Multiscale evaluations of global, national and regional digital soil mapping products in France

Blandine Lemercier, Philippe Lagacherie, Julien Amelin, Joëlle Sauter, Pascal Pichelin, Anne C Richer-De-Forges, D. Arrouays

Geoderma 425 (2022) 116052

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Multiscale evaluations of global, national and regional digital soil mapping products in France

Blandine Lemercier[a,*], Philippe Lagacherie[b], Julien Amelin[a], Joëlle Sauter[c], Pascal Pichelin[a], Anne C. Richer-de-Forges[d,e], Dominique Arrouays[d]

[a] Institut Agro, INRAE, SAS, 35000 Rennes, France
[b] INRAE, Institut Agro, IRD, LISAH34000, Montpellier, France
[c] Chambre Régionale d'Agriculture Grand-Est, 67000 Strasbourg, France
[d] INRAE, InfoSol, 45075 Orléans, France
[e] INRAE, UR Sol, 45075 Orléans, France

ARTICLE INFO

Handling Editor: Morgan Cristine L.S.

Keywords:
Digital soil mapping
GlobalSoilMap
SoilGrids
Soil survey
Pedological patterns
Evaluation

ABSTRACT

As digital soil mapping (DSM) applications have been developed at multiple extents over the two last decades, large areas of the world are now covered by several DSM products with similar resolution and targeted soil properties. Thus, from these products, end-users must carefully select the one that will best meet their needs. The aim of this study was to evaluate three DSM products obtained at different scales (global, national and regional) over three local territories of increasing area selected in three contrasting regions of France (Alsace, Brittany and Languedoc-Roussillon). Three topsoil (5–15 cm) properties were evaluated: clay content, pH in water and soil organic carbon content. Evaluations were done at both point and soil mapping unit supports, the latter corresponding to quantitative assessment of visual accordance between DSM products and conventional local soil maps of acknowledged quality. The ability to predict soil properties well increased from global to national to regional DSM products. However, none of the DSM products tested was able to predict satisfactorily at the most local (1:25,000) scale. Evaluations of DSM products using local soil maps were generally in accordance with those using points. Evaluations using local soil maps also provided additional information about the utility of DSM products for small areas with too few soil measurements to perform punctual evaluation and for issues concerning areal-support uses of DSM products. These results suggest that when focusing on local areas, users of DSM products should evaluate their performance and, if unsatisfactory, invest in development of local DSM.

1. Introduction

The concept of soil security (Koch et al., 2013; McBratney et al., 2014) considers the soil as a vital and central resource in the Earth system. Thus, information about soil is of major interest and helps address critical worldwide issues such as food security, climate change, environmental degradation, water scarcity and threatened biodiversity (Sanchez et al., 2009). Since the beginning of this century, digital soil mapping (DSM) (McBratney et al., 2003) has been advocated as a promising solution to provide, at acceptable cost, soil information adapted to users' needs. DSM is defined as "the creation and population of spatial soil information systems by numerical models inferring the spatial and temporal variations of soil types and soil properties from soil observation and knowledge and from related environmental variables" (Lagacherie and McBratney, 2006). In a recent study, Arrouays et al. (2021) reviewed how soil mapping and DSM over large areas contribute to soil security at scales ranging from national to global.

Production of soil information using DSM techniques started in the 1970s (Webster and Burrough, 1974), accelerated in the 1990s (Skidmore et al., 1991; Favrot and Lagacherie, 1993; Moore et al., 1993) and has expanded considerably since 2000. In the past decade, a dynamic and growing community has developed and shared DSM procedures (Malone et al., 2017), which has helped DSM enter a new era of operationalisation. Thus, DSM has left laboratories and is encountering end-users' needs. Due to technological advancements in statistical modelling, computational capacity and Earth-system descriptors, as well as the efforts of many research teams, spatial soil information can now be delivered worldwide, whether soil surveys exist or not. The result is a

* Corresponding author.
E-mail address: blandine.lemercier@agrocampus-ouest.fr (B. Lemercier).

https://doi.org/10.1016/j.geoderma.2022.116052
Received 7 January 2022; Received in revised form 5 July 2022; Accepted 10 July 2022
Available online 1 August 2022

growing collection of DSM products of varying resolution and extent. The worldwide *GlobalSoilMap* programme (Arrouays et al., 2014) contributed to this increase in the number of DSM products by providing specifications for the horizontal dimension (2D-resolution), vertical dimension (6 standard depth intervals down to 200 cm) and uncertainty calculations for a set of functional soil properties. *GlobalSoilMap*'s specifications, which were widely adopted for producing digital soil maps, were a major step in enhancing soil information, which is now available worldwide (Poggio et al., 2021). This information improves consideration of the role of soil in the functioning of soil ecosystems and dissemination of soil information outside of soil science. Indeed, the *GlobalSoilMap* format corresponds to the specifications of most bio-physical modelling tools and meets the expectations that many end-users have of ready-to-use data (Voltz et al., 2020).

As DSM applications have been developed at multiple scales (Grunwald, 2009; Minasny and McBratney, 2016; Poggio et al., 2021), many areas in the world are now covered by several DSM products with similar resolution and identical targeted soil properties. For example, Rossiter et al. (2021) inventoried three DSM products for the USA. Similarly, in some regions of France, a user of soil information can choose from the global DSM product SoilGrids version 2.0 (Poggio et al., 2021), the national *GlobalSoilMap* product (Mulder et al., 2016) and a regional *GlobalSoilMap* product (Vaysse and Lagacherie (2015) in Languedoc-Roussillon; Ellili-Bargaoui et al. (2020b) in Brittany).

Although they share the same "scorpan" paradigm (McBratney et al., 2003), the approaches for developing these DSM products differ greatly in soil input data, covariates and machine- learning algorithms. This logically results in large differences in soil property predictions among DSM products. These differences are revealed, for example, by inconsistencies in the global soil organic carbon (SOC) map (FAO and ITS, 2020) at country borders (e.g. USA-Canada border), as each country applied its own DSM approach to map SOC stocks. Rossiter et al. (2021) recently provided quantitative assessments of these differences in the USA. This study revealed that the DSM products differed greatly in estimates of soil pH values and uncertainties (provided for POLARIS and SoilGrids 2.0), both in range and spatial distribution. Richer-de-Forges et al. (in press) performed an evaluation of four DSM products of particle-size distribution prediction developed at various scales (Global, continental, national, and sub-national) using ca. 3,200 hand-feel soil texture data. This study conducted in an area of 550 km^2 of Central France showed that prediction performance decreased between global and sub-national predictions. Consequently, available DSM products are expected to map soil properties of a given territory differently.

From these products, end-users must carefully select the one that will best meet their needs. If they do not, errors in these products could propagate through the decision process or modelling chain, which could degrade study results (Arrouays et al., 2020, Lagacherie et al., 2021). End-users are not always aware of the quality of the data they use, however, and they need guidelines for evaluating DSM products based on assessing the products' uncertainty.

GlobalSoilMap specifications include estimates of uncertainty for each predicted property at each location and each depth layer (Heuvelink, 2014), which could theoretically serve to assess the quality of *GlobalSoilMap* products. However, these estimates of uncertainty should be taken with caution, as recent studies (Lagacherie et al., 2020; Helfenstein et al., 2021) showed significant differences between overall uncertainties calculated from the modelling procedure and those calculated from estimated independent punctual datasets. Moreover, end-users usually focus on study areas much smaller than the spatial extents of DSM products. Gomez and Coulouma (2018) observed large differences in uncertainty among spatial extents for soil properties predicted by spectrometry. It is therefore also expected that reduction in spatial extent strongly influences uncertainties in soil property predictions of DSM products.

To estimate uncertainties, DSM model outputs are usually used. An alternative method is to perform evaluation using an independent validation dataset (Brus et al., 2011; Brus, 2014). However, digital soil maps are usually evaluated using point-support comparisons of observed vs predicted values of soil properties, which is not satisfactory for two main reasons. First, unevenly distributed and non-probabilistic spatial sampling, as usually encountered when using legacy data, generate biases in uncertainty estimates (Lagacherie et al., 2020; Helfenstein et al., 2021). Second, as end-users manage areal (landscape) units rather than points (Rossiter et al., 2021), they are usually much more interested in uncertainty estimates for areal units, which avoids having to consider unmanageable short-range variations in soil properties.

Comparing DSM products to existing high-quality conventional soil surveys can be an alternative to point-support evaluations where such information exists. Conventional soil maps are spatially exhaustive, which solves the above-mentioned problem with spatial sampling while offering many possibilities for spatial aggregation of DSM products. Recently, Rossiter et al. (2021) compared DSM products to soil surveys visually and quantitatively based on spatial aggregations of both at different pixel sizes. Instead of pixels, the elementary information unit of a soil map – the soil mapping unit (SMU) – can be used as the geographical support for comparing DSM products and soil maps (Bishop et al., 2015). Indeed, the SMU represents an optimized areal-support that was delineated by the soil surveyor to best represent the soil variations occurring at a given scale (that depends on the soil map) while filtering the soil variations occurring at more local scales. Following this idea, mean values of predicted soil properties from spatial aggregations of DSM products at the SMU level can be compared to those observed in SMUs. The latter is usually approximated by measuring a "representative profile" of the SMU that is carefully selected by the soil surveyor. Indeed, Leenhardt et al. (1994) observed that representative profiles approximated values of soil properties better at all sites in SMUs than randomly-selected soil profiles did. For complex SMUs, the mean value of a soil property can be approximated by a weighted mean of the values in representative profiles of each soil type unit (STU) in the SMU. Where conventional polygonal soil maps exist, comparisons based on areal units can be useful to quantify and objectivise visual comparisons between DSM products and local soil maps.

The aim of this study was to evaluate three DSM products obtained at different scales (global, national and regional) over three local territories of increasing sizes covered by conventional soil surveys selected within three contrasting regions of France: Alsace, Brittany and Languedoc-Roussillon. Predictions of three topsoil (5–15 cm) properties were evaluated: clay content, pH in water and SOC content. Evaluations were performed using both point and SMU supports, the latter corresponding to quantitative assessment of visual accordance between DSM products and local soil maps.

