Geoderma 409 (2022) 115567

Contents lists available at ScienceDirect

Geoderma

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

Digital mapping of *GlobalSoilMap* soil properties at a broad scale: A review

Songchao Chen[a,b,c,*], Dominique Arrouays[a], Vera Leatitia Mulder[d], Laura Poggio[e], Budiman Minasny[f], Pierre Roudier[g,h], Zamir Libohova[i], Philippe Lagacherie[j], Zhou Shi[k], Jacqueline Hannam[l], Jeroen Meersmans[m], Anne C. Richer-de-Forges[a], Christian Walter[b]

[a] INRAE, Unité InfoSol, Orléans 45075, France
[b] SAS, INRAE, Institut Agro, Rennes 35042, France
[c] ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou 311200, China
[d] Soil Geography and Landscape Group, Wageningen University, Wageningen PO Box 47 6700 AA, The Netherlands
[e] ISRIC–World Soil Information, Wageningen PO Box 353 6700 AJ, The Netherlands
[f] School of Life and Environmental Sciences, Sydney Institute of Agriculture, The University of Sydney, NSW 2006, Australia
[g] Manaaki Whenua – Landcare Research, Private Bag 11052, Manawatū Mail Centre, Palmerston North 4442, New Zealand
[h] Te Pūnaha Matatini, A New Zealand Centre of Research Excellence, Private Bag 92019, Auckland, 1142, New Zealand
[i] USDA, Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE 68508, USA
[j] LISAH, University Montpellier, INRAE, IRD, Institut Agro, Montpellier 34060, France
[k] Institute of Agricultural Remote Sensing and Information Technology Application, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China
[l] School of Water, Energy and Environment, Cranfield University, Cranfield MK430AL, UK
[m] Department of Gembloux Agro-Bio Tech, University of Liège, Liège 4000, Belgium

ARTICLE INFO

Keywords:
Soil Information
GlobalSoilMap
Digital Soil Mapping
Predictive Models
Environmental Covariates
Model Performance and Uncertainty

ABSTRACT

Soils are essential for supporting food production and providing ecosystem services but are under pressure due to population growth, higher food demand, and land use competition. Because of the effort to ensure the sustainable use of soil resources, demand for current, updatable soil information capable of supporting decisions across scales is increasing. Digital soil mapping (DSM) addresses the drawbacks of conventional soil mapping and has been increasingly used for delivering soil information in a time- and cost-efficient manner with higher spatial resolution, better map accuracy, and quantified uncertainty estimates. We reviewed 244 articles published between January 2003 and July 2021 and then summarised the progress in broad-scale (spatial extent >10,000 km^2) DSM, focusing on the 12 mandatory soil properties for *GlobalSoilMap*. We observed that DSM publications continued to increase exponentially; however, the majority (74.6%) focused on applications rather than methodology development. China, France, Australia, and the United States were the most active countries, and Africa and South America lacked country-based DSM products. Approximately 78% of articles focused on mapping soil organic matter/carbon content and soil organic carbon stocks because of their significant role in food security and climate regulation. Half the articles focused on soil information in topsoil only (<30 cm), and studies on deep soil (100–200 cm) were less represented (21.7%). Relief, organisms, and climate were the three most frequently used environmental covariates in DSM. Nonlinear models (i.e. machine learning) have been increasingly used in DSM for their capacity to manage complex interactions between soil information and environmental covariates. Soil pH was the best predicted soil property (average R^2 of 0.60, 0.63, and 0.56 at 0–30, 30–100, and 100–200 cm). Other relatively well-predicted soil properties were clay, silt, sand, soil organic carbon (SOC), soil organic matter (SOM), SOC stocks, and bulk density, and coarse fragments and soil depth were poorly predicted ($R^2 < 0.28$). In addition, decreasing model performance with deeper depth intervals was found for most soil properties. Further research should pursue rescuing legacy data, sampling new data guided by well-designed sampling schemas, collecting representative environmental covariates, improving the performance and interpretability of advanced spatial predictive models, relating performance indicators such as accuracy and precision to cost-benefit and risk assessment analysis for improving decision support; moving from static DSM to dynamic DSM; and providing high-quality, fine-resolution digital soil maps to address global challenges related to soil resources.

* Corresponding author.
E-mail address: chensongchao@zju.edu.cn (S. Chen).

https://doi.org/10.1016/j.geoderma.2021.115567
Received 21 December 2020; Received in revised form 15 October 2021; Accepted 31 October 2021

1. Introduction

1.1. Background

In the 21st century, the world is experiencing grand challenges, such as large population increases, food security, land degradation, freshwater scarcity, threatened biodiversity, climate change, and sustainable development (FAO, 2011). Soil functions (e.g. producing biomass, acting as a carbon pool) are at the nexus of these challenges and are relevant to food production, climate regulation and adaptation, carbon sequestration, water filtering, and biodiversity preservation (McBratney et al., 2014; Adhikari and Hartemink, 2016; Keesstra et al., 2016). Consequently, soils have been directly linked to some of the United Nations (UN) Sustainable Development Goals (e.g. Goals 2, 3, 6, 7, 12–15) (Bouma, 2014; Keesstra et al., 2016; Evans et al., 2021; Lal et al., 2021).

To address these global and regional concerns, the demand for up-to-date and updatable soil information capable of supporting decisions at every level is increasing (Sanchez et al., 2009). Conventional soil mapping relies on soil survey; is labour intensive, time-consuming, and expensive; and heavily relies on experts' knowledge (e.g. McBratney et al., 2003; Grunwald et al., 2011; Sanchez et al., 2009; Arrouays et al., 2014a; Minasny and McBratney, 2016; Zhang et al., 2017). For a nonsoil scientist, these soil maps are difficult to interpret and use for decision-making in land management (Sanchez et al., 2009) because they are mostly based on taxonomic classification rather than quantifying soil properties. In response to these challenges, digital soil mapping (DSM, McBratney et al., 2003) has emerged over the last two decades to predict soil properties by integrating soil survey data, geographic information systems, geostatistics, terrain analysis, machine learning, remote sensing, and high-performance computing (Minasny and McBratney, 2016; Arrouays et al., 2017a).

As summarised by Lagacherie and McBratney (2006), "DSM is 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 from related environmental variables". This numerical or quantitative approach was trialled in the 1990s by McKenzie and Austin (1993), who proposed an environmental-correlation approach. The concept of DSM was later formalised as a framework for making digital soil maps, rooted in Jenny's five soil-forming factors (climate, organisms, relief, parent materials, and time; Jenny, 1941). The factors were developed into the Scorpan-SSPFe (soil spatial prediction function with spatially auto-correlated errors) framework (Eq. 1) by McBratney et al. (2003) for a quantitative spatial prediction.

$$S_a \text{ or } S_c = f(s, c, o, r, p, a, n) + e \quad (1)$$

where S_a and S_c represent the soil attributes and soil classes, respectively. s refers to soil information; c refers to climate; o refers to organisms, vegetation, fauna or human activity; r refers to relief; p refers to parent material; a refers to age or time factor; n refers to spatial or geographic position; and e are spatially correlated residuals. Most of these variables are spatially and temporally explicit.

1.2. Scientific activity

The use of DSM has increased rapidly since the 1st Global Workshop on Digital Soil Mapping organised in Montpellier, France, in 2004. The workshop was held biennially between 2006 and 2016 (Table 1). The idea of a global grid of soil properties emerged at the 2nd Global Workshop on Digital Soil Mapping held in Rio de Janeiro, in 2006, as a solution to the increasing need for soil information to address global challenges. This workshop culminated in the establishment and development of the GlobalSoilMap.net Project, the name of which was changed to *GlobalSoilMap* in 2012. This project aims to coordinate the production of a fine-resolution soil information grid for a limited set of soil properties. The main outcome of the project was the production of a consensus-based specification document for the global grid (details in Table 2, Arrouays et al., 2014a). A set of six standard depth intervals (0–5, 5–15, 15–30, 30–60, 60–100, 100–200 cm) was suggested for generating digital soil maps. The 1st *GlobalSoilMap* conference was held in Orléans, France, in 2013 (Arrouays et al., 2014b). Since then, DSM has evolved from a theoretical, academic focus to an operational status for delivering soil information to the scientific community and decision makers and policy makers (Scott et al., 2016; Grundy et al., 2020; Kidd et al., 2020) through the *GlobalSoilMap* project and Pillar 4 of the Global Soil Partnership (GSP) initiative (Minasny and McBratney, 2016; Arrouays et al., 2017a). In 2016, the International Union of Soil Sciences endorsed the Global Soil Map Working Group.

Several reviews of DSM have been conducted since 2009. Grunwald (2009) summarised the DSM progress on soil modelling from 90 articles published in two high-impact international soil science journals. Minasny et al. (2013) reviewed and discussed the advances in digital mapping of soil carbon from 40 articles, spanning field to national scales. Zhang et al. (2017) reviewed the progress in legacy soil data, environmental covariates, soil sampling, predictive models, and the applications of DSM products before 2017 and summarised their prospects. Minasny et al. (2019) reviewed 90 studies on peatland mapping in 12 countries or regions using the DSM. Lamichhane et al. (2019) reviewed DSM algorithms and covariates for soil organic carbon (SOC) mapping from 120 articles published between 2013 and 2019. Ma et al. (2019a) reviewed the relevance and synergy of pedology in DSM and discussed how DSM supports further advances in pedology. Wadoux et al. (2020) reviewed the application of machine learning algorithms in DSM. Piikki et al. (2021) performed a systematic review of the validation methods used in the DSM of continuous attributes.

In contrast with the listed reviews, this study focuses on the mapping of the 12 mandatory *GlobalSoilMap* properties (Table 2, Arrouays et al., 2014a) on a broad scale. In this case, a spatial extent greater than 10,000 km^2 is used to define broad-scale studies because this threshold corresponds to 83% of the countries and 99.99% of their total landmass. These soil property maps are produced using shared, international specifications that can be applied from country to globe. They relate to key soil information and can be used to address global environmental challenges. This review examines sampling strategies, environmental covariates, predictive models, efficient geospatial predictions, validation strategies, and uncertainty quantifications. Based on this review, we provide suggestions for further applications and developments in broad-scale DSM.

2. Methods

To assess the current progress in broad-scale DSM, we performed a

Table 1
International workshops on Digital Soil Mapping and *GlobalSoilMap*

No.	Year	Location	Event
1	2004	Montpellier, France	1st Global Workshop on Digital Soil Mapping
2	2006	Rio de Janeiro, Brazil	2nd Global Workshop on Digital Soil Mapping
3	2008	Logan, USA	3rd Global Workshop on Digital Soil Mapping
4	2010	Rome, Italy	4th Global Workshop on Digital Soil Mapping
5	2012	Sydney, Australia	5th Global Workshop on Digital Soil Mapping
6	2013	Orléans, France	1st *GlobalSoilMap* Conference
7	2014	Nanjing, China	6th Global Workshop on Digital Soil Mapping
8	2016	Aarhus, Denmark	7th Global Workshop on Digital Soil Mapping
9	2017	Moscow, Russia	2nd *GlobalSoilMap* Conference
10	2019	Santiago, Chile	Joint Workshop for Digital Soil Mapping and *GlobalSoilMap*
11	TBD	Goa, India	Joint Workshop for Digital Soil Mapping and *GlobalSoilMap*

*TBD: to be defined

Table 2
Overview of soil functional properties for *GlobalSoilMap* (Arrouays et al., 2014a).

