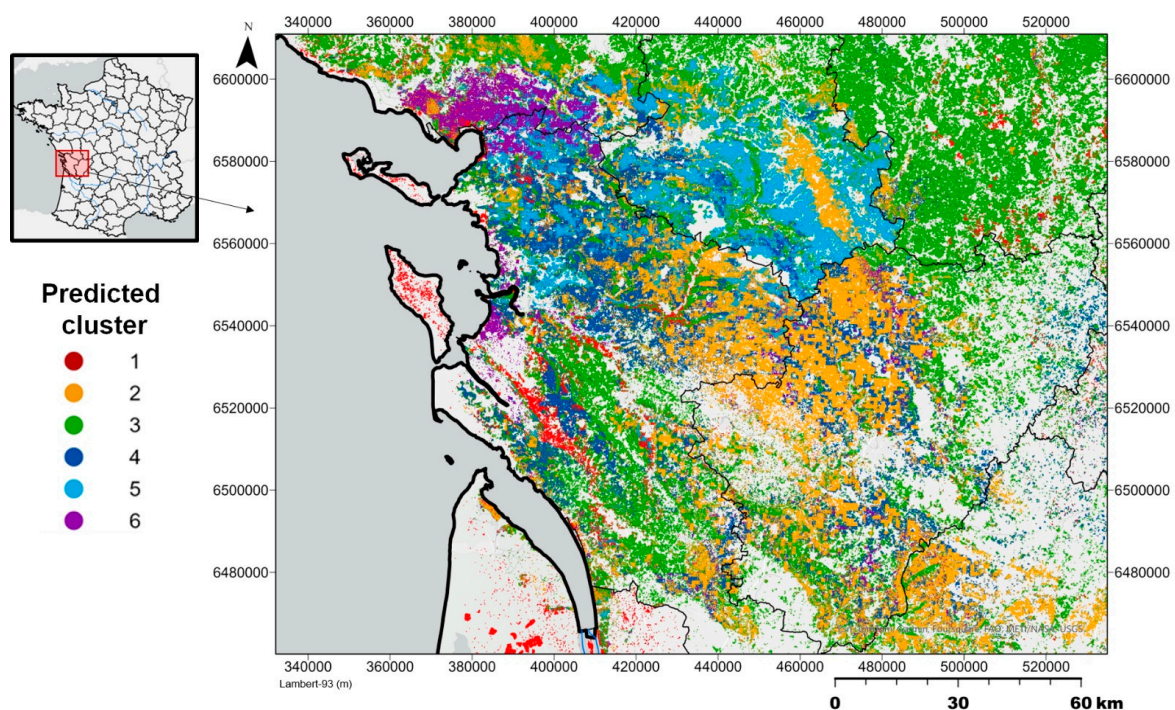
Note that the large cluster 5 patches from this map are characterized by rather extreme and homogeneous measured silt and fine silt contents when compared to the general box plot of cluster 5 silt and fine silt contents (Figure S4), as predicted by Mulder et al. [63] (see Figure 7).

Cluster 6 is mainly found in clayey and rather unstable soil structure situations. The largest patches of cluster 6 are located in the coastal marshlands in western France, where the main soils encountered are poorly drained clayey sodic soils (Solonetz, see the zoom in Figure 10) and poorly drained clayey soils (Epistagnic Regosols (Clayic)) in the Lorraine region (northeastern France).





**Figure 9.** Zoom of the map of predicted clusters on the soil map of Châteaudun [78]. Blue areas (cluster 5) are Luvisols Cambisols (Eutric) and Haplic Luvisols (Hypereutric) developed on loess deposits covering a calcareous rock and containing a large amount of silt and fine silt in the topsoil (see Figure S4). Green areas group a variety of soils mainly dominated by loamy topsoils (cluster 3).



**Figure 10.** Zoom of the map of predicted clusters in the Charentes region. Maritime sodic marshlands are in violet (cluster 6); light blue areas are mainly calcareous soils (Cambisols (Calcaric)) with high contents of silt and fine silt (cluster 5); dark blue areas correspond to non-sodic poorly drained soils (cluster 4); green areas group a variety of soils mainly dominated by loamy topsoils (cluster 3); orange areas are silty clays soils developed on calcareous hard rock (cluster 2); red areas are sandy soils (Cluster 1).



## 4. Discussion

### 4.1. Consistency of the Results and Main Outputs

We previously showed that clustering very simple information gathered in the field allowed clusters of topsoil behavior to be defined as compaction stress. These clusters were compared to several soil properties and indicators of soil structure, and spatial predictions were generated. When mapped and compared to the observed soils, the clusters displayed both consistent soil properties and geographical distribution with some of the existing national maps, such as topsoil texture, but also with more detailed maps. One striking result is that the national distribution of SCCs and soil properties remained very similar among clusters when defined by STcells and when further mapped at 90 m resolution.

The goal was to assess soil behavior to compaction; however, there was only one observation at one given time at each point, which required extrapolation techniques for spatial representation of soil behavior. We previously calculated distributions of SCCs among STcells and ran a clustering based only on the points and were able to interpret some of the clusters but with no geographic context. In this article, DSM improved the interpretation of the clusters, and in addition to confirming previous results, it identified new spatial factors (covariates) correlated to these behaviors to compaction.

Overall, we were able to map the six clusters that were defined in [51]. Most importantly, we were able to interpret controlling factors better. Cluster 1 is easily interpretable, as it corresponds to sandy soils that are the least sensitive to compaction and for which the compactness can be easily reversed by tillage. Cluster 2 is mainly composed of silty clays under hypereutric calcareous conditions; they can reach rather high compactness states, but they are reversible by tillage and/or by favorable pH for biological activity. Cluster 4 is mainly controlled by water-logging, which explains why it is not correlated to soil texture. It can be more or less easily compacted depending on the seasonal variability of soil moisture and when pressure to soil is applied.

The unexpected cluster 5 is characterized by large amounts of very fine but not deformable particles. The effect of compaction can lead to decreases in soil microporosity because high pressures can change the arrangement of these particles so that they reach their theoretical maximal BD and their minimum mean sizes of pores. However, as cluster 5 is dominated by one particle size, the theoretical maximal BD is limited when compared to soils with more heterogeneous particle sizes. Cluster 5 is also characterized by calcareous, eutric, or hypereutric conditions and rather well-drained situations. Cluster 6 is clearly linked to some clayey soils with low structural stability and are often water-saturated. Cluster 3 is still more challenging to interpret, as discussed in Sections 4.3.1 and 4.4.2.

### 4.2. Comparison with Other Studies

Most national and continental-scale maps of soil compactness or sensitivity to compaction used measurements and rather simple indicators such as PD or BDr; thresholds of porosity and/or PTFs, or pedotransfer rules (PTRs) to derive maps (e.g., [79–83]); or use models that require many properties that are most often derived from PTFs (e.g., [84–87]). Combining field observations and covariates to map soil behavior to compaction has almost never been conducted at the national scale to the best of our knowledge.