2. Materials and methods

2.1. Study areas

The three French regions studied (Fig. 1) have different climates, reliefs, geology, soils and agricultural contexts. Located in north-eastern France, Alsace (8,280 km^2) lies on the border with Germany and Switzerland. Eastern Alsace belongs to the Rhine Plain and is relatively flat, whereas western Alsace is part of the Vosges Mountains and has its highest point 1,423 m above sea level. The climate is continental, with contrasting temperatures between winter and summer. The cultivated areas are located mainly in the sedimentary Rhine Plain, whereas hillsides are frequently occupied by vineyards, and mountainous areas are mainly forested.

Located in southern France, Languedoc-Roussillon covers 27,236 km^2 on the coast of the Mediterranean Sea. Its relief varies greatly: flat and close to sea level on the coast, but rugged in the northwest due to the Massif Central Mountains and in the southwest due to the Pyrénées Mountains. Its highest elevation is 2,921 m above sea level and is located in south-eastern Languedoc-Roussillon. The region lies at geological and

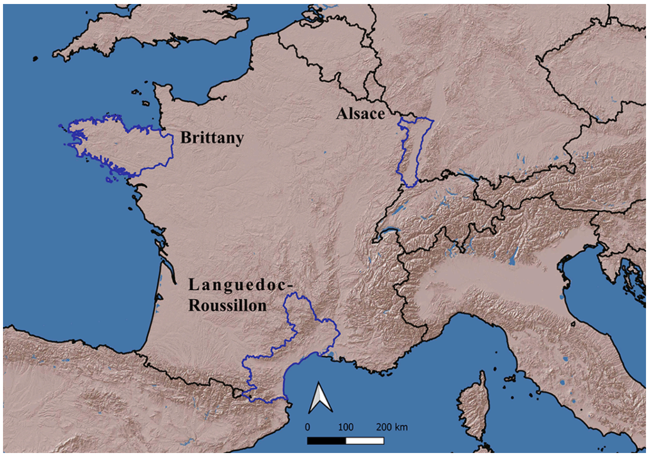

Fig. 1. Locations of the three regions studied in France.

climatic crossroads. Limestone plateaus, volcanic/granitic highlands and sedimentary plains are the main soil parent materials, and Mediterranean, temperate oceanic, temperate continental and mountain climates are observed. Consequently, this region has diverse soils and land uses.

Located in extreme western France, Brittany (27,040 km^2) is a peninsula under oceanic influence. The climate is oceanic, with mean annual rainfall ranging from 650 mm in the east to 1300 mm in the west, and mild temperatures. Brittany lies within the Armorican Massif. The relief is relatively smooth, with a maximum elevation of 382 m above sea level, and strongly correlated with the bedrock, which consists mainly of shale, sandstone and granite. Loess deposits have influenced northern Brittany, allowing for the development of fertile soils. Agriculture is oriented mainly towards integrated crop-livestock production, whereas vegetable field crops occupy the land near the northern coast.

2.2. Soil properties and layer analyzed

The values of three soil properties were compared: clay content, pH in water and SOC content. The underlying processes that influence current values of the 3 soil properties differ depending on the property considered. SOC content depends strongly on climate, soil, land use and management, whereas clay content is considered an inherent soil property. The pH has an intermediate position between SOC and clay contents as it is influenced strongly by parent material and partially modified by soil and land management. Thus, using these three attributes, we expected to cover several trends in variability across and within the study areas, as well as different soil-controlling factors that may be captured by spatial covariates.

The layer 5–15 cm, corresponding to the second standard layer considered in the *GlobalSoilMap* specifications (Arrouays et al., 2014) was analysed. It represents the best compromise between the possible layers considering both the vertical distribution of the studied properties, the fact that the information of the first layers is more reliable because based on more abundant data, and the fact that this layer frequently concerns only one soil horizon (at least in agricultural soils).

2.3. Soil data

2.3.1. Digital soil mapping products

Digital soil maps produced at global, national and regional extents were evaluated: SoilGrids 2.0 (SG2.0), the French *GlobalSoilMap* (GSM_Fr) and regional digital soil maps for Brittany (GSM_Br) and Languedoc-Roussillon (GSM_LR). The main characteristics of these digital soil maps varied (Table 1).

SG2.0 (Poggio et al., 2021) is the second version of a set of soil maps produced for the entire globe at medium spatial resolution (250 m pixel size) using a machine-learning approach, a quantile random forest algorithm trained with ca. 240,000 punctual observations unevenly distributed around the world of which 4,087 were located in mainland France (Batjes et al., 2019) and 400 global environmental covariates. This modelling procedure yields maps of values of physical and chemical soil properties, and their associated uncertainties, for six standard interval depths. Data were downloaded from https://soilgrids.org/, powered by ISRIC (last access: 12/20/2021).

Table 1
Characteristics of the four digital soil mapping products compared.

Characteristic	SoilGrids 2.0	GSM France	GSM Brittany	GSM Languedoc-Roussillon
Spatial extent	Global	France	Brittany region	Languedoc-Roussillon region
Spatial resolution (m)	250	90	50	90
Density of punctual observations (km^2 per observation)	617 (134 for France)	16–19*	36	13.5
Algorithm	Quantile random forest	Cubist regression tree	DSMART	Quantile random forest
Reference	Poggio et al. (2021)	Mulder et al. (2016)	Ellili-Bargaoui et al. (2020b)	Vaysse and Lagacherie (2017)

* depending on the soil property predicted.

GSM_Fr products (Mulder et al., 2016) were calculated at 90 m spatial resolution and for the six standard depth intervals. Input data were punctual field observations (23,822–57,915, depending on the soil property for the 5–15 cm depth interval) and 20 covariates for prevailing climate regimes and meteorological data, vegetation, topography, geology, soils and land management. The 90 % confidence intervals were estimated to provide uncertainties.

GSM_LR products were calculated at 90 m resolution for the six standard depth intervals from a set of 2,024 legacy soil profiles associated with a set of 16 covariates freely available (at least at the national level), using a quantile random forest algorithm (Vaysse and Lagacherie, 2017) that also provided local uncertainty estimates.

GSM_Br products are available for the six standard depth intervals at 50 m spatial resolution. A two-step procedure was applied. The existing 1:250,000 soil map (was first disaggregated using the DSMART algorithm (Odgers et al., 2014) and soil-landscape relationships (Vincent et al., 2018, Ellili-Bargaoui et al., 2020a) to map the three most probable STUs. Soil properties of each of the 320 STUs were then estimated for standard soil-depth intervals by applying equal-area spline functions to each STU. Finally, to map soil properties at the six standard depths, the weighted mean of each soil attribute was calculated for each grid cell from reference soil-property values of the three most probable predicted STUs, using their associated probabilities of occurrence as the weights (Ellili-Bargaoui et al., 2020b).

2.3.2. Reference data

2.3.2.1. Soil maps. The reference soil maps compared to DSM products were selected according to two criteria: i) built by experienced local soil surveyors who followed conventional pedological approaches (Jamagne et al., 1967) and ii) available in a digital format. For each of the three regions studied, three soil maps were selected to explore three scales – 1:250,000, 1:50,000 to 1:100,000 and 1:25,000 – focusing on study areas of decreasing size (Fig. 2, Table 2).

Regional (1:250,000) soil maps covered the three regions entirely.

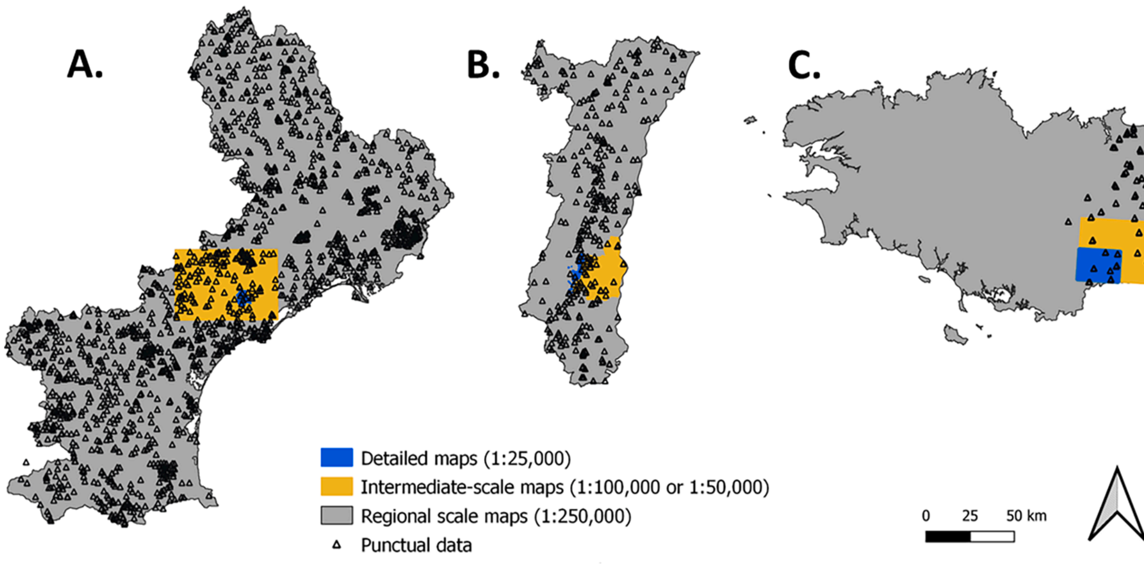

Fig. 2. Spatial extent of reference polygonal soil maps and profile data used to evaluate the digital soil mapping products in the Languedoc-Roussillon (A), Alsace (B) and Brittany (C) regions of France. In Brittany, punctual information are independent from all DSM products.

Table 2
Main characteristics of the three scales of reference soil maps used for the three regions studied: Languedoc-Roussillon (LR), Alsace (Al) and Brittany (Br).