No.	Property	Units	Precision*
1	Depth to rock	cm	N3.0
2	Plant exploitable (effective) depth	cm	N3.0
3	Organic carbon	g kg^{-1}	N4.0
4	pHx10		N3.0
5	Clay	g kg^{-1}	N3.0
6	Silt	g kg^{-1}	N3.0
7	Sand	g kg^{-1}	N3.0
8	Coarse fragments	m^3 m^{-3}	N3.0
9	Effective cation-exchange capacity	$mmol_c$ kg^{-1}	N4.0
10	Bulk density (whole soil)	Mg m^{-3}	N3.1
11	Bulk density (fine earth)	Mg m^{-3}	N3.1
12	Available water capacity	mm	N4.0

* The notation used to describe precision (e.g., N3.0) is interpreted as N = number, 3 = length of number, 0 = number of decimal digits.

literature review of DSM articles published between January 2003 and July 2021. On 19 July 2021, Web of Science and Scopus were queried using several expressions applied to the topic (i.e. title, abstract, and keywords) of the articles. The search expressions were "digital soil mapping" OR "globalsoilmap" OR "soilgrids" OR "soil-landscape modelling" OR "soil predictive modelling" OR "predictive soil mapping" (Scull et al., 2003). Because this literature review focuses on broad-scale DSM, we retained only the articles that had a spatial extent larger than 10,000 km^2 and with at least one soil property of interest listed in *GlobalSoilMap* specifications.

We found 244 relevant articles published in English and recorded in the Web of Science and Scopus. We then extracted a list of variables to derive systematic plots for the results section. Table 3 shows the 26 variables recorded for this review.

3. Current status and trends

3.1. Frequency of articles per year and journal

Fig. 1a shows the annual number of articles published between January 2003 and July 2021. A few articles addressed broad-scale DSM prior to 2011 (no more than six per year), and a significant increase in publications was observed after 2012. After the publication of the 1st *GlobalSoilMap* Conference book in 2014 (13 articles), we observed a small increase in the number of publications for that year (i.e. 32 articles), with 19 articles from journal publications. Between 2016 and 2018, the number of annual publications was stable (between 22 and 24), whereas in 2019 and 2020, there was an increase in the number of publications (i.e. 37 and 31 articles). Seventeen articles were published from January to July 2021.

The number of journals relevant to broad-scale DSM is shown in Fig. 1b. Fifty-seven journals were identified, and journals with only one publication were classified as "Others." Geoderma had the greatest number of broad-scale DSM publications, with 72 articles. Science of the Total Environment and Geoderma Regional ranked second, with 17 publications each. In addition, 24.2% of the articles (59 of 244) were open access, with the majority published after 2014 (not shown in Fig. 1b).

Table 3
List of variables included in the review process

Variable
Publication year; Journal; Authors; Open access[a]; Country/sub-country;
Spatial extent (km^2); Soil sampling year; Number of sampling sites;
Soil sampling density; Soil sampling strategy[b]; Soil sampling elementary volume[c];
Validation strategy[d]; Number of sites for validation;
Spatial resolution of map; Predictive model; Scorpan factors[e];
Number of covariates; Soil property of interest; Maximum soil depth of interest;
Soil depth interval of interest; Depth standardization[f]; Model performance indicators;
Uncertainty quantification[a]; GlobalSoilMap product[a]

[a] Boolean; [b] Probability sampling, non-probability sampling, mixture or not provided; [c] Soil genetic horizons or fixed depth intervals; [d] Data split, independent validation or not provided; [e] Soil, climate, organisms, relief, parent materials, age, position; [f] Depth weighted, equal area spline or not provided.

3.2. Spatial distribution of articles

Fig. 2 presents the sum of articles at the national/sub-national scale, from which the continental or global studies were excluded (three in Africa, one in North America, one in North and South America, one in the Middle East, 11 in Europe, and seven worldwide). The figure shows that DSM has been used to deliver soil information worldwide. Among these countries, China, France, Australia, and the United States were the most active, with the highest number of publications at 43, 29, 25, and 24, respectively. We classified these publications according to continents: 35.2% of the articles (86 articles) were published in Europe, representing 6.8% of the total landmass (Antarctica included). Representing 29.5%, 5.9%, and 16.5% of the total landmass, Asia, Oceania, and North America accounted for 24.2%, 10.7%, and 13.1% of the articles, respectively. Africa and South America published less than 10% of the publications (17 and 13 articles), although they accounted for 20.4% and 12% of the total landmass, respectively.

3.3. Soil sampling

Trends in the spatial extent and soil sampling density are presented in Fig. 3. The sampling density varied from 1 to 0.0001 sample km^{-2} for the 244 articles. Sampling density tended to decrease with increasing spatial extent. This result also showed that the articles reported with high sampling density (>0.1 sample km^{-2}) were mostly from Europe (e.g. Denmark, Sweden, Italy, Belgium, Northern Ireland, France, Estonia, the Netherlands, Scotland).

Fig. 4 shows the soil sampling year reported in 244 articles, 40.2% (98) of which did not provide relevant information when producing digital soil maps. The sampling year was a long time span, and the oldest samples were from the 1920s. Among the 146 articles reporting a sampling year, 24.7% had soil samples collected before the 1980s, and 50% used soil data collected after the 2000s.

The soil sampling design used for broad-scale DSM is shown in Fig. 4. More than half (52.4%) of the articles did not report relevant information, partially because soil databases were compiled from various sources of legacy soil information for different purposes; thus we can reasonably assume that they are nonprobability sampling designs (i.e. using the legacy data available). Fig. 4b also shows that probability and nonprobability sampling were used in 20.2% and 22.6% of the articles, respectively. A small proportion of the studies (4.8%) had a soil database that included samples collected from both probability and non-probability sampling designs. For the most recent sampling campaigns (data not shown), there is a trend to use probability sampling more frequently and focus on filling geographical gaps based on legacy soil data or maps (e.g. uncertainty-guided sampling).

3.4. Soil property and maximum depth of interest

Soil properties mentioned in the *GlobalSoilMap* specifications are shown in Fig. 5a. SOC and soil organic matter (SOM) content had the highest number of articles (126). If the 63 articles on mapping SOC stock were added to this number, 77.5% of the articles (189 of 244) addressed SOC or SOM. They were followed by articles on soil particle size fraction, with 69, 57, and 39 articles related to clay, sand, and silt mapping, respectively. Soil pH was another soil property of high interest (49 articles). Other soil properties, namely, bulk density (BD), available water capacity (AWC), soil depth (SD), and cation-exchange capacity (CEC), were relatively less predicted in the broad-scale DSM.

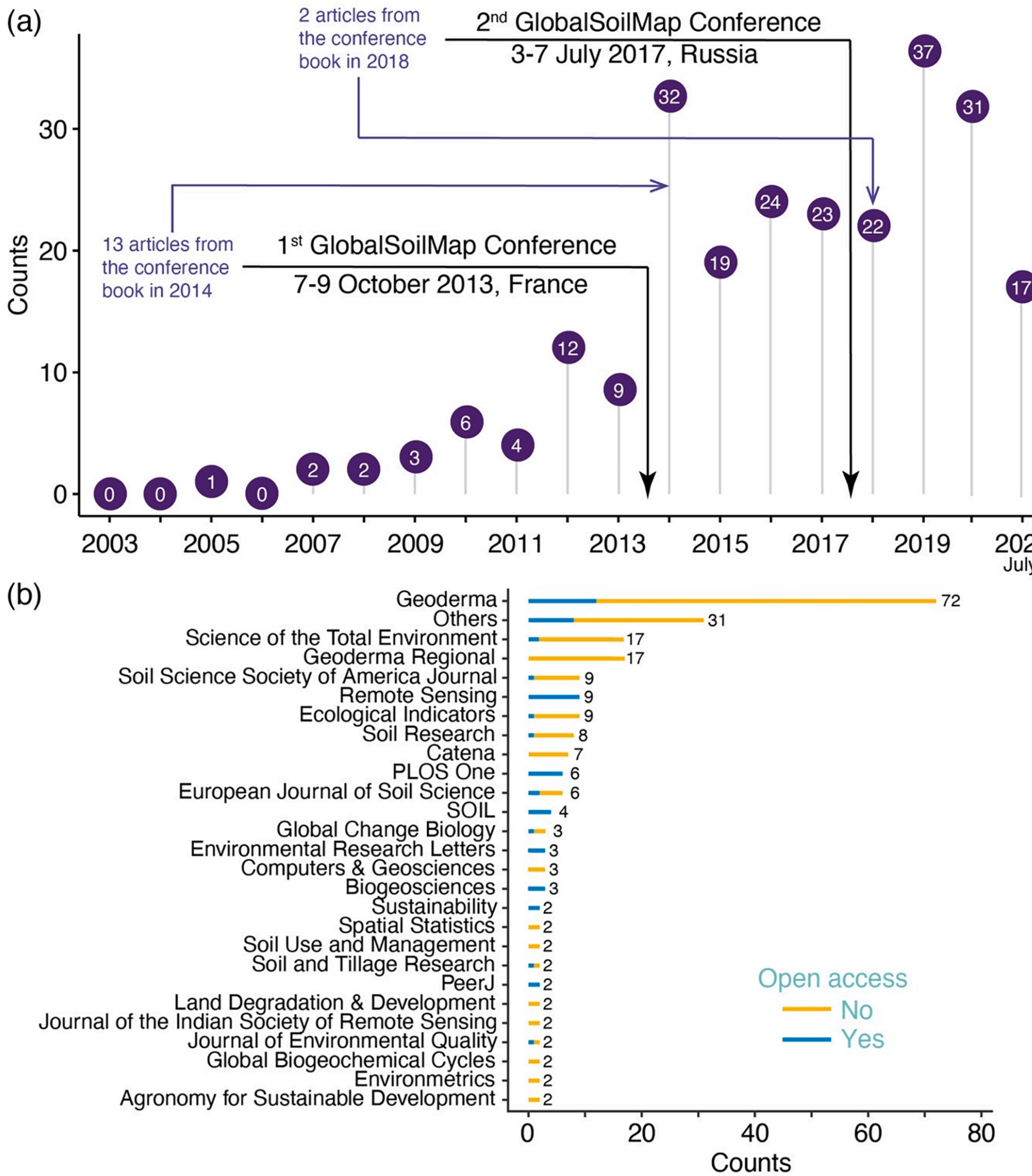

Fig. 1. Number of articles in digital soil mapping at a broad scale from January 2003 to July 2021 (a) and number of articles per journal (b). In (a), the first and second *GlobalSoilMap* Conferences were held in 2013 and 2017, resulting in two conference books published in 2014 and 2018. The articles in these books are included. In (b), whether the article is open access or not is indicated by colour. The category "Others" contains the journals with only one record among all the articles in digital soil mapping at a broad scale.

The maximum soil depth of interest for the soil maps produced is shown in Fig. 5b. Thirty-seven articles (15.2%) did not report the exact depth of the produced maps. These articles either provided broad categorical descriptors of soil layers (i.e. topsoil and subsoil) or reported soil genetic horizons (i.e. A horizon, B horizon). More than half the studies focused on topsoil mapping with a maximum depth of 30 cm, and 21.7% of the studies produced digital soil maps for soil depths greater than 100 cm. Among these studies, 98% reported values to a depth of 200 cm.

3.5. Environmental covariates and spatial resolution of map products

Fig. 6 shows the frequency of environmental covariates (classified by Scorpan factors) used in broad-scale DSM. Relief was used in 204 articles and was the top covariate. The organisms and climate covariates were ranked second and third in 201 and 191 articles, respectively. They were followed by soil and parent material, used in 159 and 106 articles, respectively. Position (geographic coordinates) was used in 20 articles, and age was rarely used (3 articles).