Jones et al. [79] estimated and mapped the inherent susceptibility of the subsoil to compaction using “relatively stable soil properties”, i.e., texture and packing density. They emphasized the limitations of the European soil and climatic databases that they used. The ST was estimated to be from only five expert-based classes, and the BD was derived from expert-based PTRs. Therefore, the estimates of PD they used, depending on both BD and clay content, were highly uncertain. In addition, Jones et al. used a PTR based on classes of ST and PD to derive four classes of susceptibility to compaction and used the dominant soil type of large soil map units to derive their map. Though this map could be considered as a starting point, it had many flaws, which were fully acknowledged and discussed by the authors.

Hollis et al. [81] further developed empirically derived PTFs and PTRs for predicting bulk density for European soils. Thanks to the growing number of available BD measurements in European topsoils [18], similar approaches were used to derive BD from PTRs and to provide pan-European maps of topsoil BD and PD [80,88]. All these maps exhibited a high uncertainty due to accumulated errors propagating from point BD measurements to inherent uncertainties in PTFs and PTRs models, as well as further uncertainties in mapping the input parameters of the PTFs and PTRs. In addition, they did not provide classes of topsoil behavior to compaction but rather classes of static inherent sensitivity to compaction. Indeed, they were most often used to derive stocks of elements [83,89] rather than to assess compaction risks in Europe.

Other attempts to map soil compaction risks included the combination of many spatial data and their use in mechanical models. Van den Akker et al. [90] mapped soil compaction risk in the Netherlands by combining several information sources (e.g., a land use database, an inventory of commonly used heavy machinery, a soil map at 1:50,000 scale and its associated database with descriptions of typical soil profiles), into a mechanical model [85].

D'Or and Destain [91] used precompression stress ( $P_c$ ) values in the subsoil, derived from measurements and PTFs, to map compaction risk in Wallonia depending on the vertical stress created by a wheel of agricultural or forest machinery. Schjønning and Lamandé [92] estimated soil strength from only three soil parameters: clay content, dry bulk density, and matric potential. By comparing mechanical stresses with soil strength, Lamandé et al. [93] evaluated and mapped the risk of subsoil compaction for specific machinery and soil conditions in Europe, using as soil inputs the 1:1,000,000 soil geographical base of Europe [71] and the SPADE8 database [94]. However, they pointed out several limitations linked to missing input data, BD values outside of the validity domain of the PTF predicting soil strength, and discrepancies at boundaries between countries. More dynamic approaches included spatially explicit modeling, such as using models incorporating crop type data, weather information, soil data, machinery information, crop simulation models, and various PTFs [86,87]. Though all these modeling approaches enabled substantial progress in compaction risk mapping, one should consider that they are heavily data-demanding, and their complexity conveys inherent error propagation.

### 4.3. Limitations

Some limitations of this study come from the rather low number of observations and time differences when the observations were made. We dealt with them in the study by Richer-de-Forges et al. [51]. In this section, we discuss only limitations that are specific to DSM.

#### 4.3.1. One Largely Dominant Cluster

Cluster 3 largely dominated, whereas cluster 6 was rather small. This has statistical implications for the robustness of our results. The Kappa, Tau, and OA indices were rather low (0.2, 0.34, and 0.45, respectively). Cohen [95] suggested the Kappa result be interpreted as follows: values  $\leq 0$  as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement. In this study, the influence of the unbalanced number of points is clearly shown by the increase from Kappa to Tau and to OA. Note that we are not comparing exactly the same classes. One classification comes from the hierarchical clustering of the STcells, and the other one comes from the point prediction of clusters using RF. Therefore, the Kappa and Tau are just indices of agreements between two different methods, none of which is considered the “truth”. Thus, they should not be interpreted as indicators of the prediction performance.

#### 4.3.2. Covariates

As is the case for most of the broad-scale DSM products, the ST fractions that we used as covariates were inherently smoothed, which may have contributed to capturing only

rather extreme situations in other clusters compared to cluster 3. However, improving predictions of ST fractions may be feasible by assembling DSM predictions realized at a more local scale (e.g., [96]). In addition, we did not search if some more relevant covariates could be used in the DSM approach. As cluster 4 is clearly related to water-logging, we could have introduced some DEM derivatives linked to water accumulation (e.g., curvatures, water indices, water flow accumulation, wetness index). These covariates could also have been efficient in better mapping cluster 6. Indeed, our set of covariates is still rather limited and, in particular, we did not add/test some sensor data that could come from remote sensing (RS; e.g., [97]). Some soil covariates could also have been useful, as shown by the large amount of fine silt in cluster 5. However, there is no topsoil fine silt content high-resolution DSM product for France yet. This lack of covariates advocates for producing DSM predictions of more detailed classes of soil particle-size distribution.

#### 4.4. Possible Improvements

##### 4.4.1. Refining the Digital Soil Mapping with New Covariates

As mentioned in Section 4.3.2, testing additional covariates could be useful for the DSM step. One first step could be to incorporate more relevant relief covariates. Producing and adding maps of more detailed classes of soil particle-size distribution could also help to map some clusters better. This is likely true for cluster 5 and fine silt. This might also be the case for cluster 1, as sandy soils may behave differently when they are mainly composed of coarse sand than when they are a mixture of different sand sizes. Overall, these maps could be useful not only for predicting compaction but also for other behaviors, such as slaking and crusting [98], run-off, and sensitivity to erosion [99]. They may also reveal some processes such as aeolian deposition [100,101], alluvial sedimentation [102], or pedogenesis [103].

In addition to relief-related covariates, RS could be used to characterize soil compactness. For example, roughness indices are often used in RS studies (e.g., [29]). They are often considered noise to eliminate or filter when they are not related to the attribute of interest (e.g., [104–106]). It could be worth testing some of these indices and their temporal changes from high spatial, temporal, and spectral resolution sensors. High-resolution RS products could also be used to derive intra-annual variability from soil moisture. The latter may also be useful in better predicting clusters 4 and 6.

##### 4.4.2. Looking Deeper in the Largest Cluster

We could have further explored cluster 3 composition and behavior by sub-clustering it using the same method that we used for the entire mainland France and then using DSM to refine sub-cluster interpretation. However, this objective is out of the scope of the present study, which was mainly to develop a proof-of-concept of how SCCs derived from on-field information can be used and how clustering and DSM are complementary for interpreting results.

#### 4.5. Perspectives and Prospects

##### 4.5.1. Adding other Information and Moving from Mapping Behaviors to Monitoring Risks

Field SCC data will potentially increase with time as the database is continuously growing. The main point will be to make relevant choices in filtering the data, whether our aim is to map what we could consider as rather stable and inherent soil properties effects or to monitor changes due to climate, crop rotations and agricultural practices, and mechanical pressure on the soil. For broad-scale mapping at time  $t$ , a preliminary large-scale delineation of “agrarian” regions could be a solution, as suggested by Bellón et al. [107]. For monitoring, we might consider that changes in crop rotations and some agricultural practices will soon be available from RS products (e.g., [108–111]).