Scale	1:250,000			1:100,000 – 1:50,000			1:25,000		
Region	LR	Al	Br	LR	Al	Br	LR	Al	Br
Area (km^2)	27,156	8,329	27,376	1,761	554	2,156	49	42	540
No. of SMUs	396	82	341	–	–	–	–	–	–
No. of STUs	732	377	320	159	41	72	29	42	75
No. of polygons	3,996	3,471	2,043	3,126	882	4,226	325	251	4114
Mean area (ha)	680	240	1,340	73	63	51	15	17	13
SD area (ha)	1,598	1,177	3,031	171	286	554	26	28	25

SMU: Soil map unit; STU: Soil type unit; SD: Standard deviation.

They followed specifications of the French National Soil Survey (INRAE InfoSol, 2014), and their soil databases were built following the national format DoneSol (INRAE InfoSol, 2021), which ensured that the databases for each French region had the same structure. Soil information is spatially represented in polygons with crisp boundaries. SMUs group together one or usually several polygons and are defined as areas with homogeneous soil-forming factors, such as morphology, geology and climate. SMUs are complex, as each contains known proportions of several STUs, which are described in the semantical DoneSol database by a set of soil layers that represent vertical variation in the soil (Laroche et al., 2014).

Intermediate-scale soil maps covered part of each region (Richer-de-Forges et al., 2014). In Languedoc-Roussillon, the 1:100,000 pedological map of Lodève (Bonfils, 1993) covers 8 % of the region. In Alsace, the 1:50,000 soil map of the “Rhin-Vignoble-Grand Ballon” territory (Sauter J., pers. comm.) covers 7 % of the region. In Brittany, the 1:100,000 soil map of Janzé (Rivière et al., 2011) covers 8 % of the region.

In each region, a local and detailed (1:25,000) polygonal soil map was chosen, nested within the intermediate-scale maps: the Peyne soil map in Languedoc-Roussillon (Coulouma et al., 2008), the vineyard area of the “Rhin-Vignoble-Grand Ballon” territory in Alsace (Sauter J., pers. comm.) and the Pipriac soil map in Brittany (Rivière et al., 1984). The SMUs of intermediate-scale and detailed maps were simple, meaning that only one STU was described in each SMU and polygon, even if in reality they may have inclusions of others STUs that were not recorded.

2.3.2.2. Punctual measurements. Punctual measurements of soil properties of interest for this study were available for the three regions. For Languedoc-Roussillon and Alsace, legacy datasets of punctual measurements (1,696 and 346, respectively), collected in the past few decades to improve pedological knowledge and map soil at multiple scales, were used. For Brittany, an independent evaluation dataset of 135 punctual measurements sampled according to a probabilistic sampling procedure (Ellili Bargaoui et al., 2019) was used.

2.4. Data processing

To make comparison of digital soil maps and reference data feasible, the following pre-processing procedures were applied to the datasets, if necessary:

- Data cleaning to discard cells with no value in gridded maps
- Standardizing units for the three soil properties in all databases
- Assigning a unique Coordinate Reference System (CRS) to all vectors and rasters: WGS84 – EPSG (European Petroleum Survey Group) code: 4326
- Adjusting the spatial extents of soil maps using regional masks to ensure compatible extents

To process data to compare DSM products to the reference datasets, the first step (pre-processing) consisted of extracting information about soil properties from the punctual and polygonal databases (Fig. 3). Soil profiles were vertically divided into pedological horizons, each of which had values for clay content, pH and SOC content. Horizons overlapped the 5–15 depth interval were selected, and weighted means of soil properties were calculated using the proportion of each layer’s thickness in the target depth interval as the weights. The same procedure was applied to the layers of STUs for intermediate-scale and detailed maps,

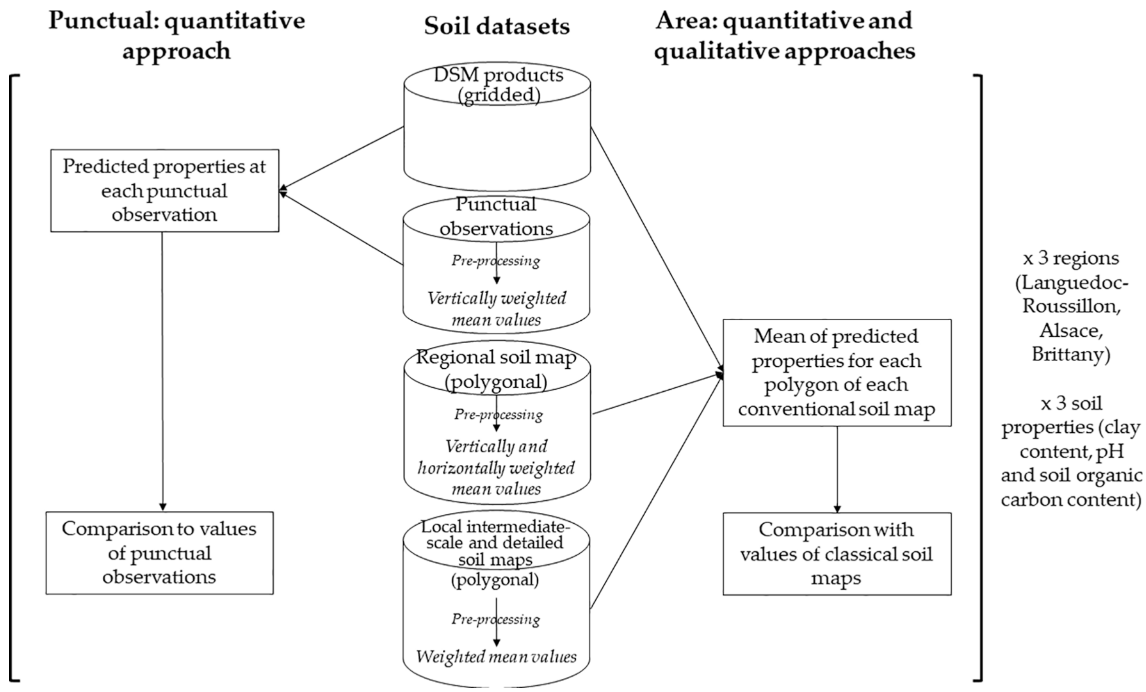

Fig. 3. Workflow for evaluating the accordance between digital soil mapping (DSM) products and legacy soil profiles or conventional soil maps.

with values extracted from the representative profiles of the STU. For complex SMUs of regional polygonal soil maps, weighted means of clay content, pH and SOC content were calculated for each STU as before, and then an additional calculation yielded horizontally weighted means, using the probability of occurrence of each STU in each SMU as the weight (Vincent et al., 2018).

The comparisons were performed using both point and areal supports. Point-support comparisons consisted of quantitatively comparing global and national DSM products to punctual soil measurements. To this end, values of SG2.0 and GSM_Fr were extracted for each profile location. As regional DSM products could not be compared to the punctual soil measurements that had been used as inputs to develop the products, different strategies were applied. For GSM_LR, results of cross-validation from a previous evaluation (Vaysse and Lagacherie, 2017) were considered. For GSM_Br, we used the independent set of 135 punctual measurements.

Areal-support comparisons consisted of comparing, for each map polygon, zonal means of a soil property derived from the DSM products to the estimated value of the property given by the map following the procedure described previously. For each region, areal-support comparisons were performed for the three soil-map scales considered. Data pre-processing and processing were performed using QGIS 3.12 software (QGIS Development Team, 2018), along with the Python (Van Rossum and Drake, 2009) and R 3.6.2 (R Core Team, 2019) languages.

2.5. Statistical indicators for quantitative comparison

For quantitative comparison, the similarity between datasets was assessed using several statistical indicators: the slope (SL), intercept and correlation coefficient of the linear regression between observed and predicted soil properties. In addition, the mean error (ME, Eq. (1)) was used to identify potential bias, and the root mean squared error (RMSE, Eq. (2)) measured the difference between values of two datasets. Two indicators of the overall quality of adjustment were calculated: the mean squared error skill score (SS_{mse}, Eq. (3)), which represents the amount of variance explained by the prediction (Wilks, 2011; Nussbaum et al., 2018), and Lin's concordance correlation coefficient (LCCC, Eq, 4).

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (2)$$

$$SS_{mse} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(y_i - \frac{1}{n}\sum_{i=1}^{n}y_i\right)^2} \quad (3)$$

$$LCCC = \frac{2\rho\sigma_{\hat{y}}\sigma y}{\sigma_{\hat{y}}^2 + \sigma_y^2 + (\bar{\hat{y}} - \bar{y})^2} \quad (4)$$

where n is the number of soil samples, y_i the observed values, $\hat{y}_i$ the predicted values, $\bar{\hat{y}}$ and $\bar{y}$ the means of the predicted and observed values, respectively, σ^2 their variance and ρ the correlation between predicted and observed values.

For brevity, only the indicators that illustrated the largest difference between the DSM products tested are presented: ME, RMSE, LCCC and SS_{mse} for point-support comparisons, and LCCC and SL for areal-support comparisons. Statistical indicators were calculated using R software 3.6.2 (R Core Team, 2019).

2.6. Objectivizing visual comparisons of the spatial distribution of soil properties' values

To supplement the punctual evaluation, DSM products were also evaluated by visual assessment of the similarity of the spatial organization of DSM products values aggregated to the map unit polygons to that of high-quality soil maps used as references. Thus DSM products aggregated to the polygons of the reference maps were compared to values given by the reference maps in order to evaluate if the spatial patterns highlighted with the reference maps are reproduced by the aggregated DSM products or not. To strengthen and objectivise the visual comparison, a quantitative and reproducible decision rule was established to allow us to determine whether the soil geography structure of a DSM product were similar to those of the reference map or not.

The decision rule consists in the combination of the two statistical indicators used for areal-support comparison, associated with thresholds defined through trial-and-error:

- LCCC related to the ability of DSM products to reproduce the polygon's representative value observed from the conventional soil maps, with a threshold of 0.25 set;
- SL represented the range of DSM products' values averaged by polygons to the reference values from conventional maps, i.e. the degree of contrast of DSM products relatively to conventional soil maps. A threshold of 0.35 was set.