Fig. 7 presents the resolution of the maps produced under different spatial extents. The map resolution ranged from 10 m to 12 km. There

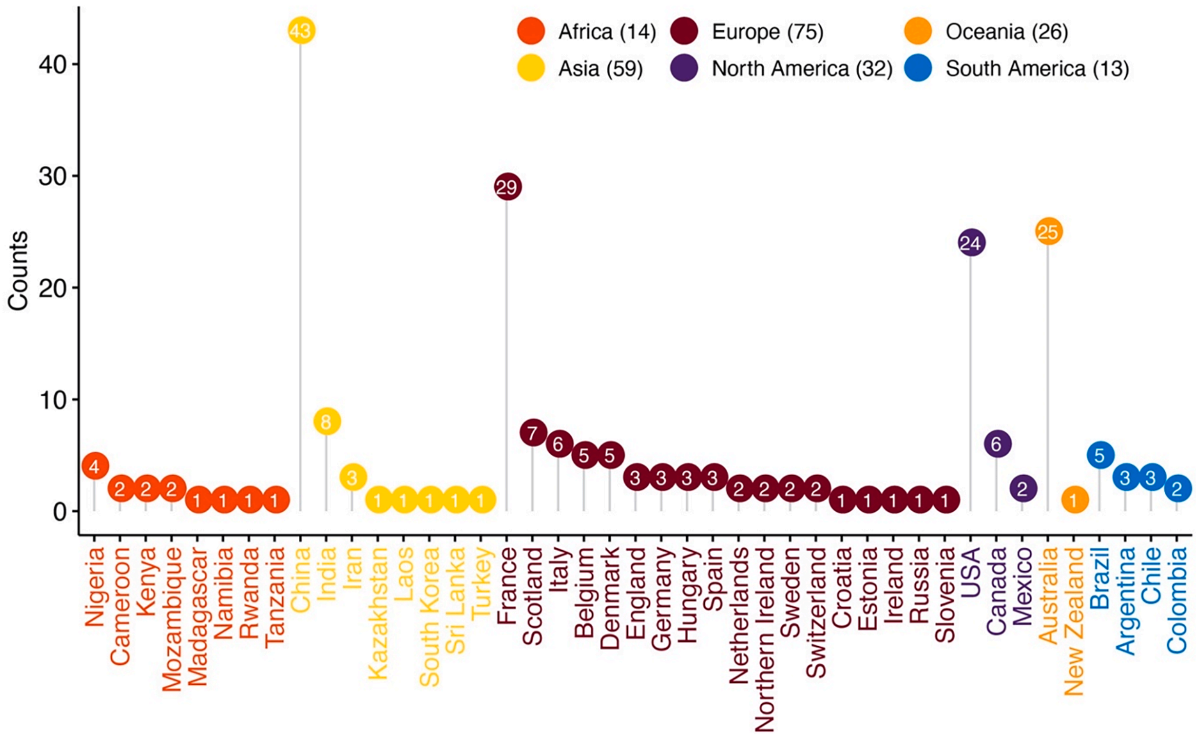

Fig. 2. Numbers of articles by country at national/sub-national scale with their sums at continental scale. Articles at continental and global scales (11 in Europe, 3 in Africa, 1 in North America, 1 in North and South America, 1 in the Middle East, and 7 for Globe) are excluded. The proportions of the countries and landmass covered by digital soil mapping for each continent are indicated.

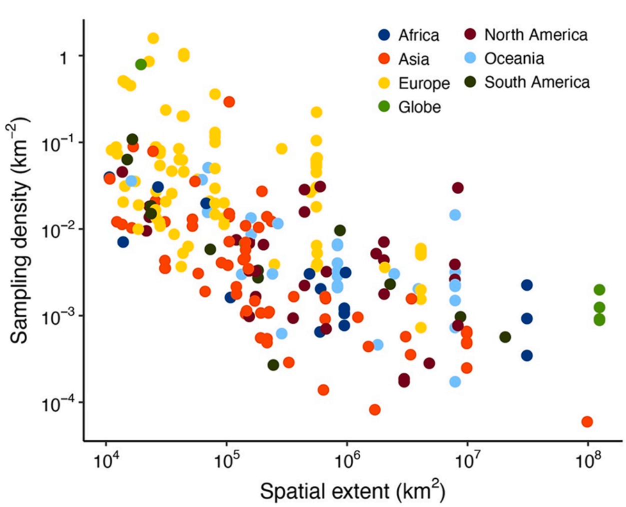

Fig. 3. Trend between spatial extent and sampling density (number of soil sampling sites km^{-2}).

was no clear tendency between the resolution of the produced map and spatial extent. Indeed, 70.9% of the articles produced digital soil maps with a rather fine spatial resolution ($\leq$250 m), which is close to the spatial resolution (3 arc-second or approximately 90 m) mentioned in the *GlobalSoilMap* specifications.

3.6. Predictive models

Frequencies of the predictive models are shown in Fig. 8a. The predictive models were divided into seven groups: (1) geostatistics, (2) linear, (3) regression kriging, (4) nonlinear, (5) nonlinear plus kriging of residuals, (6) disaggregation, and (7) model averaging. The geostatistical model (1) refers to pure interpolation in which no covariates are used (e.g. ordinary kriging, block kriging). A linear model (2) is based on a deterministic regression model using linear relationships (e.g. multiple linear regression, partial least squares regression). Regression kriging (3) combines linear models (deterministic) with geostatistics (residuals). Nonlinear models can be deterministic only (4) (e.g. cubist, random forest, support vector machine, neutral network, Bayesian model, Gaussian process regression, integrated nested Laplace approximation with stochastic partial differential equation) or combined with a geostatistical component (5).

The disaggregation model (6) refers to the spatial disaggregation by downscaling soil map units into soil classes or types (e.g. DSMART, Odgers et al., 2014a). Model averaging (7, or ensemble modelling) refers to soil property predictions that combine multiple models or map products.

As indicated in Fig. 8a, the number of articles related to the broad-scale DSM was low prior to 2013, dominated by geostatistics and linear models. Since 2014, a large increase was found for nonlinear models with/without kriging, which peaked in 2019 and 2020 (i.e. 30 and 28 articles, respectively). Disaggregation and model averaging were used in broad-scale DSM since 2012 and 2015, respectively. Usage of model averaging showed an increasing trend, reaching a peak of four articles in 2020, and disaggregation was relatively less used in broad-scale DSM.

We grouped the papers into 2D, 2.5D, and 3D models. A 2D model indicates that the predictive model does not account for vertical variation in soil properties; thus, only one soil layer is mapped. A 2.5D (or pseudo-3D) model uses harmonised soil properties for each specified depth interval (e.g. depth-weighted, equal-area spline), and then a predictive model is fitted independently for each depth interval. A 3D model uses depth as a covariate in modelling or predicting the parameters of a depth function spatially. The results showed that many studies (63.6%) applied 2D models, and a small percentage (8%) used 3D

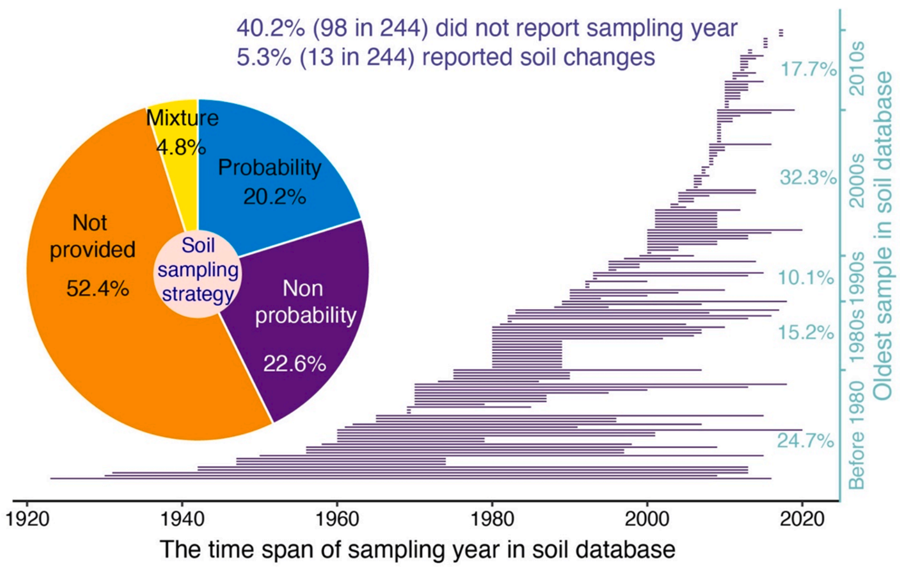

Fig. 4. The time span of soil sampling year and soil sampling strategy (pie plot) reported for the 244 articles. For soil sampling year, it excludes 40.2% of the articles (98 of 244) that did not report sampling year. The percentage of oldest samples in the soil database shown on the right (y axis) is proportional to the articles reporting sampling year. In soil sampling strategy (pie plot), the category "Mixture" represents the articles compiling soil data from both probability and non-probability sampling.

models for predictive mapping.

3.7. Validation strategy, performance indicators, and uncertainty estimation

Data splitting was the most frequent validation strategy used to evaluate map accuracy (86.9 %). Data splitting comprises single random holdback (i.e. divide data into the calibration and validation sets only once), stratified random holdback, cross-validation (i.e. *k*-fold cross-validation, leave-one-out cross-validation, repeated *k*-fold cross-validation), and other strategies (e.g. bootstrap, Kennard-Stone, nearest-neighbour distance), which accounted for 45.5%, 4.5%, 37.7%, and 2.9%, respectively (nine articles applied two validation strategies). Only 8.9% of the studies used independent validation. However, 5.8% of the articles did not show any validation procedure for evaluating the prediction performance of digital soil maps.

Fig. 8b presents the performance indicators used in the model evaluation. The coefficient of determination (R^2) and root mean square error (RMSE) were the most frequently used performance indicators (more than 180 articles). In third place (87 articles) was the indicator mean error (ME). Lin's concordance correlation coefficient (CCC) and mean absolute error were used for model evaluation in 55 and 53 articles, respectively. Other indicators, namely, the prediction interval coverage probability (PICP, assessing the quality of uncertainty estimates), the ratio of performance to deviation (RPD), adjusted R^2 (R^2_{adj}), the mean and median of standardised squared prediction errors ($\overline{\theta}$ and θ), the spatial structured variance ratio (SSVR), and the ratio of performance to inter-quantile (RPIQ), were used less frequently.

More than 56% of the studies provided uncertainty estimates (i.e. uncertainty maps) associated with the predicted soil properties produced by DSM.

3.8. Model performance of soil properties

In this study, we used R^2 was used to assess and compare the prediction performance across studies because R^2 was the most frequently used indicator (80.3%) in broad-scale DSM studies. R^2 is also dimensionless, which allows a comparison of the accuracy of different soil properties. However, the method for calculating R^2 differed by the author: some used the R^2 calculated by the square of the correlation coefficient between observations and predictions, and others used the R^2 by using the percentage of explained variance of the target soil property. In our study, we focused on R^2 expressed as the percentage of explained variance to represent performance. For studies that did not report R^2 expressed by percentage of explained variance but had Lin's CCC, we attempted to estimate R^2 by the CCC collected in this study. The fitted 3rd order polynomial regression model between the CCC and R^2 explained 90% of the variance ($R^2 = 0.90$); therefore, it is suitable to estimate R^2 from CCC using this model.

Fig. 9 presents the performance of different soil property maps using DSM that are grouped in three depth intervals (0–30, 30–100, and 100–200 cm), except for SD. Most soil properties showed a decreasing trend in performance with increasing depth intervals, except for pH, AWC, and CEC. Additionally, the performance of 3D models (indicated by R^2) for most soil properties was greater than that of the 2.5D models, except for the BD. Soil pH had the best performance with a median R^2 of 0.60, 0.63, and 0.56 at depths of 0–30, 30–100, and 100–200 cm, respectively. A significant decrease in performance (0.49, 0.28, and 0.14) was observed for SOC/SOM over the three sequential depth intervals. SOC stocks had a similar performance as a median R^2 of 0.51, 0.35, and 0.06 at 0–30, 30–100, and 100–200 cm depth intervals, respectively. BD also decreased with depth, with a median R^2 of 0.56 at 0–30 cm, to 0.47 at 30–100 cm, and finally to 0.40 at 100–200 cm. Soil particle size fractions (i.e. clay, silt, and sand) had a median R^2 of approximately 0.50 in 0–30 cm, 0.40 in 30–100 cm, and 0.29 in 100–200 cm. Performance for AWC was higher at 0–30 cm ($R^2 = 0.34$) and 30–100 cm ($R^2 = 0.42$) than that ($R^2 = 0.27$) at 100–200 cm. CEC had a median R^2 between 0.34 and 0.40 at three depth intervals. In general, coarse fragments and SD were poorly predicted, with a median R^2 of less than 0.28.

3.9. *GlobalSoilMap* products

Fig. 10 shows the frequency of *GlobalSoilMap* products, which fulfilled two requirements:

(1) The depth intervals defined by *GlobalSoilMap* (i.e. 0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm).