Spatial information on mechanical pressure on soil is not easy to derive directly from RS. However, one may think that if some agricultural systems are delineable from RS

(e.g., [107,109,112]), the loads and periods of machinery interventions may be assessed and classified. Another solution is an expert-based delineation that could rely on agricultural statistics at a rather fine resolution when available (e.g., in France [113]). As suggested and reviewed by Peng et al. [114], RS products offer a high potential for mapping soil moisture at high resolution. This will enable the capture of seasonal effects and matching soil moisture status and pressures from agricultural machinery. As suggested by Khuwald et al. [87], another possible advance would be to gather high-resolution data from global navigation satellite systems received and recorded by modern farm vehicles to monitor traffic intensity and traffic loads.

#### 4.5.2. Moving to Deeper Horizons

Subsoil compaction is a severe problem mainly due to its persistence, and its effects may even be permanent [115–117]. Moreover, soil compaction has been a rising concern since the adoption of reduced tillage and/or no-till agricultural practices (e.g., [118–120]). Therefore, it would be worth testing our approach in subsoil. We already stressed the limitations that could arise from field observations in a previous study [51]. Obviously, we should explore the database to check if our method can be applied to subsoil.

#### 4.5.3. Better Communicating on the Map by Using Loss-Function Approaches and Spatially Targeting the Audience

The patterns of the map suggest that they could be used to prioritize public intervention to enhance or preserve soil structure. Relevant recommendations could be provided at the soil–scape scale. We classified clusters without giving weight to the consequences of misclassification. From a practical point of view, we could have used weights to take into account the losses due to misclassification [121,122]. The use of such loss functions could enable better communication to end-users. Whatever the case may be, communicating the impacts of compaction appears to be a prerequisite to the adoption of good practices for sustainable soil management. Indeed, the spatial distribution and patterns of most of the clusters (except cluster 3) suggest that efficient communication on good practices could be modulated and implemented at a broad soil–scape scale.

## 5. Conclusions

Incorporating cost-effective and multi-temporal field observation data into DSM is one of the ways forward to increase the number of observations and the relevance of moving from DSM to DSA to capture complex soil responses to various pressures. We previously used about 14,500 field observations of topsoil compactness classes (SCCs) collected by qualified soil surveyors in mainland France. We used the clustering of the SCCs to differentiate various topsoil behaviors from compaction. In this article, we successfully mapped these clusters using a DSM approach.

The main outputs of this study are as follows:

1. The SCC clustering, along with a choice of relevant covariates, enabled DSM to predict the location of clusters and their corresponding SCC proportions.
2. A map of soil behavior to compaction was produced. This map was interpretable from both geographical and topographic as well as pedological conditions.
3. DSM helped to discover patterns of clusters that provided additional explanations and controlling factors to soil behaviors to compaction.
4. The use of covariates enabled us to highlight the effects of some of the pedological and geomorphological characteristics that were not captured when using the soil database alone.
5. The patterns that we identified are operational for providing recommendations at the soil–scape scale. Therefore, the maps are DSA tools that help decision-making.
6. Such approaches, involving rescuing and gathering simple observations that are routinely performed by soil surveyors, are expected to have a great potential for digital soil assessment of complex soil properties and behaviors.

The main limitations of this study are linked to the lack of real ground “truth” on soil behavior to compaction and to the smoothing induced by DSM at a national scale. This smoothing is inherently linked to the covariates used and to the large area covered by this DSM.

Improvements could be obtained by gathering other independent information on soil compactness changes with time and loading and by running DSM at more local scales. Considering the large amounts of field observations that are available in some national databases, this study paves the way for their better use in revealing and mapping soil behaviors and their controlling factors.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/land13071014/s1>. Figure S1. Projection of the soil compactness class observation points in the texture triangle (after [51]). Figure S2. Maps of the covariates retained for digital soil mapping after recursive feature elimination. Figure S3. Maps of the covariates retained for digital soil mapping after recursive feature elimination (continue). Figure S4. Measured particle-size distribution of the cultivated topsoils encountered in the blue area of the soil map of Châteaudun [78]; in blue: mean values; in red: ranges covered after excluding the 5% lowest and the 5% highest values.

**Author Contributions:** Conceptualization, A.C.R.-d.-F. and D.A.; methodology, A.C.R.-d.-F. and D.A.; data curation, A.C.R.-d.-F.; writing—original draft preparation, A.C.R.-d.-F.; writing—review and editing, A.C.R.-d.-F., S.C., H.B., D.A., Z.L. and D.E.B.; visualization, A.C.R.-d.-F.; supervision, D.A. and H.B.; project administration, H.B. and D.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data may be available upon reasonable request. However, according to French legislation, the coordinates of the points cannot be shared.

**Acknowledgments:** A.C. Richer-de-Forges is the coordinator of the Centre d’Etudes Scientifiques Theia “Cartographie Numérique des Sols” supported by the CNES, France. D. Arrouays and H. Bourenanne are members of this Centre d’Etudes Scientifiques Theia “Cartographie Numérique des Sols”. Most of the French point soil data were collected or rescued thanks to funding from the national ‘Groupement d’Intérêt Scientifique Sol’ involving the French ministries in charge of agriculture and the environment, the French Agency of the ecological transition (ADEME—l’Agence de la transition écologique); the French National Research Institute for Agriculture, Food and Environment (INRAE—Institut national de recherche pour l’agriculture, l’alimentation et l’environnement); the French National Institute for Sustainable Development (IRD—l’Institut de recherche pour le développement); the French Geological Survey (BRGM—Service géologique national); and the French National Institute of Geographic and Forest Information (IGN—l’Institut national de l’information géographique et forestière). In most cases, these national fundings were complemented by French Regions or more local agencies and organizations. We thank all the people involved in sampling the sites and populating the French DoneSol database. We thank the people involved in maintaining and improving the structure and the accessibility of DoneSol.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Lagacherie, P.; McBratney, A.B.; Voltz, M. (Eds.) *Digital Soil Mapping: An Introductory Perspective*, 1st ed.; Developments in Soil Science; Elsevier: Amsterdam, The Netherlands; Boston, MA, USA, 2006; ISBN 978-0-444-52958-9.
2. McBratney, A.B.; Mendonça Santos, M.L.; Minasny, B. On Digital Soil Mapping. *Geoderma* **2003**, *117*, 3–52. [CrossRef]
3. Jenny, H. *Factors of Soil Formation: A System of Quantitative Pedology*; McGraw-Hill: New York, NY, USA, 1941; ISBN 978-0-486-68128-3.
4. Minasny, B.; McBratney, A.B. Digital Soil Mapping: A Brief History and Some Lessons. *Geoderma* **2016**, *264*, 301–311. [CrossRef]
5. Grunwald, S. Multi-Criteria Characterization of Recent Digital Soil Mapping and Modeling Approaches. *Geoderma* **2009**, *152*, 195–207. [CrossRef]
6. Biswas, A.; Zhang, Y. Sampling Designs for Validating Digital Soil Maps: A Review. *Pedosphere* **2018**, *28*, 1–15. [CrossRef]
7. Arrouays, D.; McBratney, A.; Bouma, J.; Libohova, Z.; Richer-de-Forges, A.C.; Morgan, C.L.S.; Roudier, P.; Poggio, L.; Mulder, V.L. Impressions of Digital Soil Maps: The Good, the Not so Good, and Making Them Ever Better. *Geoderma Reg.* **2020**, *20*, e00255. [CrossRef]