3. Results

3.1. Preliminary results

Soil properties varied among regions and scales (Table 3). At regional level (SG2.0, GSM_Fr, GSM_LR, GSM_Br and 250), Languedoc-Roussillon, then Alsace exhibited higher clay contents and pH than Brittany. In addition, the clay content and pH values were much less scattered in Brittany than in the other regions. Regarding SOC, the values were higher in Brittany than in Languedoc-Roussillon and Alsace, but a greater dispersion of the data was noted for Languedoc-Roussillon, again at regional level. The detailed soil map in Languedoc-Roussillon covered an area with higher pH and lower SOC than reported from other maps. In Alsace, soils in the area covered by intermediate and detailed soil maps were more clayey and basic and contained less SOC than the region as a whole. In Brittany, mean and median values the three soil properties from the intermediate-scale map were similar to regional values and the detailed-scale map exhibited a slightly higher level of clay content. For point values, standard deviations and interquartile ranges highlighted a greater dispersion of values for all regions, scales and data sources, with the exception of SOC in Brittany and pH in Alsace.

3.2. Evaluation of punctual predictions

According to statistical indicators of the point-support evaluation, for most of the comparisons tested (18 of 24, i.e. 75 %), DSM products exhibited only moderate biases in predicting clay content ($\leq$20 g.kg^{-1}), pH ($\leq$0.2) and SOC content ($\leq$10 g.kg^{-1}) (Table 4). Four exceptions were observed for SG2.0, which strongly overestimated clay content in Languedoc-Roussillon and Alsace and pH in Languedoc-Roussillon and Alsace. One exception was observed for GSM_Fr, which strongly underestimated clay content in Brittany and SOC content in Languedoc-Roussillon. The regional DSM exhibited no strong biases, regardless of the property or region.

SG2.0 generally had poor prediction performances for most of the comparisons tested (SS_{mse} < 0.20 for 7 of 9 comparisons), except for predicted pH in Languedoc-Roussillon and Brittany, capturing a small but significant part of the variability (SS_{mse} = 0.24 and 0.30, respectively) (Table 4). GSM_Fr predicted pH slightly better than SG2.0 in Languedoc-Roussillon and Brittany (SS_{mse} = 0.56 and 0.35, respectively), while predicting pH well in Alsace (SS_{mse} = 0.38). When tested, the regional DSM products (GSM_LR and GSM_Br) outperformed SG2.0 and GSM_Fr, with a large increase in SS_{mse} and LCCC for all properties and regions, except for SS_{mse} for clay content and pH in Brittany. However, regional DSM products still predicted clay content poorly (SS_{mse} < 0.20), although LCCC for clay content was satisfactory in Brittany (LCCC = 0.59). The DSM products predicted the soil properties with differing prediction quality: overall, pH was predicted best, followed by SOC content and then clay content (Table 4).

Table 3
Descriptive statistics of clay content, pH in water and soil organic carbon content for the 5–15 cm depth interval for the four digital soil mapping products (SoilGrids 2.0 (S2.0), French GlobalSoilMap (GSM_Fr) and regional GlobalSoilMap for Brittany (GSM_Br) and Languedoc-Roussillon (GSM_LR), conventional soil maps at regional (250), intermediate (100 and 50) and detailed scales (25) and punctual information.

	Languedoc-Roussillon							Alsace						Brittany						
Clay content (g.kg^{-1})	SG2.0	GSM_Fr	GSM_LR	250	100	25	points	SG2.0	GSM_Fr	250	50	25	points	SG2.0	GSM_Fr	GSM_Br	250	100	25	points
Mean	260.1	191.4	191.6	212.4	172.7	236.9	198.9	279.2	206.7	204.5	243.6	236.2	228	209.1	175.3	183.5	165.1	189.7	201.6	213.7
Median	261	185.6	190.3	200.5	170.0	250	180.0	279	204.4	201.0	240.0	250.0	210.0	208.0	176.3	175.8	180.6	184.9	203.0	195.0
SD	35.7	43.7	42.8	78.3	89.4	80.4	108.7	37.8	60.2	90.5	45.7	71.5	110.0	26.3	19.9	43.8	31.9	44.3	32.0	84.4
IQR	46.0	68.0	63.4	90.0	100.0	140.0	131.5	49.0	78.7	86.5	90.0	100.0	127.1	31.0	22.8	53.9	35.9	34.5	43	83.1
pH in water																				
	SG2.0	GSM_F	GSM_LR	250	100	25	points	SG2.0	GSM_F	250	50	25	points	SG2.0	GSM_F	GSM_Br	250	100	25	points
Mean	7.0	6.9	7.0	6.9	6.6	7.9	6.0	6.2	6.2	6.3	7.4	7.5	6.8	5.9	5.7	5.7	5.7	5.5	5.6	6.1
Median	7.0	7.5	7.8	7.2	6.5	8.2	6.6	6.5	6.6	6.6	7.7	8.0	6.1	6.0	5.8	5.6	5.6	5.1	5.7	6.0
SD	0.6	1.3	1.2	1.4	1.3	0.7	2.6	0.7	1.2	1.4	0.6	0.8	1.1	0.3	0.5	0.4	0.4	1	0.3	1.0
IQR	0.9	3.2	3.6	2.5	2.3	1.1	3.2	2.1	2.3	2.6	1.5	1.1	1.6	0.3	0.4	0.3	0.2	0.9	0.2	0.9
Soil organic carbon content (g.kg^{-1})																				
	SG2.0	GSM_F	GSM_LR	250	100	25	points	SG2.0	GSM_F	250	50	25	points	SG2.0	GSM_F	GSM_Br	250	100	25	points
Mean	34.1	20.8	26.4	30.1	29.4	10.2	39.8	34.9	24.2	21.6	12.3	15.2	20.1	35.5	26.4	30.1	33.1	29.4	27.1	27.9
Median	30.4	19.2	21.2	24	20.9	8.2	17.5	27.9	19.7	16.4	10.8	13.9	14.1	34.2	25.6	28.0	27.9	25.9	25.7	22.0
SD	17.1	11.1	18.0	25.8	32.6	4.8	84.4	18.8	12.1	13.4	3.4	4.4	23.0	9.5	5.6	34.9	20.7	11.6	34.3	27.3
IQR	22.4	14.1	22.8	23.6	21.1	7.0	26.2	20.5	13.8	16.8	3.3	4.6	10.6	14.2	8.0	14.8	11.8	22.7	14.8	13.0

SD: standard deviation; IQR: Interquartile range.

Table 4
Quantitative indicators of comparison between observed punctual values and predicted values from global (SG2.0), national (GSM_Fr) and regional (GSM_LR for Languedoc-Roussillon and GSM_Br for Brittany) digital soil mapping products for clay content, pH in water and soil organic carbon content for the 5–15 cm depth interval. Results for GSM_LR came from Vaysse and Lagacherie (2017).

		Languedoc-Roussillon			Alsace		Brittany		
Property	Indicator	SG2.0	GSM_Fr	GSM_LR	SG2.0	GSM_Fr	SG2.0	GSM_Fr	GSM_Br
Clay content ($g.kg^{-1}$)	ME	53.7	−13.3	−14	58.8	−8.3	20	−30.6	−3.6
	RMSE	118.4	98.7	96	124.7	105.6	76.8	82.7	77.6
	SS_{mse}	−0.25	0.13	0.17	−0.29	0.08	0.16	0.03	0.14
	LCCC	0.1	0.26	/	0.07	0.27	0.29	0.17	0.59
pH in water	ME	0.3	0.2	0.1	0.12	0.03	0.6	−0.1	−0.2
	RMSE	1.2	1	0.82	1.6	0.9	0.8	0.8	0.8
	SS_{mse}	0.24	0.56	0.67	−1.06	0.38	0.3	0.35	0.32
	LCCC	0.46	0.74	/	0.12	0.62	0.32	0.46	0.64
Soil organic carbon content ($g.kg^{-1}$)	ME	−7.5	−19.5	−9	6	−1.5	1.2	−6	0.5
	RMSE	78.8	82	41	23.6	22.5	24.6	26.9	23.1
	SS_{mse}	0.13	0.06	0.26	−0.05	0.04	0.16	−0.01	0.26
	LCCC	0.14	0.1	/	0.24	0.19	0.21	0.05	0.47

/: not provided by Vaysse and Lagacherie (2017).
ME: Mean error; RMSE: Root mean squared error; SS_{me}: Mean squared error skill score; LCCC: Lin's concordance correlation coefficient.

3.3. Comparisons of DSM products to soil maps

3.3.1. Example of pH in Languedoc-Roussillon

The decision rule established to objectivise the visual assessment of the similarity of the predicted versus conventional soil spatial distribution (section 2.6) is illustrated using the example of pH in Languedoc-Roussillon, which had the greatest range of performances among DSM products (Table 4). Using this decision rule, the ability of DSM products to reproduce soil pH patterns of reference maps in Languedoc-Roussillon (Fig. 4) was classified as follows:

- o SG2.0 compared to 1:250,000 and 1:100,000 maps: partial reproduction of polygons' representative values (LCCC > 0.25) but low contrast (SL < 0.35)
- o GSM_Fr compared to 1:250,000 and 1:100,000 maps: partial reproduction of polygons' representative values (LCCC > 0.25) and satisfactory contrast (SL > 0.35)
- o SG2.0 and GSM_Fr compared to 1:25,000 maps: no reproduction of polygons' representative values (LCCC < 0.25) and low contrast (SL < 0.35)
- o GSM_LR compared to 1:250,000 and 1:100,000 maps: partial reproduction of polygons' representative values (LCCC > 0.25) and satisfactory contrast (SL > 0.35)
- o GSM_LR compared to the 1:25,000 map: no reproduction of polygons' representative values (LCCC < 0.25) and low contrast (SL < 0.35)

3.3.2. Evaluation of soil pattern accordance

We applied the decision rule to all 72 comparisons of the experiments (i.e. 3 properties × 3 regions × 3 soil reference maps of various scales × 2 products (SG2.0 and GSM_Fr) + 3 properties × 2 regions × 3 scales × 1 product (GSM_LR or GSM_Br)).