(2) Uncertainty estimates was provided associated with digital soil map predictions.

In total, 21.7% of the articles (53 of 244) were classified as *GlobalSoilMap* products, all of which were published after 2012. An increasing trend was found for *GlobalSoilMap* products, which increased from three articles in 2012–2013 to 11 articles in 2014–2015, and reached 13 articles biennially from 2016 to 2021.

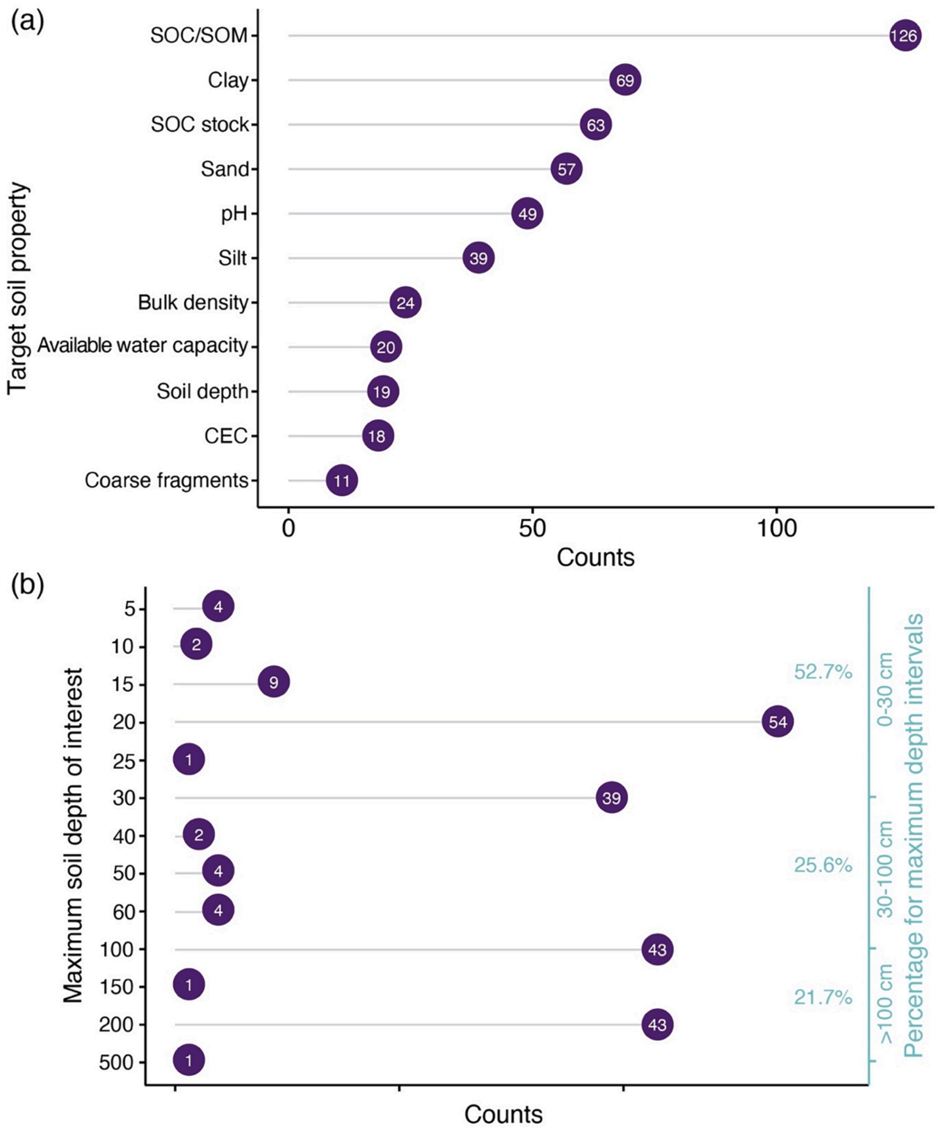

Fig. 5. Frequency of target soil properties (a) and maximum soil depth of interest for the produced soil property maps (b). Note that only 12 soil properties recommended by *GlobalSoilMap* specifications are listed. SOC stock is listed separately from SOC/SOM content as it integrates other soil properties such as bulk density, coarse fragments and soil depth. For maximum soil depth of interest (b), the counts exclude 15.2% of the articles (37 of 244) with missing reporting soil depth. The percentage of maximum depth intervals shown on the right (y axis) is proportional to the articles reporting maximum depth of interest.

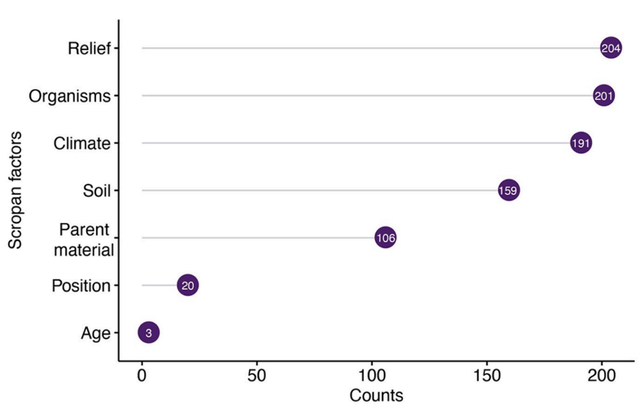

Fig. 6. Frequency of environmental covariates used in broad-scale digital soil mapping.

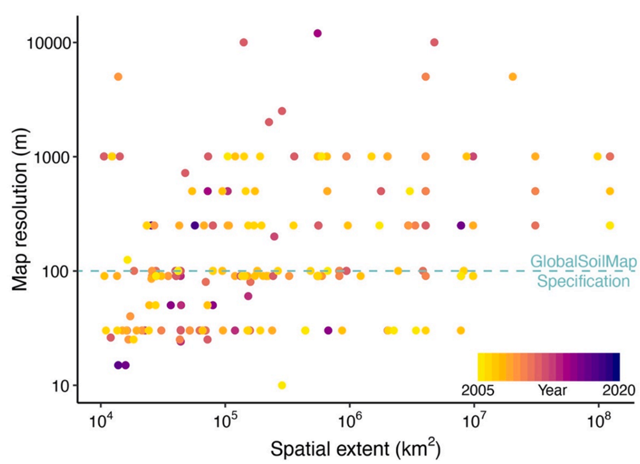

Fig. 7. Relationship between map resolution and spatial extent.

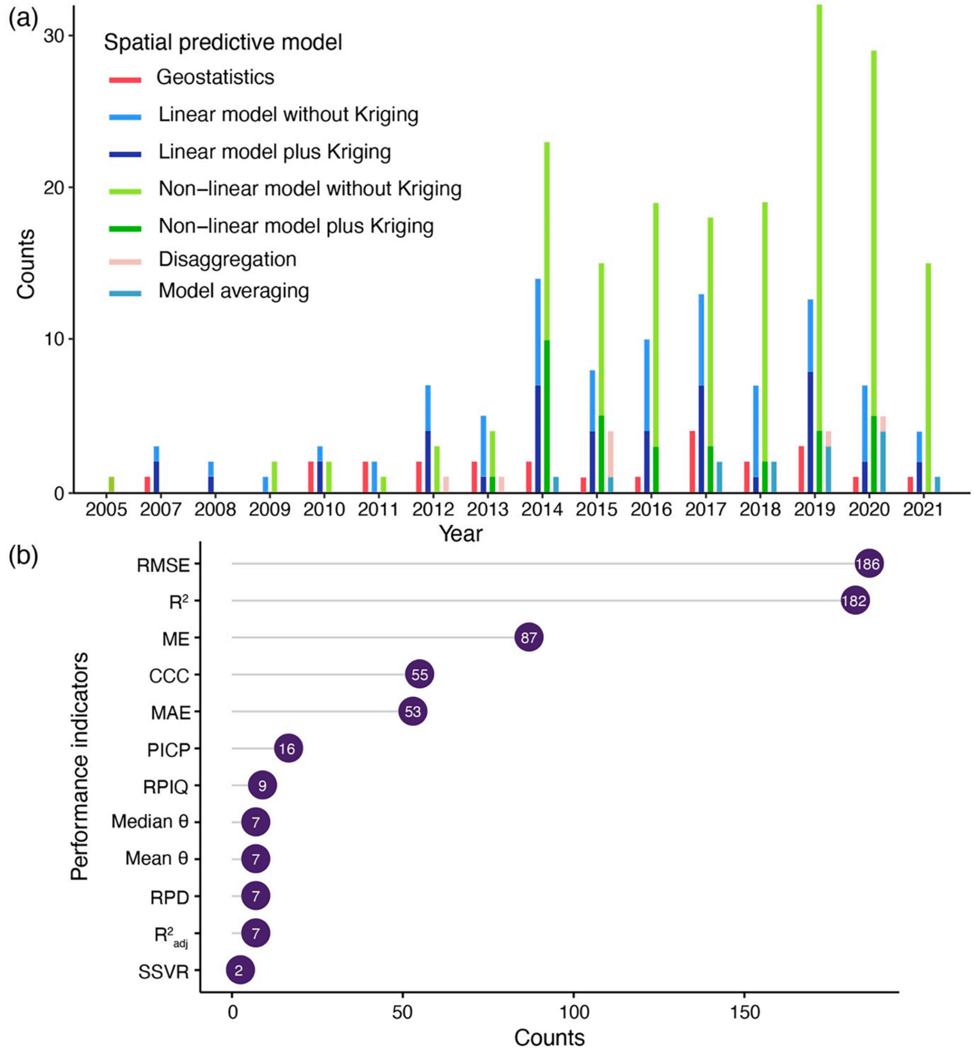

Fig. 8. Type of prediction models (a) and frequency of performance indicators in model evaluation (b). Note that many articles compared the model performance using several types of spatial prediction models (a) and therefore the sum of all the counts is far more than the total number of articles.

4. The future, remaining challenges, and perspectives

This section is presented in four parts: i) trends in current and future products (i.e. publications and maps), ii) how to use legacy data and gather new soil information, iii) issues in modelling and spatial prediction, and iv) performance evaluation and uncertainty estimation.

4.1. Current trends

4.1.1. Diversity challenges

Notably, 43.4% of the articles were published in three journals (i.e. Geoderma, Science of the Total Environment, and Geoderma Regional). The benefit of this phenomenon is the exchange of ideas in a limited environment, making it easier for the DSM community to use state-of-the-art methodologies and approaches and keep abreast of the most recent updates. Consequently, we observed that the studies only attempted to communicate with soil science communities. Therefore, expanding the use of DSM products for soil property maps by other disciplines and, ultimately, stakeholders require publication and communication in multidisciplinary and open-access journals. Indeed, the DSM community is moving toward this target, for example, the diversification of journals increased from less than four before 2011 to more than 14 after 2018, expanding into nonsoil science journals and open-access articles.

Four countries (China, France, Australia, and the United States) have published nearly half the articles related to broad-scale (>10,000 km^2) DSM for the required soil properties. This result may be due to their huge land mass and/or their state of progress in delivering *GlobalSoilMap* products. In the future, we expect articles on broad-scale DSM to originate from other countries as part of the *GlobalSoilMap* initiative.

We may expect that the links between some institutes will be reinforced and expanded through the activities of the *GlobalSoilMap* Working Group and future workshops and meetings. In addition, we expect that some exchanges of researchers and DSM practitioners will reinforce these links and that capacity building will result in new partners. We also posit that the cluster constituted by the UN-FAO GSP will grow as they attempt to deliver products for other soil properties in addition to SOC, and if the effort in training and capacity building is pursued.