8. Chen, S.; Arrouays, D.; Leatitia Mulder, V.; Poggio, L.; Minasny, B.; Roudier, P.; Libohova, Z.; Lagacherie, P.; Shi, Z.; Hannam, J.; et al. Digital Mapping of GlobalSoilMap Soil Properties at a Broad Scale: A Review. *Geoderma* **2022**, *409*, 115567. [[CrossRef](#)]
9. Guevara, M.; Olmedo, G.F.; Stell, E.; Yigini, Y.; Aguilar Duarte, Y.; Arellano Hernández, C.; Arévalo, G.E.; Arroyo-Cruz, C.E.; Bolivar, A.; Bunning, S.; et al. No Silver Bullet for Digital Soil Mapping: Country-Specific Soil Organic Carbon Estimates across Latin America. *SOIL* **2018**, *4*, 173–193. [[CrossRef](#)]
10. Wadoux, A.M.J.-C.; Minasny, B.; McBratney, A.B. Machine Learning for Digital Soil Mapping: Applications, Challenges and Suggested Solutions. *Earth-Sci. Rev.* **2020**, *210*, 103359. [[CrossRef](#)]
11. Searle, R.; McBratney, A.; Grundy, M.; Kidd, D.; Malone, B.; Arrouays, D.; Stockman, U.; Zund, P.; Wilson, P.; Wilford, J.; et al. Digital Soil Mapping and Assessment for Australia and beyond: A Propitious Future. *Geoderma Reg.* **2021**, *24*, e00359. [[CrossRef](#)]
12. Arrouays, D.; Leenaars, J.G.B.; Richer-de-Forges, A.C.; Adhikari, K.; Ballabio, C.; Greve, M.; Grundy, M.; Guerrero, E.; Hempel, J.; Hengl, T.; et al. Soil Legacy Data Rescue via GlobalSoilMap and Other International and National Initiatives. *GeoResJ* **2017**, *14*, 1–19. [[CrossRef](#)]
13. Arrouays, D.; Richer-de-Forges, A.C.; Héliès, F.; Mulder, V.L.; Saby, N.P.A.; Chen, S.; Martin, M.P.; Román Dobarco, M.; Follain, S.; Jolivet, C.; et al. Impacts of National Scale Digital Soil Mapping Programs in France. *Geoderma Reg.* **2020**, *23*, e00337. [[CrossRef](#)]
14. Hengl, T.; Mendes de Jesus, J.; Heuvelink, G.B.M.; Ruiperez Gonzalez, M.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global Gridded Soil Information Based on Machine Learning. *PLoS ONE* **2017**, *12*, e0169748. [[CrossRef](#)]
15. Hengl, T.; Leenaars, J.G.B.; Shepherd, K.D.; Walsh, M.G.; Heuvelink, G.B.M.; Mamo, T.; Tilahun, H.; Berkhout, E.; Cooper, M.; Fegraus, E.; et al. Soil Nutrient Maps of Sub-Saharan Africa: Assessment of Soil Nutrient Content at 250 m Spatial Resolution Using Machine Learning. *Nutr. Cycl. Agroecosyst.* **2017**, *109*, 77–102. [[CrossRef](#)]
16. Poggio, L.; de Sousa, L.M.; Batjes, N.H.; Heuvelink, G.B.M.; Kempen, B.; Ribeiro, E.; Rossiter, D. SoilGrids 2.0: Producing Soil Information for the Globe with Quantified Spatial Uncertainty. *SOIL* **2021**, *7*, 217–240. [[CrossRef](#)]
17. Grundy, M.J.; Rossel, R.A.V.; Searle, R.D.; Wilson, P.L.; Chen, C.; Gregory, L.J. Soil and Landscape Grid of Australia. *Soil Res.* **2015**, *53*, 835–844. [[CrossRef](#)]
18. Orgiazzi, A.; Ballabio, C.; Panagos, P.; Jones, A.; Fernández-Ugalde, O. LUCAS Soil, the Largest Expandable Soil Dataset for Europe: A Review. *Eur. J. Soil Sci.* **2018**, *69*, 140–153. [[CrossRef](#)]
19. Chaney, N.W.; Minasny, B.; Herman, J.D.; Nauman, T.W.; Brungard, C.W.; Morgan, C.L.S.; McBratney, A.B.; Wood, E.F.; Yimam, Y. POLARIS Soil Properties: 30-m Probabilistic Maps of Soil Properties Over the Contiguous United States. *Water Resour. Res.* **2019**, *55*, 2916–2938. [[CrossRef](#)]
20. Liang, Z.; Chen, S.; Yang, Y.; Zhou, Y.; Shi, Z. High-Resolution Three-Dimensional Mapping of Soil Organic Carbon in China: Effects of SoilGrids Products on National Modeling. *Sci. Total Environ.* **2019**, *685*, 480–489. [[CrossRef](#)]
21. Batjes, N.H.; Ribeiro, E.; Van Oostrum, A. Standardised Soil Profile Data to Support Global Mapping and Modelling (WoSIS Snapshot 2019). *Earth Syst. Sci. Data* **2020**, *12*, 299–320. [[CrossRef](#)]
22. Suleymanov, A.; Richer-de-Forges, A.C.; Saby, N.P.A.; Arrouays, D.; Martin, M.P.; Bispo, A. National-Scale Digital Soil Mapping Performances Are Related to Covariates and Sampling Density: Lessons from France. *Geoderma Reg.* **2024**, *37*, e00801. [[CrossRef](#)]
23. Grimley, D.A.; Arruda, N.K.; Bramstedt, M.W. Using Magnetic Susceptibility to Facilitate More Rapid, Reproducible and Precise Delineation of Hydric Soils in the Midwestern USA. *Catena* **2004**, *58*, 183–213. [[CrossRef](#)]
24. Viscarra Rossel, R.A.; Cattle, S.R.; Ortega, A.; Fouad, Y. In Situ Measurements of Soil Colour, Mineral Composition and Clay Content by Vis–NIR Spectroscopy. *Geoderma* **2009**, *150*, 253–266. [[CrossRef](#)]
25. Soriano-Disla, J.M.; Janik, L.J.; Viscarra Rossel, R.A.; Macdonald, L.M.; McLaughlin, M.J. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Appl. Spectrosc. Rev.* **2014**, *49*, 139–186. [[CrossRef](#)]
26. Asgari, N.; Ayoubi, S.; Dematté, J.A.M. Soil Drainage Assessment by Magnetic Susceptibility Measures in Western Iran. *Geoderma Reg.* **2018**, *13*, 35–42. [[CrossRef](#)]
27. Gholizadeh, A.; Saberioon, M.; Viscarra Rossel, R.A.; Boruvka, L.; Klement, A. Spectroscopic Measurements and Imaging of Soil Colour for Field Scale Estimation of Soil Organic Carbon. *Geoderma* **2020**, *357*, 113972. [[CrossRef](#)]
28. Lagacherie, P.; Arrouays, D.; Bourennane, H.; Gomez, C.; Nkuba-Kasanda, L. Analysing the Impact of Soil Spatial Sampling on the Performances of Digital Soil Mapping Models and Their Evaluation: A Numerical Experiment on Quantile Random Forest Using Clay Contents Obtained from Vis-NIR-SWIR Hyperspectral Imagery. *Geoderma* **2020**, *375*, 114503. [[CrossRef](#)]
29. Richer-de-Forges, A.C.; Chen, Q.; Baghdadi, N.; Chen, S.; Gomez, C.; Jacquemoud, S.; Martelet, G.; Mulder, V.L.; Urbina-Salazar, D.; Vaudour, E.; et al. Remote Sensing Data for Digital Soil Mapping in French Research—A Review. *Remote Sens.* **2023**, *15*, 3070. [[CrossRef](#)]
30. Collard, F.; Kempen, B.; Heuvelink, G.B.M.; Saby, N.P.A.; Richer-de-Forges, A.C.; Lehmann, S.; Nehlig, P.; Arrouays, D. Refining a Reconnaissance Soil Map by Calibrating Regression Models with Data from the Same Map (Normandy, France). *Geoderma Reg.* **2014**, *1*, 21–30. [[CrossRef](#)]
31. Odgers, N.P.; Sun, W.; McBratney, A.B.; Minasny, B.; Clifford, D. Disaggregating and Harmonising Soil Map Units through Resampled Classification Trees. *Geoderma* **2014**, *214–215*, 91–100. [[CrossRef](#)]
32. Nauman, T.W.; Thompson, J.A. Semi-Automated Disaggregation of Conventional Soil Maps Using Knowledge Driven Data Mining and Classification Trees. *Geoderma* **2014**, *213*, 385–399. [[CrossRef](#)]