For most comparisons (49 of 72, i.e. 68 %), soil patterns of the GSM products were not in accordance with those in regional, intermediate-scale or detailed soil maps (LCCC < 0.25) (Table 5). Conversely, a few comparisons (13 of 72, i.e. 18 %) were in partial accordance with soil patterns in soil maps, with satisfactory contrast (SL > 0.35). For the remaining comparisons (14 %), the DSM products represented some of the soil patterns in soil maps (LCC > 0.25), but with low contrast (SL < 0.35) (i.e. variability underestimated).

Where they were tested, regional DSM products always outperformed national and global products, regardless of the region or soil property. DSM products did not reproduce reference values of all soil properties equally well. Patterns of pH were clearly reproduced better - only 12 of 24 (50 %) disagreements - than those of clay – 19 of 24 (79 %) disagreements and SOC contents – 18 of 24 (75 %) disagreements.

The ability of DSM products to represent soil patterns of reference maps decreased as the scale at which soil surveyors had delineated them increased (Fig. 5). In addition, there was a clear hierarchy among DSM products: GSM_Fr outperformed SG2.0, and regional DSM products outperformed national DSM products.

3.3.3. Relations between punctual evaluation and comparison of soil patterns

Results of punctual evaluations of the DSM products at the regional scale (Table 4) were related to those of evaluations of the ability of DSM products to reproduce soil variations given by reference maps (Table 5), based on the thresholds used to classify the comparisons. DSM products that failed to predict soil variations in accordance with soil maps (LCCC < 0.25) also failed to predict soil properties at punctual sites (SS_{mse} < 0.20). Conversely, all DSM products that captured some variability in a soil property at punctual sites ($SS_{mse} \geq 0.20$) also predicted regional soil patterns in accordance with regional soil maps. Interestingly, eight DSM products that were in accordance with soil maps did not predict punctual sites well.

4. Discussion

4.1. Need for local evaluation of DSM products

The uncertainty in punctual evaluations of DSM products performed for three contrasting regions of France (Table 4) were higher than those obtained by the authors of the DSM products tested (Table 6), except for the pH map of *GlobalSoilMap* France in Languedoc-Roussillon. Other researchers observed this result with the previous version of SoilGrids (Hengl et al., 2017), such as Tifafi et al. (2018), Song et al. (2020) and Dharumarajan et al. (2021). This result may be due to the size of area studied, which biases the percentage of variance explained, as a larger area contains more contrasting situations. Thus, when a DSM product is used for spatial extents that differ from that for which it was created, it should be re-evaluated with local references.

4.2. Possible biases due to legacy data

The legacy soil data used in this study consisted of point data and soil-property maps, which themselves were based on measured point values. However, this information was acquired over a few decades. While clay content can be assumed not to vary over this period, pH or SOC content may do so, depending mainly on soil use and management. Thus, the probable change in these two properties may have introduced bias when evaluating the DSM products. This change influenced mainly Brittany, where the difference in date between the legacy data used to

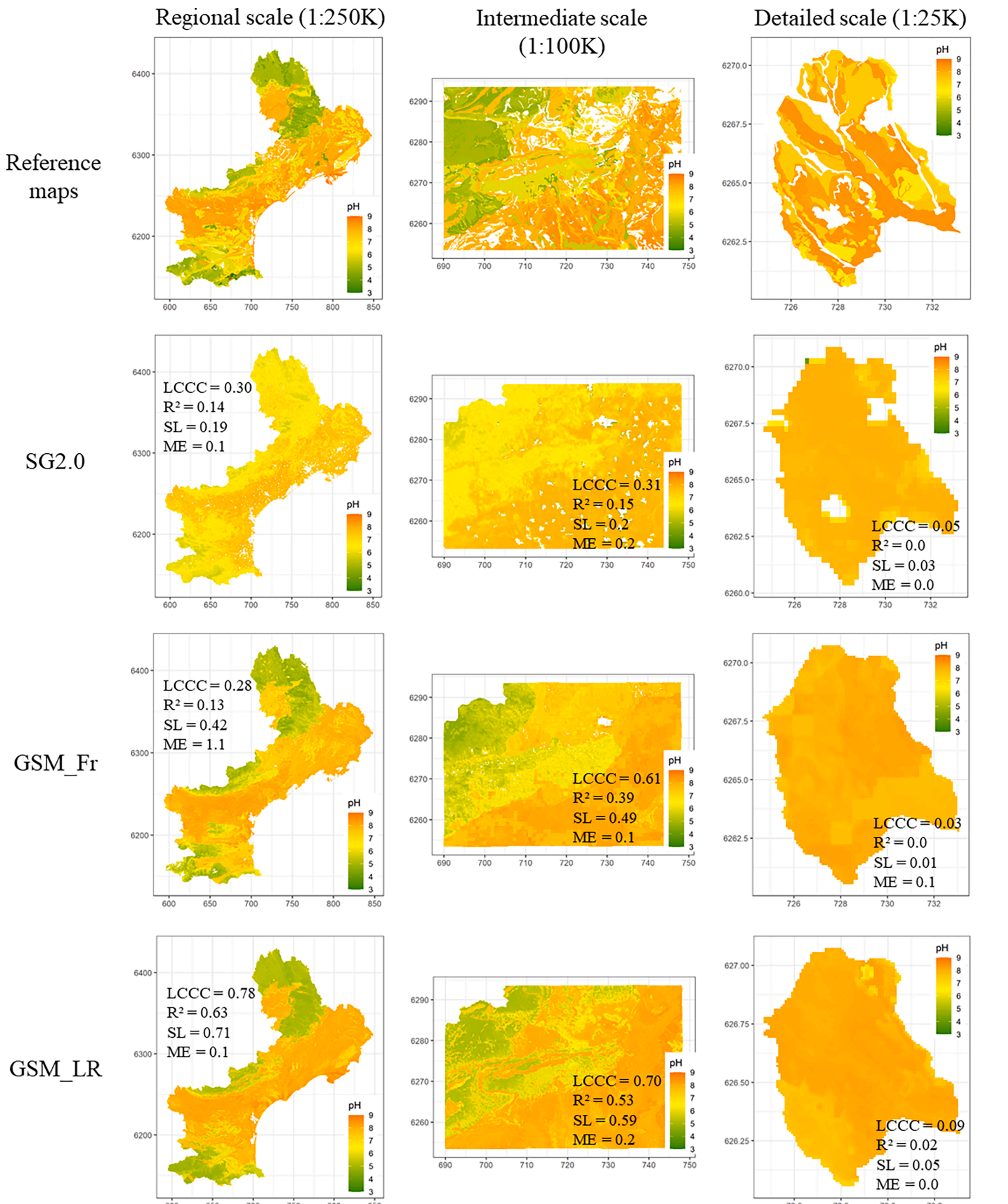

Fig. 4. Visual comparisons of soil spatial distributions and indicators of accordance for the example of pH in Languedoc-Roussillon from reference maps and tested digital soil mapping products at three scales: (left column) regional; (middle column) intermediate; and (right column) detailed. Units of geographic coordinates are kilometres. LCCC: Lin's concordance correlation coefficient; R2: Correlation coefficient of the linear regression; SL: Slope of the linear regression; ME: Mean error.

create the DSM products and the independent dataset used in this study was large. We estimate that this bias is weak, however, because the magnitude of change is smaller over time than over space, and spatial structures with contrasting values are conserved over time, as verified for pH in France using the national soil test database (Saby et al., 2017).

For Languedoc-Roussillon and Alsace, we estimate that the bias of legacy data is weak because the point data used for the evaluation were also included in the datasets used to calibrate the national DSM model

Table 5
Degree of accordance between digital soil mapping products and reference soil maps across soil properties and scales for the French regions of Languedoc-Roussillon (LR), Brittany (Br) and Alsace (Al). Lin's concordance correlation coefficient (LCCC) assesses reproduction of polygons' representative values given by the reference maps, and the slope of the linear regression (SL) assesses the degree to which the range of soil variability is similar. Colours indicate whether LCCC or SL lies above (green) or below (pink) the satisfactory threshold (0.25 and 0.35, respectively). Grey: not compared.

Scale	Region	Clay content						pH in water						Soil organic carbon content					
		SG2		GSM_Fr		GSM_Reg		SG2		GSM_Fr		GSM_Reg		SG2		GSM_Fr		GSM_Reg	
		LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL	LCCC	SL
Regional	LR	0.07	0.05	0.25	0.16	0.32	0.2	0.3	0.19	0.28	0.42	0.78	0.71	0.54	0.38	0.32	0.2	0.53	0.36
	Br	0.1	0.13	0.12	0.21	0.99	0.94	0.25	0.18	0.31	0.28	0.93	0.97	0.16	0.09	0.09	0.05	0.93	0.9
	Al	0.03	0.03	0.23	0.15			0.45	0.3	0.62	0.55			0.34	0.61	0.45	0.38		
Intermediate	LR	-0.01	-0.01	0.14	0.08	0.24	0.13	0.31	0.2	0.61	0.49	0.7	0.59	0.23	0.13	0.18	0.11	0.23	0.14
	Br	0.02	0.02	0.06	0.11	0.39	0.28	0.02	0.01	0.06	0.04	0.07	0.05	0.04	0.02	0.04	0.02	0.07	0.05
	Al	0.03	0.03	0.01	0.01			0.11	0.15	0.52	0.45			0.03	0.12	0.02	0.01		
Detailed	LR	0.11	0.07	0.02	0.02	0.07	0.06	0.05	0.03	0.03	0.01	0.09	0.05	0.09	0.1	0.16	0.1	0.11	0.06
	Br	0.07	0.06	0.12	0.11	0.34	0.28	0.05	0.05	0.2	0.18	0.16	0.1	0.04	0.02	0.01	0	0.03	0.01
	Al	-0.01	-0.01	0.06	0.06			-0.03	-0.06	0.11	0.1			0.01	0	-0.03	-0.03		

GSM_Reg: regional DSM products.