4.1.2. Increase in broad-scale digital soil maps

We observed a large increase in the number of publications over time. This increase began at the end of the 2000s, which may be partially attributed to the article in Science by Sanchez et al. (2009) and the activities of the *GlobalSoilMap* consortium. A peak was observed in 2014, corresponding to the publication of the 1st *GlobalSoilMap* Conference book. If we exclude this exceptional peak, the number of publications continues to increase. This increase occurs because some countries (e.g. Cameroon, China, Chile, Denmark, France, Hungary, India, Nigeria,

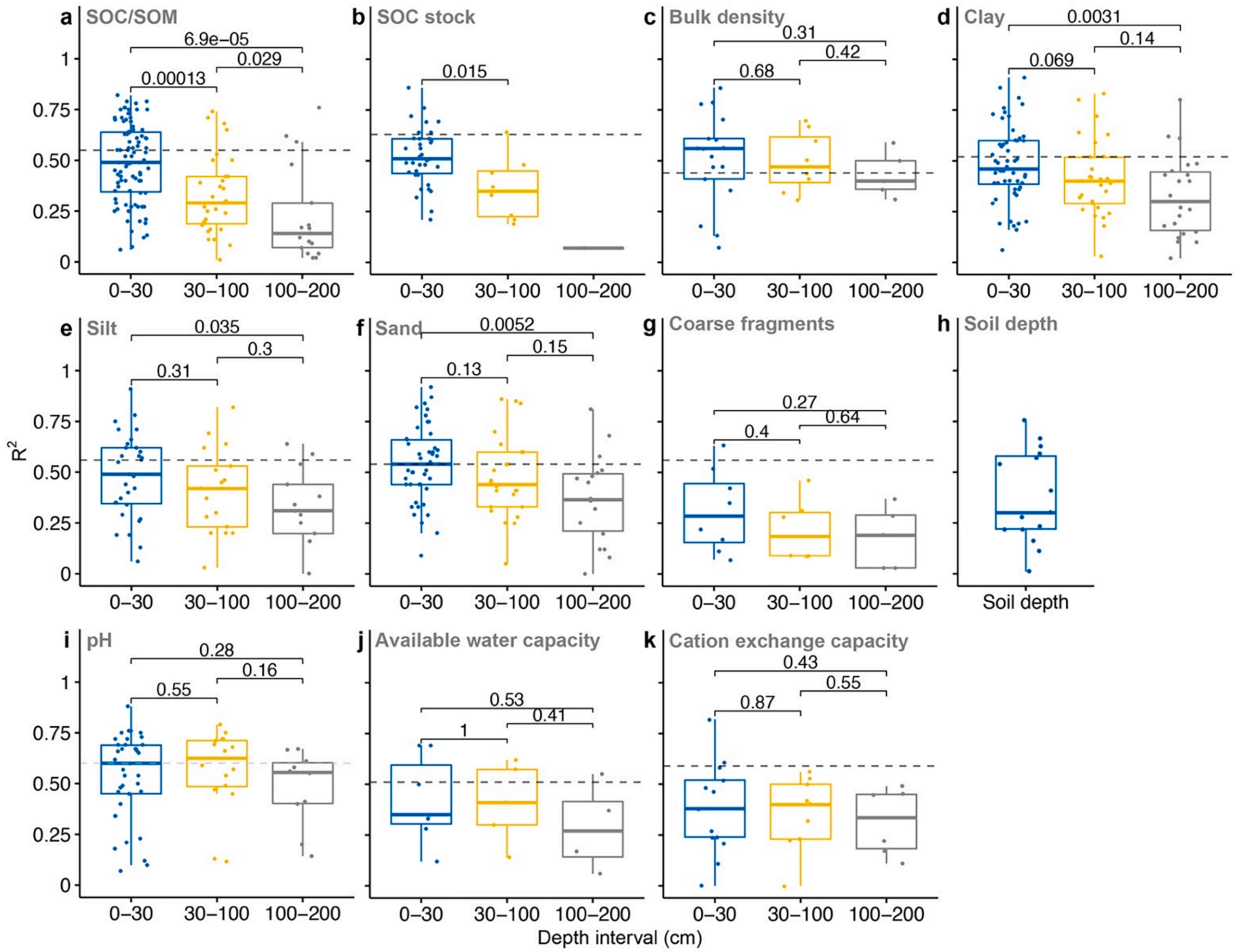

Fig. 9. Performance of DSM products for different soil properties at three depth intervals (0–30, 30–100 and 100–200 cm, except for soil depth). The dash horizontal grey line indicates the median performance of DSM products for relevant soil properties using 3D models (not built on specific depth intervals separately). The p values between different depth intervals are calculated by Mann-Whitney test.

Scotland, New Zealand) published complete *GlobalSoilMap* products (the United States, Australia) and released some selected *GlobalSoilMap* soil properties (Adhikari et al., 2013; Akpa et al., 2014; Viscarra Rossel et al., 2015; Mulder et al., 2016; Padarian et al., 2017; Poggio and Gimona, 2017; Ramcharan et al., 2018; Dharumarajan et al., 2019; Laborczi et al., 2019; Liang et al., 2019; Roudier et al., 2020; Silatsa et al., 2020; Reddy et al., 2021; Liu et al., 2021). What is likely is that the increasing trend in the number of maps will continue because some required soil properties remain under development and more countries (e.g. The Netherlands, Canada, Russian Federation, Iran, some countries in Africa, many countries in South and Central America) are joining the *GlobalSoilMap* project or similar national/continental initiatives (Guevara et al., 2018; Pfeiffer et al., 2020; Zeraatpisheh et al., 2020; Helfenstein et al., 2021). This trend may be amplified by the increasing number of available environmental covariates (e.g. satellite imagery) with time, and efforts from the GSP to apply DSM methods to deliver soil property products. The strength of the trend is not certain, however, because it may depend on national soil mapping programme commitments and/or political situations or conflicts.

4.2. How to use legacy data and acquire new soil information?

4.2.1. Current sources for input soil data

In most cases, the soil information used from DSM is from legacy data; thus, the sampling design cannot be controlled. Furthermore, for large territories, soil data may be assembled from local soil surveys, which can result in clustered sampling and unharmonised databases. Although the technical and theoretical challenges for selecting the appropriate soil sampling schemes are well elaborated in the DSM community, the practical and operational challenges regarding how to implement them at global scales are not well defined. Additionally, most future input data will probably remain beyond the researchers' control, either because the data will be historical or because new data collected by others will mainly be used (i.e. with purposes other than mapping or performing some type of statistical sampling).

Soil data used in broad-scale DSM dates back to the 1920s. Consequently, most of the available soil data used for modelling soil properties do not consider changes with time because of the scarcity of soil information (Fig. 4a). Although this practice may be acceptable for relatively stable soil properties, such as SD and particle size fractions, because that there is no evidence of erosion or redeposition in the short term (i.e. 100 years), it introduces large uncertainty for dynamic soil properties that may change over short time scales (from years to decades), such as topsoil SOC, pH, and CEC.

4.2.2. Improvement of imperfect soil data

Considering the imperfections (e.g. spatial coverage, sampling density, strategy) of data sources, developing techniques that mitigate the

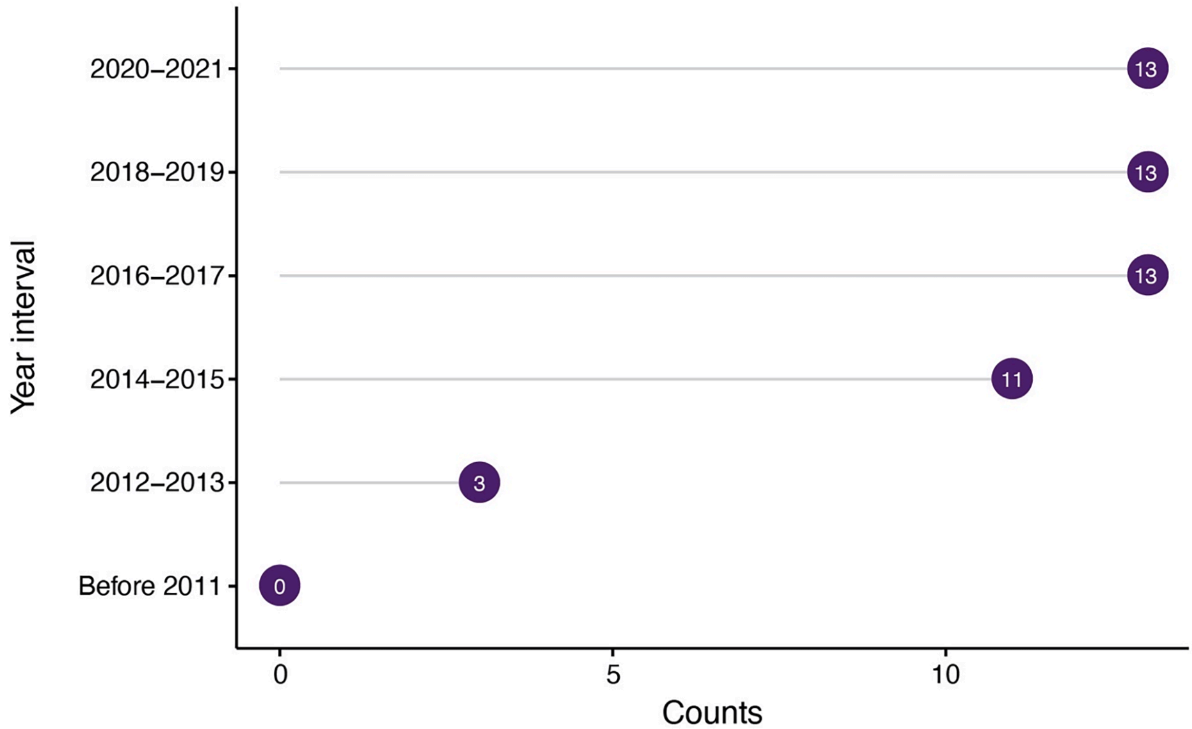

Fig. 10. Frequency of *GlobalSoilMap* like articles, which produced soil maps at *GlobalSoilMap* standard depths (0–5, 5–15,15–30, 30–60, 60–100, 100–200 cm) with quantified uncertainty estimates at a fine resolution (from 25 m to 500 m).

unevenness and clustering of soil data is necessary. One solution to improve imperfect soil data is to correct and harmonise input data. A method to solve discrepancies in data density is declustering legacy data, either spatially or using covariates. In other words, clustered data can be weighted differently according to their spatial or covariate clusters and soil data collected from multiple sensors, platforms, and surveys (Richer-de-Forges et al., 2017; An et al., 2018; Liu et al., 2020a; Taghizadeh-Mehrjardi et al., 2020b). Another solution is to add new soil observations by using specific sampling techniques that correct the initial sampling design (Carré et al., 2007).

In solving the inconsistency of soil sampling time, two solutions may be helpful: (1) only use the soil data within a given period (i.e. 2001–2010) in model training and (2) incorporate the sampling year as a covariate and build a space–time model (2D + T or 3D + T). The first strategy was adopted by Stockmann et al. (2015) to map SOC stocks at a global scale in the 1960s, 1980s, 1990s, and 2000s and to assess spatial-temporal changes in SOC. The second strategy was recently developed by Heuvelink et al. (2020) in modelling topsoil SOC change for Argentina by using a machine learning-based space–time method and by Sun et al. (2021) in modelling SOM change by using integrated nested Laplace approximation with the stochastic partial differential equation.

Another cause of the heterogeneity of legacy data can be related to different sampling protocols (e.g. between legacy soil profile databases and the grid-based soil monitoring network in France) and/or different laboratory methods (e.g. Walkley-Black method, dry combustion for SOC measurement) changing over time. One solution is to develop pedotransfer functions to make the results comparable, but this may require many soil observations where various methods have been applied (Ciampalini et al., 2013; Louis et al., 2014). The use of pedotransfer functions and differences resulting from multiple laboratories that measure soil properties based on different quality control standards is a potential error source for predictions when using soil data of different vintages from different countries for global DSM products (Libohova et al., 2019; Hu et al., 2021).