33. Holmes, K.W.; Griffin, E.A.; Odgers, N.P. Large-Area Spatial Disaggregation of a Mosaic of Conventional Soil Maps: Evaluation over Western Australia. *Soil Res.* **2015**, *53*, 865–880. [[CrossRef](#)]
34. Møller, A.B.; Malone, B.; Odgers, N.P.; Beucher, A.; Iversen, B.V.; Greve, M.H.; Minasny, B. Improved Disaggregation of Conventional Soil Maps. *Geoderma* **2019**, *341*, 148–160. [[CrossRef](#)]
35. Lázaro-López, A.; González-SanJosé, M.L.; Gómez-Miguel, V. Disaggregation of Conventional Soil Maps: A Review. *Soil Res.* **2021**, *59*, 747–766. [[CrossRef](#)]
36. Terribile, F.; Coppola, A.; Langella, G.; Martina, M.; Basile, A. Potential and Limitations of Using Soil Mapping Information to Understand Landscape Hydrology. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3895–3933. [[CrossRef](#)]
37. Terribile, F.; Agrillo, A.; Bonfante, A.; Buscemi, G.; Colandrea, M.; D’Antonio, A.; De Mascellis, R.; De Michele, C.; Langella, G.; Manna, P.; et al. A Web-Based Spatial Decision Supporting System for Land Management and Soil Conservation. *Solid Earth* **2015**, *6*, 903–928. [[CrossRef](#)]
38. Bonfante, A.; Bouma, J. The Role of Soil Series in Quantitative Land Evaluation When Expressing Effects of Climate Change and Crop Breeding on Future Land Use. *Geoderma* **2015**, *259–260*, 187–195. [[CrossRef](#)]
39. Bonfante, A.; Terribile, F.; Bouma, J. Refining Physical Aspects of Soil Quality and Soil Health When Exploring the Effects of Soil Degradation and Climate Change on Biomass Production: An Italian Case Study. *SOIL* **2019**, *5*, 1–14. [[CrossRef](#)]
40. Nield, S.J.; Boettinger, J.L.; Ramsey, R.D. Digitally Mapping Gypsic and Natric Soil Areas Using Landsat ETM Data. *Soil Sci. Soc. Am. J.* **2007**, *71*, 245–252. [[CrossRef](#)]
41. Jafari, A.; Finke, P.A.; Vande Wauw, J.; Ayoubi, S.; Khademi, H. Spatial Prediction of USDA-Great Soil Groups in the Arid Zarand Region, Iran: Comparing Logistic Regression Approaches to Predict Diagnostic Horizons and Soil Types. *Eur. J. Soil Sci.* **2012**, *63*, 284–298. [[CrossRef](#)]
42. Evans, D.M.; Hartemink, A.E. Digital Soil Mapping of a Red Clay Subsoil Covered by Loess. *Geoderma* **2014**, *230–231*, 296–304. [[CrossRef](#)]
43. Richer-de-Forges, A.C.; Saby, N.P.A.; Mulder, V.L.; Laroche, B.; Arrouays, D. Probability Mapping of Iron Pan Presence in Sandy Podzols in South-West France, Using Digital Soil Mapping. *Geoderma Reg.* **2017**, *9*, 39–46. [[CrossRef](#)]
44. Zhi, J.; Zhang, G.; Yang, F.; Yang, R.; Liu, F.; Song, X.; Zhao, Y.; Li, D. Predicting Mattic Epipedons in the Northeastern Qinghai-Tibetan Plateau Using Random Forest. *Geoderma Reg.* **2017**, *10*, 1–10. [[CrossRef](#)]
45. Zhi, J.; Zhou, Z.; Cao, X. Exploring the Determinants and Distribution Patterns of Soil Mattic Horizon Thickness in a Typical Alpine Environment Using Boosted Regression Trees. *Ecol. Indic.* **2021**, *133*, 108373. [[CrossRef](#)]
46. Jiang, Z.-D.; Wang, Q.-B.; Libohova, Z.; Adhikari, K.; Brye, K.R.; Sun, Z.-X.; Sun, F.-J.; Jiang, Y.-Y.; Owens, P.R. Fe-Mn Concentrations in Upland Loess Soils in Mid-Continental North America: A Step towards Dynamic Soil Survey. *Catena* **2021**, *202*, 105273. [[CrossRef](#)]
47. Assami, T.; Chenchouni, H.; Hadj-Miloud, S. Mapping the Petrogypsic Horizon Occurrence Probability in the Sahara Desert Using Predictive Models. *Eurasian Soil Sc.* **2024**, *57*, 551–561. [[CrossRef](#)]
48. FAO. *Guidelines for Soil Description*, 4th ed.; Food and Agriculture Organization of the United Nations: Rome, Italy, 2006; ISBN 978-92-5-105521-2.
49. Baize, D.; Jabiol, B. *Guide pour la Description des sols Nouvelle Édition*; Editions Quae: Paris, France, 2012; ISBN 978-2-7592-1035-0.
50. Schoneberger, P.J.; Wysocki, D.A.; Benham, E.C.; Soil Survey Staff. *Field Book for Describing and Sampling Soils. Version 3.0*; U.S. Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center: Lincoln, NE, USA, 2012.
51. Richer-de-Forges, A.C.; Arrouays, D.; Libohova, Z.; Chen, S.; Beaudette, D.E.; Bourennane, H. Revealing Topsoil Behavior to Compaction from Mining Field Observations. *Land* **2024**, *13*, 909. [[CrossRef](#)]
52. Thierion, V.; Vincent, A.; Valero, S. Theia OSO Land Cover Map 2021 (Version 1) [Data Set]. 2022. Zenodo. Available online: <https://commons.datacite.org/doi.org/10.5281/zenodo.6538909> (accessed on 3 July 2024).
53. Jamagne, M. *Grands Paysages Pédologiques de France; Collection Synthèses*, 1st ed.; Éditions Quae: Versailles, France, 2011; ISBN 978-2-7592-1036-7.
54. Tibshirani, R.; Walther, G.; Hastie, T. Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2001**, *63*, 411–423. [[CrossRef](#)]
55. Ishwaran, H.; Kogalur, U.B. Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC), R Package Version 3.2.0. 2023. Available online: <https://cran.r-project.org/web/packages/randomForestSRC/randomForestSRC.pdf> (accessed on 3 July 2024).
56. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
57. Liaw, A.; Weiner, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22. Available online: <https://CRAN.R-project.org/doc/Rnews/> (accessed on 3 July 2024).
58. Beaudette, D.E.; Roudier, P.; Brown, A. Package R aqp: Algorithms for Quantitative Pedology 2023. Version 2.0. Available online: <https://CRAN.R-project.org/package=aqp> (accessed on 3 July 2024).
59. Ma, Z.; Redmond, R.L. Tau Coefficients for Accuracy Assessment of Classification of Remote Sensing Data. *Photogramm. Photogramm. Eng. Remote Sens.* **1995**, *61*, 435–439.
60. Rossiter, D. Technical Note: Statistical Methods for Accuracy Assessment of Classified Thematic Maps. 2004. 46p. Available online: [https://www.researchgate.net/publication/228802780\\_Technical\\_Note\\_Statistical\\_methods\\_for\\_accuracy\\_assesment\\_of\\_classified\\_thematic\\_maps](https://www.researchgate.net/publication/228802780_Technical_Note_Statistical_methods_for_accuracy_assesment_of_classified_thematic_maps) (accessed on 3 July 2024).