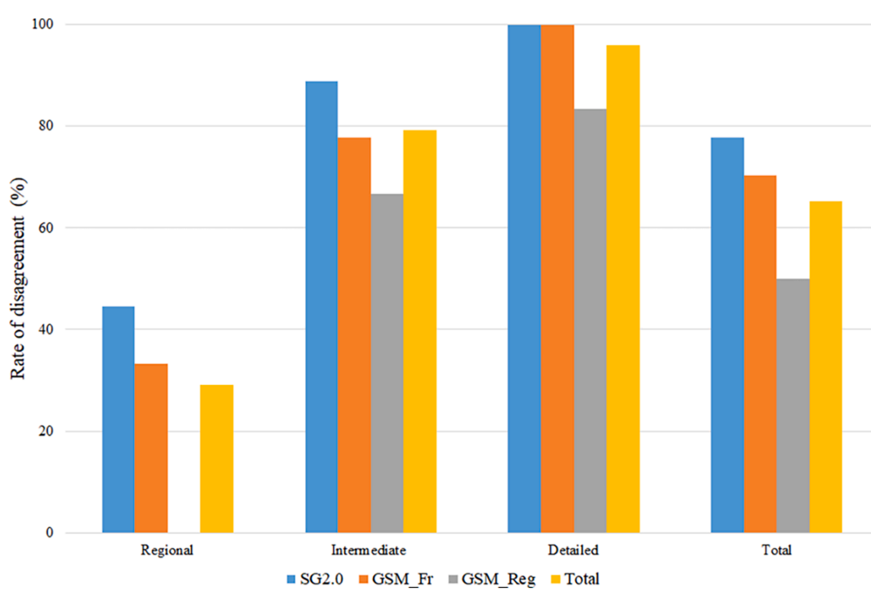

Fig. 5. Overall percentages of global (SG2.0), national (GSM_Fr), regional (GSM_Reg) and overall digital soil mapping products not in accordance with soil patterns of soil maps across scales (Lin's concordance correlation coefficient < 0.25).

Table 6
Proportion of variance of predicted soil properties explained in articles on SoilGrids version 2 (Poggio et al., 2021) and GlobalSoilMap for France (Mulder et al., 2016).

Property (5–15 cm)	SoilGrids	*GlobalSoilMap* France
Clay content	0.42	0.30
pH in water	0.68	0.48
Soil organic carbon content	0.39	0.20

and that for Languedoc-Roussillon. However, this raises another issue: this approach may have overestimated the prediction quality when comparing DSM products to these punctual measurements. However, this influenced only comparisons of the regional DSM product of Languedoc-Roussillon, as regional datasets of soil measurements represented too little of all input datasets of the national GSM product to influence its uncertainty assessment significantly.

4.3. Utility of evaluating DSM products using local soil maps

In this study, we compared global, national and regional (where available) DSM products to local soil maps. Quantitative indicators (LCCC and SL) helped determine whether the soil patterns shown by the DSM products were consistent with the patterns shown by the soil maps. Although visual comparisons have been performed to assess the quality of DSM products (Hengl et al., 2017), to our knowledge, until recently, no study attempted to objectivise these comparisons using quantitative indicators. Rossiter et al. (2021) recently compared DSM products to conventional soil maps for square areas using a set of quantitative indicators that expressed the similarity/dissimilarity between products and characterised their predicted spatial pattern. However, none of these indicators was explicitly related to the visual comparisons between soil maps and DSM products or evaluations at punctual sites.

Evaluating DSM products with soil maps using quantitative indicators that express the similarity of predicted soil patterns was relevant, as good agreement was found with conventional evaluation using spatial sets of punctual soil-property measurements (section 3.3.3). However, soil maps delineate soil classes, not soil properties. Furthermore, it is well known that soil maps have their own uncertainties in predicting soil properties (Marsmann and De Gruijter, 1986; Leenhardt et al., 1994) and that their model of spatial representation of soil variability (the choropleth map) distorts reality severely (Burrough, 2006). Thus, conventional soil maps should not be considered as ground truth but rather as another (inevitably distorted) picture of spatial variations, with which DSM products should be in partial accordance. This view matches well with the low thresholds of LCCC and SL used in this study to identify partial accordance between soil maps and DSM products.

Finally, when available, using existing local soil maps to evaluate DSM products considerably increases the possible ways to evaluate DSM products. This approach is a solution to the limitation of conventional point-support evaluations, which are the worst-case scenario for assessing the prediction quality of DSM products (Bishop et al., 2015). In this perspective, the LCCC of comparisons with soil maps (Table 5) were generally higher than those of comparisons with punctual measurements (Table 4). Similarly, some DSM products that did not predict punctual sites well still exhibited partial accordance with conventional regional soil maps, for example the national *GlobalSoilMap* for SOC content in Languedoc-Roussillon and Alsace. This result could be useful to avoid rejecting DSM products for applications that consider areal units and thus do not need good punctual estimates but only good representation of the trends of soil variation over an area. In addition, using existing local soil maps as a reference is helpful to evaluate DSM products for small areas for which too few profile measurements may exist to perform traditional validation, as shown in this study.

4.4. Performances of DSM products developed at different scales

Results for the three regions clearly showed that *GlobalSoilMap* France outperformed SoilGrids 2.0 for predicting soil properties and showing realistic soil spatial distribution at the regional scale. Likewise, regional GSM products, when available, outperformed the national GSM product. We thus observed an increase in the quality of DSM products, as they were developed at a more local scale. These results are consistent with those of Richer-de-Forges et al. (in press) who pointed out that the predictions of texture classes deriving from the regional model were the closest to the manually assessed soil texture classes, followed by *GlobalSoilMap* France and then SoilGrids 2.0. However these results do not agree with those of Rossiter et al. (2021), who observed no increase in quality from SoilGrids 2.0 (global scale) to the national GSM product for the USA. As there is no evidence that SoilGrids 2.0 used significantly more profiles from the USA than from France (see Poggio et al. (2021), Figs. 3 and 4), these differences could be explained by differences in the geographical supports considered in the present study and those of Rossiter et al. (2021) when comparing DSM products to soil maps. Differences in complexities of soil patterns between France and the USA may also have an influence. As Minasny et al. (2010) highlight, mainland France has one of the highest diversity of soil types in Europe and the rest of the world.

Finally, the comparison to local soil patterns in soil maps revealed that none of these DSM products was relevant for use at local extent despite an apparent high resolution that might suggest the opposite to end-users. Indeed, confusion between spatial resolution and accuracy is frequent, and accompanying predicted values with confidence intervals or, even better, probability function distributions, helps to clarify the concepts. These results for comparison of local soil patterns provided by DSM products to local soil maps agree with those of Rossiter et al. (2021). DSM techniques therefore need to be applied to small extents ("local DSM") to satisfy users' needs at a local scale (Arrouays et al., 2017). Doing so would require great effort to populate DSM models with soil data and relevant environmental covariates over such small extents, so that the models are able to capture local and usually small variations.

In the future, such results will need to be complemented by comparisons of uncertainty assessments provided by the DSM products that will constitute an important output for the end users.

4.5. Drivers of GSM product performances

Despite sharing a common methodological framework, the performances of DSM products differed greatly across properties, regions, scales of development and scales of application. Some possible causes of these differences include the following:

- **Targeted soil property:** pH was generally predicted better than clay or SOC contents. This result agreed with the global and national evaluations provided by the authors of the DSM products (Table 6). This is also consistent with the results of a review of 244 broad-scale DSM studies published between 2003 and 2021, where pH was found to be the best predicted property among the 12 mandatory soil properties for *GlobalSoilMap* (Chen et al., 2022). Indeed, the spatial variations in pH seems to be influenced more by drivers acting over long distances, such as lithology, which are more likely to be captured by calibration datasets of low spatial density, whereas spatial distributions of clay and SOC contents seem to be influenced more by drivers acting over short distances, such as local topography and agricultural practices. However, pH may have been predicted relatively poorly in Brittany because it usually varies little there, as most of its soils are acidic due to the bedrock, except on the northern coast, due to the presence of decarbonated Aeolian silts, which are parent materials of local marine origin, and the influence of amendments, particularly on vegetable field crops sensitive to acidity. Therefore, our results for Brittany should not be interpreted as revealing poor predictions.
- **Spatial density of the calibration dataset**: The model for SoilGrids 2.0 was calibrated with a less dense dataset (1 profile/600 km^2 globally and 1 profile /134 km^2 for mainland France) than that for GSM France (1 profile/15–19 km^2). Many recent studies have demonstrated the large impact of this spatial density on the performances of DSM models (Somarathna et al., 2017; Wadoux et al., 2019; Lagacherie et al., 2020; Loiseau et al., 2021). This lower density of calibration datasets may partly explain the lower performance. Thus, collecting additional soil data to densify calibration datasets is a key issue for improving DSM applications, especially if the goal is to capture local soil variations.
- **The scale of development of the DSM product**: Calibrating models over areas larger than those at which they are subsequently analyzed generates distortions in the modelled soil-landscape rules, which become visible when comparing predicted soil patterns. For example, in pH maps of Alsace, local soil spatial structures that result from soil-landscape relationships that are specific to a mapped area, such as the presence of a narrow zone of a basic fluviosol in the Rhine Valley (Fig. 6, blue ellipse), were not predicted by DSM products calibrated with datasets that cover larger extents. This occurs more frequently as the extent increases (global instead of national), as shown by a zone of neutral pH, which *GlobalSoilMap* France predicts well, but SoilGrids 2.0 does not (Fig. 6, black ellipse). The opposite can also occur; for example, SoilGrids 2.0 and *GlobalSoilMap* France clearly delineated an acidic area that does not correspond to any soil pattern recognised at the regional level (Fig. 6, red ellipse). This area, which corresponds to a forest, strongly influenced predictions of SoilGrids 2.0 and *GlobalSoilMap* France. A soil-landscape