In a review, Arrouays et al. (2017b) indicated that approximately 800,000 legacy soil profiles were rescued in countries that responded to their survey and that this number of soil profiles is probably and largely underestimated. Despite the substantial success of DSM and data rescue efforts, the majority of soil data remain lost or only available in hardcopy format. However, new emerging deep learning techniques (e.g. image analysis, text recognition) are promising for converting data from hardcopy to digital format, speeding up soil data rescue. In a recent paper, Bui et al. (2020) demonstrated the necessity of considering the positional error (50 to 100 m) for data collected before 2000. We agree that this error can be a substantial source of data uncertainty in field-scale studies, and it often remains in the same order of magnitude as most of the covariates (e.g. relief, climate) used for broad-scale DSM, especially for global studies. Moreover, the data collected in the past are precious regarding their use to assess soil changes at different periods. Another advantage of using legacy data could be the use of field measured soil texture categories by soil surveyors; Malone and Searle (2021a, 2021b) presented a new approach to converting these data into quantitative estimates of soil particle size fractions by soil texture centroids and then integrating them into DSM modelling to improve map accuracy.

4.2.3. *Suggestions for new data collection*

Regarding future sampling campaigns, Brus (2019) asserted that there is no single best sampling design for DSM, and which is the best depends on the DSM technique. To produce digital soil maps, systematic grid sampling or regular geographical coverage sampling (even adding supplementary closer samples to better capture local variability) is recommended when no environmental covariate is available (Marchant and Lark, 2007; Walvoort et al., 2010; Wadoux et al., 2019a). However, these sampling schemes, particularly regular grids, may be cost-prohibitive. These two sampling designs are also preferred for constructing soil-monitoring networks to evaluate various soil properties simultaneously.

In the presence of environmental covariates, stratified random sampling, if sufficiently dense, offers an efficient manner to cover the feature and geographical spaces of some soil properties of interest and provides efficient statistical estimates (De Gruijter et al., 2006; Minasny et al., 2013). Proposed by Minasny and McBratney (2006), the conditioned Latin hypercube sampling (cLHS) is a modified version of a stratified sampling design that enables the selection of sampling locations based on the distribution values of a set of environmental covariates (feature space) and has been applied in many DSM studies (Mulder et al., 2013; Pahlavan Rad et al., 2014; Thomas et al., 2015; Omuto and Vargas, 2015). Other methods of stratification to maximise sampling efficiency (i.e. covering similar spatial variation with lowest cost) have been advocated, such as feature space coverage sampling with *k*-means (Brus et al., 2007; Yang et al., 2016; Brus, 2019; Ma et al., 2020; Wadoux et al., 2019b). In addition, covering not only the feature space but also the geographical space is necessary (Lagacherie et al., 2020). The best sampling design could result from a trade-off depending on the field conditions and relations between the target properties and covariates. If the relationship between covariates and the target property is strong, more effort can be put into sampling using the covariate feature space; if this relationship is weak, the preference may be to cover the geographical space. In any of these cases, leaving entirely "empty" areas is unsuitable because some spatial structures not captured by the covariates might be missed. Moreover, to use regression kriging, regularly spread points are necessary. For regions that have already used legacy data for DSM, the outcomes could be used to design efficient supplementary sampling campaigns in locations with greater uncertainty.

In addition to conventional laboratory analysis, significant efforts have been dedicated to cost-effective soil measurement using soil-sensing techniques (e.g. soil spectroscopy, hyperspectral remote sensing) to significantly increase the density of soil data (or pseudodata) for DSM (Viscarra Rossel et al., 2011; Lagacherie and Gomez, 2018; Lagacherie et al., 2019). Although these studies have been mainly conducted to a much smaller spatial extent, we expect that these soil-sensing techniques will be available to a wider spatial extent with the development of national and global soil spectral libraries and soil-sensing platforms and systems (Shi et al., 2015; Viscarra Rossel et al., 2016, 2017; Demattê et al., 2018; Seybold et al., 2019). Some challenges remain: i) integration of the prediction errors from sensors in DSM modelling and ii) cost-efficient data fusion between sensors having different wavelengths and precisions.

Other new datasets collected outside soil surveys and DSM applications, such as data collected by farmers or through citizen science approaches, can also be utilised and integrated into DSM studies. Román Dobarco et al. (2017) and Caubet et al. (2019) have provided examples of how to incorporate information collected on farmers' demand for merging digital maps of soil texture using ensemble modelling. Citizen science (participation of nonspecialists in scientific research) has been proposed by Rossiter et al. (2015) in DSM studies. They stated that "citizens can contribute primary observations or note discrepancies on existing maps" and that the major issues are how to stimulate citizens to participate and how to integrate observations from citizens and professionals.

Fig. 5b shows a large gap in soil information for SDs greater than 30 cm. Ascertaining soil properties at depth is essential to understanding and the complete accounting for ecosystem service provision and response. The SOC stocks in the first 30 cm represent less than half and one-third of that stored in the first 100 and 200 cm, respectively (Batjes, 1996). Pries et al. (2017) showed that warming of the whole soil (down to 100 cm) revealed a larger soil respiration response than in many in situ experiments focused on topsoil. Deep carbon may also play a major role in C cycling, because it was shown to be more recalcitrant than topsoil carbon (Balesdent et al., 2018). Considering crops and plants, some essential parameters, such as SD and AWC, cannot be correctly assessed by monitoring only the topsoil. Moreover, climate change (e.g. anomalously warm droughts) will probably influence deep soil properties, especially through changes in water behaviour in soil, impacting crops and plants (Montagne and Cornu, 2010; Schlaepfer et al., 2017; Goulden and Bales, 2019). In many cases, rooting depth will exceed 100 cm, and this is one of the main reasons why the *GlobalSoilMap* specifications include the characterisation of soil properties down to 200 cm. Deep soil properties are also linked to other important soil functions, such as water filtering, mediating pollutants, habitat support for animals, and human activities (e.g. infrastructure) (Baveye et al., 2016; Jónsson et al., 2017; Kelleway et al., 2017). However, because of limited resource availability, difficulty in deep soil sampling, and shallow soil conditions, only a few legacy soil data provide information for the 100–200 cm layer, and few current initiatives aim to develop soil databases at this depth. Data to this depth would allow predictions with acceptable precision but also support decisions that consider soil as a whole. Therefore, there is an urgent need to include data to greater SDs when building soil databases to improve soil monitoring, modelling, and functioning in further research and applications (Lal, 2018). Proximal soil sensors (e.g. electromagnetic induction, ground-penetrating radar) reveal information in subsoil, and these data are mainly used in local surveys; thus, further research is necessary to upscale them to larger areas and to search for more sensors relevant to subsoil.

4.3. Spatial modelling and prediction

4.3.1. Soil information of interest and its prediction performance

The first step of global soil mapping is to map essential soil properties related to flows, changes, and stocks of water and nutrients in soils. The 12 soil properties defined by *GlobalSoilMap* are central to nearly all modelling, monitoring, and forecasting exercises involving soils. We fully agree that mapping the threats to soil quality is also necessary (Montanarella et al., 2016), but we should not *put the cart before the horse*, and mapping these essential soil properties is a necessary precondition for soil threats and for soil functions, services, and security (McBratney et al., 2014).

In broad-scale DSM, studies on mapping SOC/SOM content and SOC stocks account for the largest proportion (77.5%). Because of the need to ensure food security and adapt to climate change, the potential of soil to store or sequester additional SOC has received considerable attention in recent years (e.g. Wiesmeier et al., 2014, 2020; McNally et al., 2017; Minasny et al., 2017; Zomer et al., 2017; Chen et al., 2018, 2019a; Lal et al., 2018; Chenu et al., 2019; Ma et al., 2021a). The UN-FAO GSP published their first global soil map on SOC stocks (GSOCmap) to establish a baseline. This map is used as a baseline to monitor soil conditions, identify degraded areas, set restoration targets, explore SOC sequestration potentials, support the reporting of greenhouse gas emission reporting, and make evidence-based decisions on adapting to and mitigating climate change. In addition, a global SOC sequestration potential map was prepared using the UN-FAO GSP. Soil scientists that recognised the significant role of SOC in the global C cycle and ecosystem services (Koch et al., 2013; Adhikari and Hartemink, 2016; Rumpel et al., 2018; Amelung et al., 2020); however, there remains a large uncertainty associated with global SOC estimates, and it consequently promotes the need to map soil carbon stocks (IPCC, 2013; Scharlemann et al., 2014). DSM can be a useful tool in the broad-scale mapping of SOC sequestration and storage potential. DSM can be an efficient decision-making platform for implementing proper, sustainable management practices and identifying areas with high potential for sequestering atmospheric carbon or for protecting soils to avoid CO_2 release into the atmosphere (Akpa et al., 2016; Chen et al., 2018, 2019b; Minasny et al., 2019; Martin et al., 2021).

Soil particle size fractions (i.e. clay, silt, and sand) are the second most frequently studied and are important for soil hydrologic model parameters (i.e. AWC and soil moisture), erosion, biogeochemical, and crop modelling.SD and thickness are among the most poorly predicted soil properties, mainly for four reasons: (1) high spatial heterogeneity, (2) difficulty in modelling with the most used Scorpan factors, (3) high cost of measurement, and (4) different definitions (Lacoste et al., 2016).

Right-censored SD (when the measured SD is less than the actual SD) is also notable because it underestimates SD for deep soil in modelling. Several solutions are available for correcting this limitation: (1) estimating the pseudo SD at the censored location by the simulated beta function from SD observations (Kempen et al., 2015); (2) integrating pseudo-observations generated by expert knowledge in modelling, which is applicable for deep soils and bare rocks (Shangguan et al., 2017); (3) using random survival forest to produce a probability map of exceeding a given SD (Chen et al., 2019c); and (4) integrating data mining with binary models representing both rock outcrops and deep soil (Malone and Searle, 2020).

Coarse fragments were almost the least predicted soil properties. This finding is probably due to its high spatial variability. Because of the importance of coarse fragments in weight-to-volume conversion for soil water, SOC, nutrients, or trace elements, improving its predictive ability is necessary. A more accurate measurement of coarse fragments would require large sample volumes. In addition to the larger volumes required for such analysis, in many routine soil surveys and laboratory measurements, coarse fragments from a soil profile are often only visually estimated. In some cases, they are measured in the field for fractions between 20 and 70 mm (in diameter) but measured in the laboratory for fractions between 2 and 20 mm (Soil Survey Staff, 2014), which may be another reason for its poor prediction accuracy. Regardless of the method of coarse fragment estimation, one challenge is to integrate the measurement error propagated in DSM predictions (Román Dobarco et al., 2019b).

Plant exploitable (effective) depth, also called rootable depth, is rarely mapped in broad-scale DSM. This phenomenon is due to two reasons: (1) rootable depth is not recorded and/or directly related to any soil property that can be observed during soil field surveys, and (2) rootable depth varies between plant species. Leenaars et al. (2018) provided a good reference for plant exploitable (effective) depth by estimating the rootability index for maize. This index expresses the adequacy (0–100%) of the 10 selected soil factors to support root growth relative to optimal root growth. A threshold index of 20% for each soil factor describes inadequate conditions for plant exploitable (effective) depth.

Currently, DSM has a very unbalanced vision of the soil that mainly focuses on SOC and nearly ignores key properties such as SD or coarse fragments. In the future, DSM should address these properties, soil types, and soil quality indicators that match user demands. Many national and global concerns (e.g. food security, water security, human health, mitigation, adaptation to climate change, biodiversity protection) require other soil information. This is also relevant for mapping threats to soils (e.g. erosion, salinity, loss in SOC, contamination, nutrient imbalance, loss of biodiversity, compaction). Thus, applying broad-scale mapping to other soil properties, such as electrical conductivity, soil moisture, Na^{2+}, N, P, K, S, trace elements and persistent organic pollutants, infiltration capacity, structural stability, and biological properties (e.g. the abundance, diversity, and activity of soil organisms) is necessary. Some soil structure descriptions and related properties are often stored in databases as qualitative variables, but they are rarely used for DSM. In addition, Richer-de Forges et al. (2019) stated that based on a survey of users' needs in France, some indicators of soil properties, such as structural stability, macro-porosity, water infiltration, and rooting potential, are often asked by end-users as they integrate information on many soil functional properties. Maps of functional soil properties rather than maps of directly analytical soil variables are necessary, which emphasises the need to move from DSM to digital soil assessment (DSA) (e.g. Carré et al., 2007; Finke, 2012; Minasny et al., 2012; Kidd et al., 2015, 2020; Arrouays et al., 2020b). Finally, there remains large potential for coupling process knowledge, pedology, and DSM (Finke, 2012; Ma et al., 2019a, 2019b; Wadoux, 2019) and for mapping services rendered by soils (e.g. Dominati et al., 2010; Robinson et al., 2014; Turner et al., 2016; Kidd et al., 2020). In summary, future DSM should involve stakeholder requirements to produce products that decision makers (not only soil scientists) need and want to use.