61. Rossiter, D.G.; Zeng, R.; Zhang, G.-L. Accounting for Taxonomic Distance in Accuracy Assessment of Soil Class Predictions. *Geoderma* **2017**, *292*, 118–127. [CrossRef]
62. Gini, C. Measurement of Inequality of Incomes. *Econ. J.* **1921**, *31*, 124–125. [CrossRef]
63. Mulder, V.L.; Lacoste, M.; Richer-de-Forges, A.C.; Arrouays, D. GlobalSoilMap France: High-Resolution Spatial Modelling the Soils of France up to Two Meter Depth. *Sci. Total Environ.* **2016**, *573*, 1352–1369. [CrossRef] [PubMed]
64. Johannes, A.; Matter, A.; Schulin, R.; Weisskopf, P.; Baveye, P.C.; Boivin, P. Optimal Organic Carbon Values for Soil Structure Quality of Arable Soils. Does Clay Content Matter? *Geoderma* **2017**, *302*, 14–21. [CrossRef]
65. Prout, J.M.; Shepherd, K.D.; McGrath, S.P.; Kirk, G.J.D.; Haefele, S.M. What Is a Good Level of Soil Organic Matter? An Index Based on Organic Carbon to Clay Ratio. *Eur. J. Soil Sci.* **2021**, *72*, 2493–2503. [CrossRef]
66. Poeplau, C.; Don, A. A Simple Soil Organic Carbon Level Metric beyond the Organic Carbon-to-clay Ratio. *Soil Use Manag.* **2023**, *39*, 1057–1067. [CrossRef]
67. Joly, D.; Brossard, T.; Cardot, H.; Cavailhes, J.; Hilal, M.; Wavresky, P. Les Types de Climats En France, Une Construction Spatiale. *Cybergeo* **2010**. [CrossRef]
68. INRAE. *Base de Données Géographiques des sols de France à 1/1 000 000 Version 3.2.8.0, 10/09/1998*; INRAE: Paris, France, 2018. [CrossRef]
69. BRGM-Centre Scientifique et Technique Indice de Développement et de Persistance des Réseaux (IDPR) 2014. Available online: <https://infoterre.brgm.fr/actualites/mise-jour-couche-idpr-indice-developpement-persistance-reseaux> (accessed on 3 July 2024).
70. OpenTopography. Shuttle Radar Topography Mission (SRTM) Global. 2013. [CrossRef]
71. King, D.; Burrell, A.; Daroussin, J.; Le Bas, C.; Tavenier, R.; Van Ranst, E. *European Land Information Systems for Agro-Environmental Monitoring*; Office for Official Publications of the European Communities: Luxembourg, 1995; p. 284.
72. Le Bas, C.; Barthes, L.; Boutefroy, I.; Fort, J.-L.; Scheurer, O.; Darracq, S.; Lacassin, J.-C.; Sauter, J.; Schwartz, C. Utilisation des données sols d'I.G.C.S. en France: Un état des lieux. *Etude Et Gest. Des Sols* **2004**, *11*, 299–305.
73. Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
74. Rossiter, D.G. *Assessing the Thematic Accuracy of Area-Class Soil Maps*. 2001. Available online: <https://api.semanticscholar.org/CorpusID:14557075> (accessed on 3 July 2024).
75. Fantappiè, M.; L'Abate, G.; Schillaci, C.; Costantini, E.A.C. Digital Soil Mapping of Italy to Map Derived Soil Profiles with Neural Networks. *Geoderma Reg.* **2023**, *32*, e00619. [CrossRef]
76. Jones, C.A. Effect of Soil Texture on Critical Bulk Densities for Root Growth<sup>1</sup>. *Soil Sci. Soc. Am. J.* **1983**, *47*, 1208–1211. [CrossRef]
77. Augusto, L.; Bakker, M.R.; Morel, C.; Meredieu, C.; Trichet, P.; Badeau, V.; Arrouays, D.; Plassard, C.; Achat, D.L.; Gallet-Budynek, A.; et al. Is 'Grey Literature' a Reliable Source of Data to Characterize Soils at the Scale of a Region? A Case Study in a Maritime Pine Forest in Southwestern France. *Eur. J. Soil Sci.* **2010**, *61*, 807–822. [CrossRef]
78. Isambert, M. *Carte des sols de Châteaudun: 1-9 au 1/100 000 1984*; Editions QUAE: Versailles, France, 1984.
79. Jones, R.J.A.; Spoor, G.; Thomasson, A.J. Vulnerability of Subsoils in Europe to Compaction: A Preliminary Analysis. *Soil Tillage Res.* **2003**, *73*, 131–143. [CrossRef]
80. Panagos, P.; De Rosa, D.; Liakos, L.; Labouyrie, M.; Borrelli, P.; Ballabio, C. Soil Bulk Density Assessment in Europe. *Agric. Ecosyst. Environ.* **2024**, *364*, 108907. [CrossRef]
81. Hollis, J.M.; Hannam, J.; Bellamy, P.H. Empirically-derived Pedotransfer Functions for Predicting Bulk Density in European Soils. *Eur. J. Soil Sci.* **2012**, *63*, 96–109. [CrossRef]
82. Brus, D.J.; Van Den Akker, J.J.H. How Serious a Problem Is Subsoil Compaction in the Netherlands? A Survey Based on Probability Sampling. *SOIL* **2018**, *4*, 37–45. [CrossRef]
83. De Rosa, D.; Ballabio, C.; Lugato, E.; Fasiolo, M.; Jones, A.; Panagos, P. Soil Organic Carbon Stocks in European Croplands and Grasslands: How Much Have We Lost in the Past Decade? *Glob. Chang. Biol.* **2024**, *30*, e16992. [CrossRef]
84. Defosse, P.; Richard, G. Models of Soil Compaction Due to Traffic and Their Evaluation. *Soil Tillage Res.* **2002**, *67*, 41–64. [CrossRef]
85. Van Den Akker, J.J.H. SOCOMO: A Soil Compaction Model to Calculate Soil Stresses and the Subsoil Carrying Capacity. *Soil Tillage Res.* **2004**, *79*, 113–127. [CrossRef]
86. Kuhwald, M.; Dörnhöfer, K.; Oppelt, N.; Duttman, R. Spatially Explicit Soil Compaction Risk Assessment of Arable Soils at Regional Scale: The SaSCiA-Model. *Sustainability* **2018**, *10*, 1618. [CrossRef]
87. Kuhwald, M.; Kuhwald, K.; Duttman, R. Spatio-Temporal High-Resolution Subsoil Compaction Risk Assessment for a 5-Years Crop Rotation at Regional Scale. *Front. Environ. Sci.* **2022**, *10*, 823030. [CrossRef]
88. Ballabio, C.; Panagos, P.; Monatanarella, L. Mapping Topsoil Physical Properties at European Scale Using the LUCAS Database. *Geoderma* **2016**, *261*, 110–123. [CrossRef]
89. Panagos, P.; Köningner, J.; Ballabio, C.; Liakos, L.; Muntwyler, A.; Borrelli, P.; Lugato, E. Improving the Phosphorus Budget of European Agricultural Soils. *Sci. Total Environ.* **2022**, *853*, 158706. [CrossRef]
90. Van Den Akker, J.J.H.; de Vries, F.; Vermeulen, G.D.; Hack-ten Broeke, M.J.D.; Schouten, T. *Risico Op Ondergrondverdichting in Het Landelijk Gebied in Kaart*; Alterra, Wageningen University and Research Centre: Wageningen, The Netherlands, 2013.
91. D'Or, D.; Destain, M.-F. Toward a Tool Aimed to Quantify Soil Compaction Risks at a Regional Scale: Application to Wallonia (Belgium). *Soil Tillage Res.* **2014**, *144*, 53–71. [CrossRef]