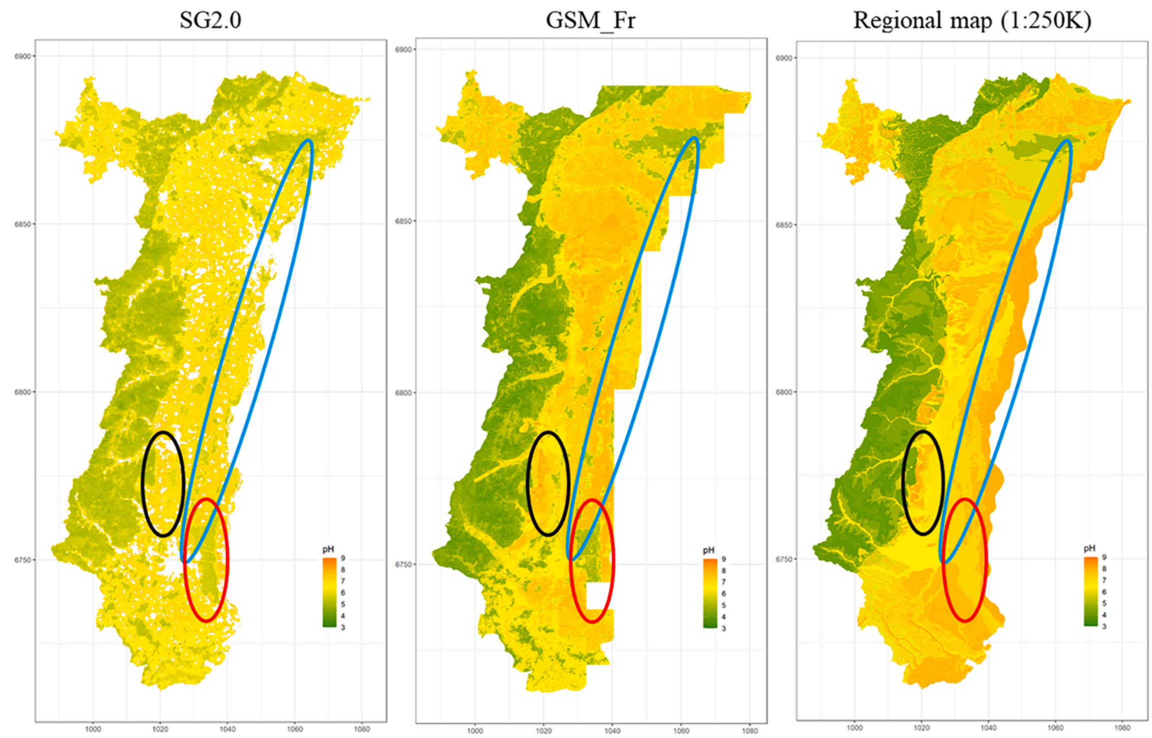

Fig. 6. Example of maps of pH in water in Alsace, France, showing prediction problems caused by differences in calibration set extents: high pH known to exist along the Rhine Valley not predicted by GSM products (blue ellipse), lower pH known to exist at the transition between the Alsace Plain and mountains not predicted by SoilGrids 2.0 (black ellipse), and a zone with low pH predicted by both SoilGrids 2.0 and GSM France but that is not an identified soil mapping unit (red ellipse). Units of geographic coordinates are kilometres.

relationship that was calibrated from soil profiles located outside of the area could have inferred a false soil pattern. Global models may not necessarily consider local variations because the variety of processes and soil-covariate relationships is too wide to model local features. More generally, the relative importance of covariates may depend on the size of the area of interest, because covariates represent the various factors controlling soil processes and these factors depend on the scale at which these processes occur. For example, the predominant effect of climate on SOC content is clear at the global scale, whereas SOC content may be driven more by land use and agricultural practices at more local scales.

5. Conclusions

This study evaluated DSM products obtained at global, national and regional levels over three French regions, at regional, intermediate and detailed scales. Topsoil values of pH in water, clay content and SOC content were evaluated for both point and SMU supports. The main lessons from these multiscale evaluations are the following:

- Users should evaluate DSM products themselves as soon as they focus on a study area much smaller than the initial area covered by the DSM products.
- Comparing DSM products to existing soil maps could greatly increase the potential of DSM product evaluations for small areas with too few soil measurements and for the consideration of areal-support uses of DSM products.
- Users should select a DSM product whose coverage by the calibration soil dataset and covariates best match those of their target study area. In our study, regional DSM products outperformed the national product, which outperformed the global product.
- None of the DSM products tested provided satisfactory predictions at a local scale. This raises the need to develop local DSM products to provide tools to improve local soil management.

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 on request.

Acknowledgements

The authors thank the CNES (French National Centre for Space Studies) for funding the Centre of Scientific Expertise (CES) Théia "Digital Soil Mapping", to which the authors belong. They also acknowledge the Regional Council of Brittany for financing the project "Sols de Bretagne", which contributed financially and technically to the study. A.R.d.F., P.L. and D.A. are members of the research consortium GLADSOILMAP, supported by the STUDIUM Loire Valley Institute for advanced research studies.

References

Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M., Hong, S. Y., Lagacherie, P., Lelyk, G., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M. D.L., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A., Zhang, G.L., 2014. *GlobalSoilMap*: Toward a fine-resolution global grid of soil properties. Adv. Agron. 125 https://doi.org/10.1016/B978-0-12-800137-0.00003-0.

Arrouays, D., Lagacherie, P., Hartemink, A.E., 2017. Digital soil mapping across the globe. Geoderma Regional 9, 1–4.

Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A.C., Morgan, C. L.S., Roudier, P., Poggio, L., Mulder, V.L., 2020a. Impressions of digital soil maps: The good, the not so good, and making them ever better. Geoderma Regional 20, e00255.

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., Laroche, B., Loiseau, T., Cousin, I., Lacoste, M., Ranjard, L., Toutain, B., Le Bas, C., Eglin, T., Bardy, M.,