Overall, model performance (in R^2) for most of the soil properties was not high in broad-scale DSM, which may interactively result from the following reasons: (1) the low quality of input soil data, including the unrepresentativeness of soil samples and differences in soil sampling dates; (2) the low quality of environmental covariates, including the inconsistency of resolution in the original data, the absence of data relevant to soil formation, and target properties; (3) the mismatch between point-based soil sampling and pixel-based modelling; and (4) the low predictive ability of the model. The solutions to these issues are discussed in the following sections.

4.3.2. Space–time modelling

Most DSM studies focus on predicting soil properties for a particular time frame but not on their changes. To map past changes, Meersmans et al. (2011) resampled soil profiles at the same locations as a soil survey from the 1960s. Most other studies have used soil data from several periods obtained using different sampling designs (Sun et al., 2012; Minasny et al., 2016; Schillaci et al., 2017; Song et al., 2018; Huang et al., 2019; Zhou et al., 2019; Sun et al., 2021). However, the range of the estimated trend should be compared with the map uncertainty to test if this trend is plausible or at the same order of magnitude as the cumulated errors linked to the spatial predictions at different dates.

Few studies have attempted to predict future soil changes, and most of them are related to SOC. Minasny et al. (2013) stated that two methods are used to address this issue: a dynamic mechanistic simulation model and a static empirical model. In a dynamic mechanistic simulation model, DSM is first used to estimate an initial soil state; next, the model is simulated per pixel under future climate, land use/land cover (LULC), and land management scenarios (Martin et al., 2021). In a static empirical model, future soil changes can be predicted using a fitted Scorpan model, in which the present climate, LULC, and land management are replaced by future scenarios (Yigini and Panagos, 2016; Gray and Bishop, 2016; Meersmans et al., 2016; Adhikari et al., 2019; Reyes Rojas et al., 2018). These last broad-scale DSM studies used a static empirical model, which mainly results from several constraints, such as (1) the large disconnection between DSM and mechanistic dynamics modelling, and (2) complex parameter initialisation and heavy computing, which is challenging for mechanistic dynamics modelling on a broad scale (Walter et al., 2006; Luo et al., 2016). These challenges require collaboration among scientists from multiple disciplines and improved integration of DSM and dynamic mechanistic modelling to speed up the simulation efficiency and improve the prediction accuracy (e.g. simulate observed locations by mechanistic dynamics model and then map soil information by DSM on simulated data).

4.3.3. Environmental covariates

The frequency of the Scorpan factors used in DSM is often restricted by the availability of environmental covariates (Grunwald, 2009). Benefiting from global, free, and available remote sensing data (Fig. 11), relief, organisms (mainly vegetation), and climate factors have been widely used in broad-scale DSM, and the frequency of other factors, such as soil, parent material, age, and position, have had more limited use (Fig. 6). Some proximal sensors, such as electromagnetic induction, ground-penetrating radar, and X-rays, are commonly used in field-scale soil mapping and are absent in large-scale mapping studies.

LULC, vegetation index (e.g. normalised difference vegetation index, enhanced vegetation index), and net primary productivity are mainly derived from remote sensing products. Because of substantial advances, satellites can now provide products with a high spatiotemporal resolution. Examples include the newly launched multitemporal Sentinel 1 (5–20 m resolution), Sentinel-2 (10–60 m resolution), and Sentinel 3 (300 m resolution), which has proven its high potential in DSM (Poggio and Gimona, 2017; Loiseau et al., 2019; Dharumarajan et al., 2020; Zhou et al., 2020; Zhou et al., 2021). The potential of future Sentinel-10 hyperspectral data requires further exploration.

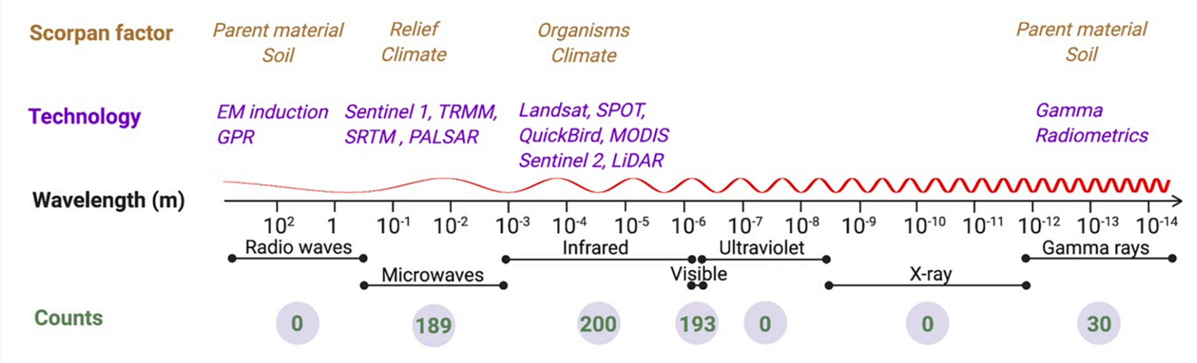

Fig. 11. Sensing techniques and their relevance to Scorpan factors. Counts indicate the frequency of different bands used as environmental covariates.

Notably, hyperspectral remote sensing can also be used as a proxy for some soil properties in topsoil for bare soil surfaces (Gholizadeh et al., 2018; Lagacherie and Gomez, 2018; Castaldi et al., 2019; Vaudour et al., 2019). Fine-resolution and multitemporal imagery offer the possibility of detecting vegetation dynamics and changes in some topsoil properties (e.g. SOC, salinity). Using a time series average (from month to decade) or a particular snapshot is a commonly used practice in DSM modelling. However, spectral images may only represent "recent" times (at least since the widespread deployment of satellites) where soil has been more disturbed by human beings. As a result, poor correspondence between covariate snapshots (the time coverage of spectra images) and the status of the soil property is expected. This limitation from satellite imagery can probably be solved by integrating other Scorpan factors, such as parent material and age, in spatial predictive modelling. Including recent LULC changes (e.g. conversion from forest or grassland to cropland) and land management practices helps improve map accuracy, especially in regions with intensive human activities. In addition, new effective covariates can also be explored in DSM, such as land surface dynamic feedback information and land surface phenology variables (e. g. Zeng et al., 2020; Yang et al., 2021).

Evidence shows that soil micro- and macro-organisms contribute significantly to global biogeochemical cycles in a changing climate (Wieder et al., 2015; Luo et al., 2016; Cavicchioli et al., 2019; Jansson and Hofmockel, 2020). These below-ground organisms are rarely used in DSM because they are much more difficult to measure than aboveground organisms, in addition to being highly variable both in space and time (Banfield et al., 2017). Global maps for soil fungi (Tedersoo et al., 2014), bacteria (Delgado-Baquerizo et al., 2018), nematodes (van den Hoogen et al., 2019), and earthworms (Phillips et al., 2019) have been published, providing good resources for broad-scale DSM practices. Global soil biodiversity maps were produced using a DSM approach. However, concerns remain about their correlation with other environmental covariates (e.g. LULC, climate). In addition, such biological indicators are highly variable in space and time (Kuzyakov and Blagodatskaya, 2015), making accurate predictions difficult.

In most studies, only the annual averages of climatic data have been considered in the DSM. We suggest that the short-term average (monthly or seasonal) of climatic data and its variance within the corresponding period may also be a useful covariate to characterise some soil properties (e.g. Keskin et al., 2019; Liu et al., 2020b) because this information provides insights into key soil processes (e.g. wetting and drying frequency, soil moisture movement). Changes in vegetation response to extreme climatic events, such as changes in NDVI or other remote sensing indices after extreme droughts, could also be very useful (Mulder et al., 2019).

In broad-scale DSM studies, soil factors are often characterised by soil class maps and/or soil texture maps derived from historical soil surveys (Grunwald, 2009). Soil information from proximal soil sensing can also be spatially interpolated and serve as a covariate for DSM. For example, the first three principle components of visible–near infrared (Vis-NIR) spectra have been used to produce an Australian three-dimensional soil grid (Viscarra Rossel et al., 2015). Other soil information, including soil moisture and soil property maps from other sources, can also be used as covariates in spatial predictive modelling (Keskin et al., 2019; Liang et al., 2019).

For broad-scale studies, parent material is mainly derived from geological maps and partially from airborne gamma-rays in some countries or regions, such as the United States, Australia, the United Kingdom, and some regions of France (Lacoste et al., 2011; Beamish, 2014; Viscarra Rossel et al., 2015; Keskin et al., 2019). Based on the statistics in Fig. 6, the usage of the parent material is low because of the absence of data sources. With technical advances, such as airborne gamma-ray spectrometry and proximal gamma-ray spectrometers, the availability of this factor is expected to increase and is already used in DSM practice (Loiseau et al., 2020; Chen et al., 2021a). Gray et al. (2016) demonstrated strong improvement in DSM predictions when the lithology variable was classified for pedologic purposes and then used for DSM. They pointed out that lithology data could have substantial potential for use in DSM. Despite some recent progress (Miller and Juilleret, 2020; Simon et al., 2021), lithology maps remain scarce in most parts of the world; thus, they require further exploration.

Despite its significant role in pedogenesis, age remains the least used Scorpan factor because of the difficulty of direct measurement at a broad scale (McBratney et al., 2003; Zhang et al., 2017). Considerable advances in technology (e.g. soil dating, material dating) and expert knowledge are necessary to derive the age factor, especially for broad-scale DSM. Geology and geomorphology, however, may help derive information on age (e.g. ancient periglacial landscapes).

Because the number of environmental covariates used for DSM has rapidly increased recently, the principle of parsimony has become more crucial than ever for covariate selection in DSM (Wadoux et al., 2020). Although the use of more environmental covariates may improve the model accuracy, especially in combination with machine learning, this approach may introduce more uncertainty from input data and make the results less interpretable from a soil science point of view. Therefore, to find the correct balance between parsimony and model performance, covariate selection is necessary in DSM, not only based on a pure statistical selection or case-based reasoning but also on their pedological relevance (Wadoux et al., 2020; Liang et al., 2021).

Notably, the importance of environmental covariates varies in different soil properties of interest (Fig. 12). SOC and SOC stocks are jointly dominated by temperature, precipitation, elevation, and parent materials in broad-scale studies, and BD is mainly controlled by soil class, temperature, precipitation, land use, topographic wetness index, and parent materials. If the soil properties of interest were soil particle size fractions, then parent materials, elevation, solar radiation, temperature, and precipitation are major controlling factors, and the use of gamma-ray data can also help improve their model performance.