92. Schjøning, P.; Lamandé, M. Models for Prediction of Soil Precompression Stress from Readily Available Soil Properties. *Geoderma* **2018**, *320*, 115–125. [[CrossRef](#)]
93. Lamandé, M.; Greve, M.H.; Schjøning, P. Risk Assessment of Soil Compaction in Europe—Rubber Tracks or Wheels on Machinery. *CATENA* **2018**, *167*, 353–362. [[CrossRef](#)]
94. Koue, P.M.; Balstrøm, T.; Breuning-Madsen, H. *Update of the European Soil Analytical Database (SPADE-1) to Version SPADE8*; EU-Joint Research Centre: Ispra, Italy, 2008.
95. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [[CrossRef](#)]
96. Brungard, C.; Nauman, T.; Duniway, M.; Veblen, K.; Nehring, K.; White, D.; Salley, S.; Anchang, J. Regional Ensemble Modeling Reduces Uncertainty for Digital Soil Mapping. *Geoderma* **2021**, *397*, 114998. [[CrossRef](#)]
97. De Souza, I.R.M.; Sano, E.E.; De Lima, R.P.; Da Silva, A.R. A Remote Sensing Approach to Estimate the Load Bearing Capacity of Soil. *Inf. Process. Agric.* **2024**, *11*, 109–116. [[CrossRef](#)]
98. Valentin, C.; Bresson, L.-M. Morphology, Genesis and Classification of Surface Crusts in Loamy and Sandy Soils. *Geoderma* **1992**, *55*, 225–245. [[CrossRef](#)]
99. Le Bissonnais, Y.; Montier, C.; Jamagne, M.; Daroussin, J.; King, D. Mapping Erosion Risk for Cultivated Soil in France. *Catena* **2002**, *46*, 207–220. [[CrossRef](#)]
100. Antoine, P.; Catt, J.; Lautridou, J.; Sommé, J. The Loess and Coversands of Northern France and Southern England. *J. Quat. Sci.* **2003**, *18*, 309–318. [[CrossRef](#)]
101. Chen, S.; Richer-de-Forges, A.C.; Leatitia Mulder, V.; Martelet, G.; Loiseau, T.; Lehmann, S.; Arrouays, D. Digital Mapping of the Soil Thickness of Loess Deposits over a Calcareous Bedrock in Central France. *Catena* **2021**, *198*, 105062. [[CrossRef](#)]
102. Eymard, A.; Richer-de-Forges, A.C.; Martelet, G.; Tissoux, H.; Bialkowski, A.; Dalmasso, M.; Chrétien, F.; Belletier, D.; Ledemé, G.; Laloua, D.; et al. Exploring the Untapped Potential of Hand-Feel Soil Texture Data for Enhancing Digital Soil Mapping: Revealing Hidden Spatial Patterns from Field Observations. *Geoderma* **2024**, *441*, 116769. [[CrossRef](#)]
103. Hardy, M.; Jamagne, M.; Elsass, F.; Robert, M.; Chesneau, D. Mineralogical Development of the Silt Fractions of a Podzoluvisol on Loess in the Paris Basin (France): Mineralogy of Silt Fractions of a Podzoluvisol on Loess. *Eur. J. Soil Sci.* **1999**, *50*, 443–456. [[CrossRef](#)]
104. Urbina-Salazar, D.; Vaudour, E.; Baghdadi, N.; Ceschia, E.; Richer-de-Forges, A.C.; Lehmann, S.; Arrouays, D. Using Sentinel-2 Images for Soil Organic Carbon Content Mapping in Croplands of Southwestern France. The Usefulness of Sentinel-1/2 Derived Moisture Maps and Mismatches between Sentinel Images and Sampling Dates. *Remote Sens.* **2021**, *13*, 5115. [[CrossRef](#)]
105. Urbina-Salazar, D.; Vaudour, E.; Richer-de-Forges, A.C.; Chen, S.; Martelet, G.; Baghdadi, N.; Arrouays, D. Sentinel-2 and Sentinel-1 Bare Soil Temporal Mosaics of 6-Year Periods for Soil Organic Carbon Content Mapping in Central France. *Remote Sens.* **2023**, *15*, 2410. [[CrossRef](#)]
106. Vaudour, E.; Gholizadeh, A.; Castaldi, F.; Saberioon, M.; Borůvka, L.; Urbina-Salazar, D.; Fouad, Y.; Arrouays, D.; Richer-De-Forges, A.; Biney, J.; et al. Satellite Imagery to Map Topsoil Organic Carbon Content over Cultivated Areas: An Overview. *Remote Sens.* **2022**, *14*, 2917. [[CrossRef](#)]
107. Bellón, B.; Bégué, A.; Lo Seen, D.; de Almeida, C.; Simões, M. A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series. *Remote Sens.* **2017**, *9*, 600. [[CrossRef](#)]
108. Bégué, A.; Arvor, D.; Lelong, C.; Vintrou, E.; Simoes, M. Agricultural Systems Studies Using Remote Sensing. In *Remote Sensing Handbook. Vol. II: Land Resources: Monitoring, Modeling, and Mapping*; Thenkabail, P.S., Ed.; CRC Press: Boca Raton, FL, USA; Taylor and Francis Group: London, UK; New York, NY, USA, 2015; pp. 113–130.
109. Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; de Abelleira, D.; Ferraz, R.; Lebourgeois, V.; Lelong, C.; Simoes, M.; Veron, S. Remote Sensing and Cropping Practices: A Review. *Remote Sens.* **2018**, *10*, 99. [[CrossRef](#)]
110. d’Andrimont, R.; Verhegghen, A.; Lemoine, G.; Kempeneers, P.; Meroni, M.; van der Velde, M. From Parcel to Continental Scale—A First European Crop Type Map Based on Sentinel-1 and LUCAS Copernicus in-Situ Observations. *Remote Sens. Environ.* **2021**, *266*, 112708. [[CrossRef](#)]
111. Ghassemi, B.; Dujakovic, A.; Zóltak, M.; Immitzer, M.; Atzberger, C.; Vuolo, F. Designing a European-Wide Crop Type Mapping Approach Based on Machine Learning Algorithms Using LUCAS Field Survey and Sentinel-2 Data. *Remote Sens.* **2022**, *14*, 541. [[CrossRef](#)]
112. Cheng, M.; Liu, X.; Sheng, H.; Yuan, Z. MAPS: A New Model Using Data Fusion to Enhance the Accuracy of High-Resolution Mapping for Livestock Production Systems. *One Earth* **2023**, *6*, 1190–1201. [[CrossRef](#)]
113. Godin, T. Evaluation Des Contraintes Engendrées Par Les Engins Dans Les Systèmes de Grandes Cultures, Viticoles et Forestiers Français. In Proceedings of the 69. IIRB Congress, Institut International de Recherches Betteravières (IIRB). BEL., (hal-02820599), Bruxelles, Belgique, 15–16 February 2006. 19p.
114. Peng, J.; Albergel, C.; Balenzano, A.; Brocca, L.; Cartus, O.; Cosh, M.; Crow, W.; Dabrowska-Zielinska, K.; Dadson, S.; Davidson, M.; et al. A Roadmap for High-Resolution Satellite Soil Moisture Applications—Confronting Product Characteristics with User Requirements. *Remote Sens. Environ.* **2021**, *252*, 112162. [[CrossRef](#)]
115. Håkansson, I.; Reeder, R.C. Subsoil Compaction by Vehicles with High Axle Load—Extent, Persistence and Crop Response. *Soil Tillage Res.* **1994**, *29*, 277–304. [[CrossRef](#)]
116. Arvidsson, J. Subsoil Compaction Caused by Heavy Sugarbeet Harvesters in Southern Sweden. *Soil Tillage Res.* **2001**, *60*, 67–78. [[CrossRef](#)]