Antoni, V., Meersmans, J., Ratié, C., Bispo, A., 2020b. Impacts of national scale digital soil mapping programs in France. Geoderma Regional 23, e00337.
Arrouays, D., Mulder, V.L., Richer-de-Forges, A.C., 2021. Soil mapping, digital soil mapping and soil monitoring over large areas and the dimensions of soil security - A review. Soil Security 5, 100018.
Batjes, N.H., Ribeiro, E., van Oostrum, A., 2019. Standardised soil profile data for the world (WoSIS Snapshot – September 2019). doi: 10.17027/isric-wdcsoils.20190901.
Bishop, T.F.A., Horta, A., Karunaratne, S.B., 2015. Validation of digital soil maps at different spatial supports. Geoderma 241–242, 238–249. https://doi.org/10.1016/j.geoderma.2014.11.026.
Bonfils, P., 1993. Carte pédologique de France à 1/100 000. Feuille de Lodève. Service d'étude des sols et de la carte pédologique de France. INRAE (Eds). [In French].
Brus, D.J., 2014. Statistical sampling approaches for soil monitoring. Eur. J. Soil Sci. 62, 779–791. https://doi.org/10.1111/ejss.12176.
Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62, 394–407. https://doi.org/10.1111/j.1365-2389.2011.01364.x.
Burrough, P.A., 2006. The display of Digital Soil data: 1976–2004. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping: An Introductory Perspective. Elsevier, Amsterdam, pp. 555–569.
Chen, S., Arrouays, D., Leatitia Mulder, V., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de-Forges, A. C., Walter, C., 2022. Digital mapping of GlobalSoilMap soil properties at a broad scale: a review. Geoderma 409, 115567.
Coulouma, G., Barthes, J.P., Robbez-Masson, J.M., 2008. Carte des sols au 1/25 000 de la basse vallée de la Peyne. Report and Map UMR LISAH (INRAE, France). [In French].
Dharumarajan, S., Kalaiselvi, B., Suputhra, A., Lalitha, M., Vasundhara, R., Kumar, K.S. A., Nair, K.M., Hegde, R., Singh, S.K., Lagacherie, P., 2021. Digital soil mapping of soil organic carbon stocks in Western Ghats, South India. Geoderma Regional 25, e00387.
Ellili Bargaoui, Y., Walter, C., Michot, D., Saby, N.P.A., Vincent, S., Lemercier, B., 2019. Validation of digital maps derived from spatial disaggregation of legacy soil maps. Geoderma 356, 113907.
Ellili-Bargaoui, Y., Malone, B.P., Michot, D., Minasny, B., Vincent, S., Walter, C., Lemercier, B., 2020a. Comparing three approaches of spatial disaggregation of legacy soil maps based on the disaggregation and harmonisation of soil map units through resampled classification trees (DSMART) algorithm. SOIL 6, 371–388. https://doi.org/10.5194/soil-6-371-2020.
Ellili-Bargaoui, Y., Walter, C., Michot, D., Lemercier, B., 2020b. Mapping soil properties at multiple depths from disaggregated legacy soil maps in the Brittany region. France. Geoderma Regional 23, e00342.
FAO and ITS, 2020. Global Soil Organic Carbon Map V1.5: Technical report. Rome, FAO. doi: 10.4060/ca7597en.
Favrot, J.C., Lagacherie, P., 1993. La cartographie automatisée des sols: une aide à la gestion écologique des paysages ruraux. Comptes Rendus de L'Academie d'Agriculture de France 79, 61–76 [In French].
Gomez, C., Coulouma, G., 2018. Importance of the spatial extent for using soil properties estimated by laboratory VNIR / SWIR spectroscopy : Examples of the clay and calcium carbonate content. Geoderma 330, 244–253. https://doi.org/10.1016/j.geoderma.2018.06.006.
Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152 (3–4), 195–207. https://doi.org/10.1016/j.geoderma.2009.06.003.
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., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., Bond-Lamberty, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE 12 (2), e0169748.
Heuvelink, G.B.M., 2014. Uncertainty quantification of GlobalSoilMap products, in: Arrouays, D., McKenzie, N.J., Hempel, J.W., Richer-de-Forges, A., McBratney, A.B (Eds.), GlobalSoilMap: Basis of the Global Spatial Soil Information System - Proceedings of the 1st *GlobalSoilMap* Conference2014. CRC press/Balkema, London, pp. 335–340.
INRAE InfoSol, 2014. Référentiel Régional Pédologique. Cahier des Clauses Techniques Générales (CCTG). De la réalisation à la labellisation des bases de données sols au 1/250 000. 27 p. https://www.gissol.fr/wp-content/uploads/2019/03/CCTG_RRP.pdf (last access: 12/20/2021). [In French].
INRAE InfoSol, 2021. DoneSol Version 3.5. Dictionnaire de données. 561 p. https://dw3.gissol.fr/fichiers/dictionnaire_donesol_igcs_3-15_31-03-2021.pdf (last access: 12/20/2021) [In French].
Jamagne, M., 1967 - Bases et techniques d'une cartographie des sols. Annales agronomiques, hors-série vol. 18, 142 p. [In French].
Koch, A., McBratney, A.B., Adams, M., Field, D., Hill, R., Crawford, J., Minasny, B., Lal, R., Abbott, L., O'Donnell, A., Angers, D., Baldock, J., Barbier, E., Binkley, D., Parton, W., Wall, D.H., Bird, M., Bouma, J., Chenu, C., Flora, C.B., Goulding, K., Grunwald, S., Hempel, J., Jastrow, J., Lehmann, J., Lorenz, K., Morgan, C.L., Rice, C. W., Whitehead, D., Young, I., Zimmermann, M., 2013. Soil Security: Solving the Global Soil Crisis. Global Policy 4 (4), 434–441. https://doi.org/10.1111/1758-5899.12096.
Lagacherie, P., McBratney, A.B., 2006. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping developments in soil science. Digital Soil Mapping—An Introductory Perspective, 316, IRD Montpellier, USJ-creen Beyrouth, pp. 3-22. doi: 10.1016/S0166-2481(06)31001-X.
Lagacherie, P., Arrouays, D., Bourennane, H., Gomez, C., Nkuba-kasanda, L., 2020. Analysing the impact of soil spatial sampling on the performances of Digital Soil Mapping models and their evaluation : A numerical experiment on Quantile Random Forest using clay contents obtained from Vis-NIR-SWIR hyperspectral imagery. Geoderma 375. https://doi.org/10.1016/j.geoderma.2020.114503.
Lagacherie, P., Buis, S., Constantin, J., Dharumarajan, S., Ruiz, L., Sekhar, M., 2021. Evaluating the impact of using digital soil mapping products as input for spatializing a crop model : The case of drainage and maize yield simulated by STICS in the Berambadi catchment (India). Geoderma 406. https://doi.org/10.1016/j.geoderma.2021.115503.
Laroche, B., Richer-de-Forges, A.C., Leménager, S., Arrouays, D., Schnebelen, N., Eimberck, M., Chenu, J.P., 2014. Le programme Inventaire Gestion et Conservation des Sols. Volet Référentiel Régional Pédologique. Etude et Gestion des Sols 21, 125-140. https://www.afes.fr/wp-content/uploads/2017/09/EGS_21_1_2110_Laroche_125_140_web.pdf [In French].
Leenhardt, M., Voltz, M., Bornand, M., Webster, R., 1994. Evaluating soil maps for prediction of soil water properties. Eur. J. Soil Sci. 45 (3), 293–301. https://doi.org/10.1111/j.1365-2389.1994.tb00512.x.
Loiseau, T., Arrouays, D., Richer-de-Forges, A.C., Lagacherie, P., Ducommun, C., Minasny, B., 2021. Density of soil observations in digital soil mapping: a study in the Mayenne region. France. Geoderma Regional 24, e00358.
Malone, B.P., Minasny, B., McBratney, A.B., 2017. Using R for Digital Soil Mapping. Springer, Cham (Switzerland), Progress in Soil Science, 262 p. doi: 10.1007/978-3-319-44327-0.
Marsmann, B.A., De Gruijter, J.J., 1986. Quality of soil maps. A comparison of soil survey methods in a sandy area. Soil Survey Papers, no. 15.
McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.
McBratney, A.B., Field, D.J., Koch, A., 2014. The dimensions of soil security. Geoderma 213, 203–213. https://doi.org/10.1016/j.geoderma.2013.08.013.
Minasny, B., McBratney, A.B., Hartemink, A.E., 2010. Global pedodiversity, taxonomic distance, and the World Reference Base. Geoderma 155, 132–139. https://doi.org/10.1016/j.geoderma.2009.04.024.
Minasny, B., McBratney, A.B., 2016. Digital soil mapping: a brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017.
Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57, 443–452. https://doi.org/10.2136/sssaj1993.03615995005700020058x.
Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Arrouays, D., 2016. *GlobalSoilMap* France: high-resolution spatial modelling the soils of France up to two meter depth. Science of the Total Environnement 573, 1352–1369. https://doi.org/10.1016/j.scitotenv.2016.07.066.
Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M.E., Papritz, A., 2018. Evaluation of digital soil mapping approaches with large sets of environmental covariates. SOIL 4, 1–22. https://doi.org/10.5194/soil-4-1-2018.
Odgers, N., McBratney, A., Minasny, B., Sun, W., Clifford, D., 2014. Dsmart: An algorithm to spatially disaggregate soil map units. In : Arrouays, D., McKenzie, N., Hempel, J., Richer-de-Forges, A., McBratney, A. (Eds.), GlobalSoilMap, CRC Press, pp. 261-266. doi: 10.1201/b16500-49.
Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.
QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.
R Core Team, 2019. R: A language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.
Richer-de-Forges, A.C., Arrouays, D., Poggio, L., Chen, S., Lacoste, M., Minasny, B., Libohova, Z., Roudier, P., Mulder, V.L., Nédélec, H., Martelet, G., Lemercier, B., Lagacherie, P., Bourennane, H., In press. Hand-feel soil texture observations to evaluate the accuracy of digital soil maps for local prediction of particle size distribution. A case study in central France. Pedosphere 32.
Richer-de-Forges, A.C., Baffet, M., Berger, C., Coste, S., Courbe, C., Jalabert, S., Lacassin, J.-C., Maillant, S., Michel, F., Moulin, J., Party, J.-P., Renouard, C., Sauter, J., Scheurer, O., Verbèque, B., Desbourdes, S., Héliès, F., Lehmann, S., Saby, N.P.A., Tientcheu, E., Jamagne, M., Laroche, B., Bardy, M., Voltz. M., 2014. La cartographie des sols à moyennes échelles en France métropolitaine. Etude et Gestion des sols 21, 25-36. https://www.afes.fr/wp-content/uploads/2017/09/EGS_21_1_2102_Richerdeforges_25_36.pdf [In French].
Rivière, J.M., Lemercier, B., Dupont, C., Rouxel, M., Walter, C., Edeline, P., Quidu, O., 2011. Carte pédologique de France à 1/100 000. Feuille de Janzé. INRAE (Eds). [In French].
Rivière, J.M., 1984. Carte pédologique de Pipriac au 1/25 000 et notice explicative. INRAE, Rennes, 117 p. [In French].
D.G. Rossiter L. Poggio D. Beaudette Z. Libohova How well does Predictive Soil Mapping represent soil geography? 2021 SOIL, under discussion An investigation from the USA 10.5194/soil-2021-80.
Saby, N.P.A., Swiderski, C., Lemercier, B., Walter, C., Louis, B.P., Eveillard, P., Arrouays, D., Goss, M., 2017. Is pH increasing in the non calcareous topsoils of France under agricultural management? A statistical framework to overcome the limitations of a soil test database. Soil Use Manag. 33 (3), 460–470.
Sanchez, P.A., Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.L., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vågen, T.G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., Zhang, G.L., 2009. Digital

Soil Map of the World. Science 325 (5941), 680–681. https://doi.org/10.1126/science.1175084.

Skidmore, A.K., Ryan, P.J., Dawes, W., Short, D., O'Loughlin, E., 1991. Use of an expert system to map forest soils from a geographical information system. Int. J. Geographical Information Sci. 5, 431–445. https://doi.org/10.1080/02693799108927867.

Somarathna, P.D.S.N., Minasny, B., Malone, B.P., 2017. More data or a better model? Figuring out what matters most for the spatial prediction of soil carbon. Soil Sci. Soc. Am. J. 81 (6), 1413–1426. https://doi.org/10.2136/sssaj2016.11.0376.

Song, X.-D., Wu, H.-Y., Ju, B., Liu, F., Yang, F., Li, D.-C., Zhao, Y.-G., Yang, J.-L., Zhang, G.-L., 2020. Pedoclimaticzone-based three-dimensional soil organic carbon mapping in China. Geoderma 363, 114145.

Tifafi, M., Guenet, B., Hatte, C., 2018. Large differences in global and regional total soil carbon stock estimates based on soilgrids, HWSD, and NCSCD: intercomparison and evaluation based on field data from USA, England, Wales, and France. Global Biochemical cycles 32 (1), 42–56. https://doi.org/10.1002/2017GB005678.

Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. Scotts Valley, California, USA: CreateSpace.

Vaysse, K., Lagacherie, P., 2015. Evaluating digital soil mapping approaches for mapping *GlobalSoilMap* soil properties from legacy data in Languedoc-Roussillon (France). Geoderma Regional 4, 20–30.

Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. Geoderma 291, 55–64. https://doi.org/10.1016/j.geoderma.2016.12.017.

Vincent, S., Lemercier, B., Berthier, L., Walter, C., 2018. Spatial disaggregation of complex soil map units at the regional scale based on soil-landscape relationships. Geoderma 311, 130–142. https://doi.org/10.1016/j.geoderma.2016.06.006.

Voltz, M., Arrouays, D., Bispo, A., Lagacherie, P., Laroche, B., Lemercier, B., Richer-de-Forges, A., Sauter, J., Schnebelen, N., 2020. Possible futures of soil-mapping in France. Geoderma Regional 23. https://doi.org/10.1016/j.geodrs.2020.e00334.

Wadoux, A.-C., Brus, D.J., Heuvelink, G.B.M., 2019. Sampling design optimization for soil mapping with random forest. Geoderma 355, 113913.

Webster, R., Burrough, P.A., 1974. Multiple discriminant analysis in soil survey. J. Soil Sci. 25, 120–134.

Wilks, D.S., 2011. Statistical Methods in the Atmospheric Sciences, Academic Press, 3rd Ed.