117. Arvidsson, J.; Trautner, A.; Van Den Akker, J.J.H.; Schjøning, P. Subsoil Compaction Caused by Heavy Sugarbeet Harvesters in Southern Sweden. *Soil Tillage Res.* **2001**, *60*, 79–89. [[CrossRef](#)]
118. Etana, A.; Håkansson, I. Swedish Experiments on the Persistence of Subsoil Compaction Caused by Vehicles with High Axle Load. *Soil Tillage Res.* **1994**, *29*, 167–172. [[CrossRef](#)]
119. Lal, R.; Vandoren, D.M. Influence of 25 Years of Continuous Corn Production by Three Tillage Methods on Water Infiltration for Two Soils in Ohio. *Soil Tillage Res.* **1990**, *16*, 71–84. [[CrossRef](#)]
120. Jabro, J.D.; Stevens, W.B.; Iversen, W.M.; Sainju, U.M.; Allen, B.L. Soil Cone Index and Bulk Density of a Sandy Loam under No-till and Conventional Tillage in a Corn-Soybean Rotation. *Soil Tillage Res.* **2021**, *206*, 104842. [[CrossRef](#)]
121. Van Orsouw, T.L.; Mulder, V.L.; Schoorl, J.M.; Van Os, G.J.; Van Essen, E.A.; Pepers, K.H.J.; Heuvelink, G.B.M. Practical Implications of the Availability of Multiple Measurements to Classify Agricultural Soil Compaction: A Case-Study in The Netherlands. *Agronomy* **2022**, *12*, 1669. [[CrossRef](#)]
122. Lark, R.M.; Knights, K.V. The Implicit Loss Function for Errors in Soil Information. *Geoderma* **2015**, *251–252*, 24–32. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.