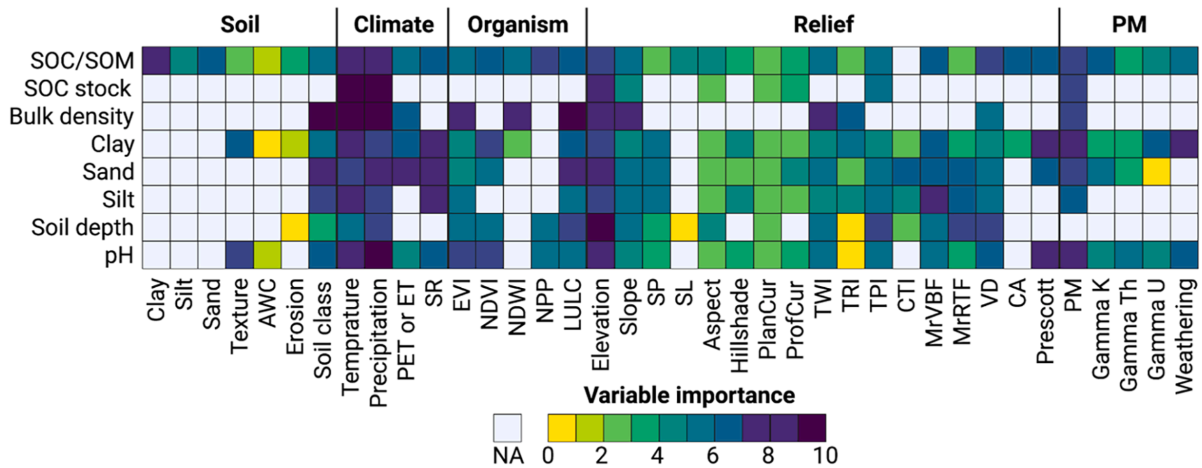

Fig. 12. Variable importance in broad-scale DSM. Considering the differences in determining the variable importance, such as correlation coefficient, t value, F value or partial R^2 using MLR, rule usage in Cubist, %IncMSE, or IncNodePurity in machine learning, the importance of each variable is ranked from 10 (most important) to 1 (least important). We excluded AWC, CEC, and coarse fragments because the variable importance for these soil properties were reported in less than five articles (not robust enough). Notably, we only keep widely used Scorpan-related environmental covariates. AWC, available water capacity; PET, potential evapotranspiration; ET, evapotranspiration; SR, solar radiation; EVI, enhanced vegetation index; NDVI; normalised difference vegetation index; NDWI, normalised difference water index; NPP, net primary production; LULC, land use and land cover; SP, slope position; SL, slope length; PlanCur, plan curvature; ProCur, profile curvature; TWI, topographic wetness index; TRI, terrain ruggedness index; TPI, topographic position index; CTI, compound topographic index; MrVBF, multi-resolution valley bottom flatness; MrRTF, multi-resolution ridge top flatness; VD, valley depth; CA, contribution area; PM, parent materials.

Precipitation and relief derivatives (i.e. elevation, topographic position index, multi-resolution ridge top flatness, and valley depth) are important for soil depth modelling, and soil pH is driven by precipitation, parent materials, the Prescott index, elevation, and soil texture. The variable importance shown in Fig. 12 can be used as a reference for selecting useful covariates in broad-scale DSM studies, and notably, the variable importance can be highly site- and depth-specific. Multiscale interactions between environmental covariates and the soil property of interest may be considered. O'Rourke et al. (2015) and Wiesmeier et al. (2019) have indicated that the drivers of SOC varied considerably from micro to global scales, and their findings are support those of Lamichhane et al. (2019), who reviewed 120 studies on DSM of SOC across scales, and the results of Adhikari et al. (2020) after investigating the environmental controllers of SOC with scale. To manage multiscale interactions between soil properties and environmental covariates, wavelet transformation, empirical mode decomposition, and the Gaussian scale space are suggested to produce multiscale covariate layers to potentially improve the prediction accuracy in DSM (e.g. Zhou et al., 2016; Behrens et al., 2018; Sun et al., 2019; Taghizadeh-Mehrjardi et al., 2021).

A commonly observed issue in DSM using field point data and covariates is that they are not linked to the same spatial support (points *vs* "pixels" or map units) and do not have the same accuracy. Thus, one challenge is how to make the best use of both data when knowing that there are discrepancies in spatial support and precision of measurements.

4.3.4. *Predictive models*

Machine learning has become the most commonly used predictive model in broad-scale DSM since 2011 (Fig. 8a). This trend has also been confirmed by Arrouays et al. (2020b) and Padarian et al. (2020), and mainly results from three reasons: (1) machine learning can manage complex nonlinear relationships between the soil property of interest and an increasing number of environmental covariates, and it thus often performs better than classic statistical models and geostatistics; (2) the rapidly increasing computing power and techniques (e.g. parallel computing, cloud computing, high-performance computing) make it more efficient to produce digital soil maps from big data using DSM than ever before; and (3) machine learning is a nonparametric method that does not require any hypothesis on distribution and stationary, which are no longer valid with large spatial extents and legacy data. However, geostatistical models remain useful because they can capture other spatial structures (e.g. diffuse contamination) better than pure machine learning models, and the use of hybrid modelling (e.g. random forest spatial interpolation, Sekulić et al., 2020) for broad-scale DSM may improve the model performance when some spatial structures have not been captured by machine learning models. An alternative to using geostatistical models, which may be difficult to undertake at a broad scale, is to consider the geographical locations as inputs of machine learning algorithms, enabling these algorithms to capture the spatial structures not explained by classical environmental covariates (Hengl et al., 2018). Nevertheless, using spatial coordinates as inputs in hierarchical models should be avoided because these models partition the input space and do not fit a trend.

A recent advance in predictive models of DSM is the introduction of deep learning (i.e. 2-dimensional convolutional neural network), explicitly described by Padarian et al. (2019), Wadoux (2019), Wadoux et al. (2019c), and Taghizadeh-Mehrjardi et al. (2020a). Deep learning opens new possibilities for predicting soil properties because (1) the input data for model training is a stack of spatial patterns, not spatial points, and (2) the trained model enables simultaneous predictions of multiple soil properties (Padarian et al., 2019).

Despite substantial advances in machine learning and deep learning, predictive models focus on prediction performance and overlook the importance of pedological knowledge for DSM and the use of DSM in understanding controlling factors of the soil property of interest. Therefore, further DSM research should further endeavour to open the "black boxes" of machine learning and deep learning (e.g. convolutional neural network) and integrate more pedological knowledge (e.g. structural equation modelling) in both the predictive model and environmental covariate selection (Angelini et al., 2017; Angelini and Heuvelink, 2018; Arrouays et al., 2020a; Ma et al., 2019b). Conversely, DSM can be used not only for predictive mapping but also for enhancing pedological knowledge (Wadoux et al., 2020; Ma et al., 2019a).

4.3.5. *Comparison between 2.5D and 3D models*

There is no general consensus on whether the 2.5D model is better than the 3D model or vice versa (Ma et al., 2021b). Each has pros and cons. In a 2.5D model, an equal-area spline is commonly used in vertical interpolation or depth interval standardisation (Ma et al., 2021b). This spline technique can be sensitive to outliers and cannot be extrapolated outside the range of observations.

Another approach is to fit a depth function (e.g. exponential function) in 3D modelling (Meersmans et al., 2009a, 2009b; Ottoy et al., 2017; Rentschler et al., 2019). Poggio and Gimona (2014) used a 3D GAM model using geographical location (x, y) and depth (z) as coordinates to fit the model. Hengl et al. (2017) directly used depth as a covariate in random forest and gradient bootstrap tree models.

Liu et al. (2016) proposed an approach to predict the 3D variation of SOM concentration by integrating a similarity-based method with depth functions. They concluded that the proposed approach is effective and accurate for 3D SOM prediction and that it overcomes the two drawbacks of the 2.5D approach: (1) the neglect of vertical soil patterns when performing horizontal soil predictions and (2) the repeated applications of depth function fittings in the mapping process, both of which may lead to prediction errors. However, Nauman and Duniway (2019) noted that 3D modelling of soil properties with strong variation with depth can result in substantial areas with much higher uncertainty that coincide with unrealistic predictions relative to 2.5D models, although 3D models performed slightly better. Roudier et al. (2020) proposed a 3D modelling approach relying on data augmentation and demonstrated that there were no differences in 2.5D and 3D models in predicting soil properties at defined depth intervals. Ma et al. (2021b) found similar results; however, using depth as a predictor in tree-based models could create "stepped" depth functions. The choice between 2.5D and 3D models may be case-specific, depending on the variation of soil properties with depth and soil sampling volume.

4.4. Performance evaluation uncertainty estimation and map resolution

4.4.1. Validation strategy

As aforementioned in the results, data splitting either in single random holdback (45.5%) or cross-validation (37.7%) was the most commonly adopted validation strategy for evaluating DSM products, which is in line with the recent finding of Piikki et al. (2021). Although more studies chose a single random holdback, it may result in nonrobust accuracy because the single randomly selected validation data may not represent the entire dataset (Lagacherie et al., 2019; Chen et al., 2021b). Therefore, we recommend repeated random holdback (i.e. 100 times) or cross-validation for reporting robust validation results. Indeed, such an approach has provided uncertainty estimates in many broad-scale DSM studies (e.g. Mulder et al., 2016; Kempen et al., 2019; Loiseau et al., 2019). In addition, spatial cross-validation needs to be further tested for clustered legacy data to avoid overly optimistic model performance due to spatial autocorrelation (Meyer et al., 2018; Ploton et al., 2020; Poggio et al., 2021).

Brus et al. (2011) recommended the use of independent validation datasets because data splitting may not provide an unbiased accuracy assessment because of the nonrandom sampled soil data. These additional independent data can be collected using a design-based sampling strategy involving probability sampling and design-based estimation. Due to the high cost of additional soil sampling, only a few broad-scale studies (9.1%) have used independent validation for map evaluation (Thomas et al., 2015; Rial et al., 2016; Vaysse et al., 2017; Ellili Bargaoui et al., 2019). Lagacherie et al. (2019) suggested that if this independent validation is not conducted with a proper sampling density, it can lead to uncertain prediction performance assessments.

4.4.2. Indicators for model evaluation

As indicated in Fig. 8b, R2 is the most commonly used indicator for model evaluation of continuous soil properties. The use of R^2 allows the comparison of the accuracy for different soil properties with various units and magnitudes; thus, a recommendation is to report it in DSM studies. However, R^2 has several limitations in interpretation because it strongly depends on the number of points used to calculate it, and it is very sensitive to the presence of extreme values or outliers. Notably, the method used to calculate R^2 differs by the authors; some may take the R^2 by the square of the correlation coefficient between observations and predictions, and others use the R^2 by using the percentage of explained variance of the target property. If different DSM studies are to be compared, a consistent method for R^2 calculation must be adopted. RPD is another indicator that eliminates the difference in units and magnitudes. However, Minasny and McBratney (2013) suggested using either RPD or R^2 because they are the same measure, and the ratio of performance to interquartile range (RPIQ) is a better indicator than RPD for data that are not normally distributed (Bellon-Maurel et al., 2010).

In addition to the aforementioned indicators, we suggest the following in DSM studies: (1) SSVR and PICP. SSVR has been used to fill the gap if none of the aforementioned indicators are relevant to the spatial structure of the produced digital maps (Poggio and Gimona, 2018). It is defined as a complement to one of the nugget to sill ratios (Kerry and Oliver, 2008). An SSVR value closer to one indicates a higher proportion of the data explained by the spatial component. PICP is an indicator that determines the efficacy of uncertainty estimates, which could be important for end-users in decision-making. If the uncertainty estimates are reasonably defined, PICP should result in 90% for a 90% prediction interval (Malone et al., 2016).

Caution should be used when employing large-scale (i.e. national, continental, and global) DSM products for studies at a local scale, because map accuracy at this scale may be much lower than global accuracy (Gomez and Coulouma, 2018). Notably, uncertainty may be underestimated because of sampling imperfections. For instance, Lagacherie et al. (2020) showed underestimations of the uncertainty of DSM predictions, especially for sparse samplings poorly covering the distribution of the target soil property and the dense samplings unevenly distributed in the geographical space.

4.4.3. Estimates of map uncertainty

More than half (137 of 244 articles) studies provided estimates of map uncertainty, of which a large percentage were published after 2015 and related to *GlobalSoilMap* products. Approaches used for uncertainty quantification can be classified into the five groups (Malone et al., 2017): (1) universal kriging prediction variance, (2) bootstrapping, (3) empirical uncertainty quantification through data partitioning and cross-validation, (4) Monte Carlo simulation, and (5) the Bayesian approach. Groups 1, 4, and 5 were mainly used in geostatistical models, and Groups 2 and 3 were mainly used for machine learning models. The produced map and its associated uncertainty are useful in decision-making for end-users and allow quantification of uncertainty propagation in some secondary soil information (e.g. SOC stocks, AWC) and digital soil assessment (Finke, 2012; Poggio and Gimona, 2014; Román Dobarco et al., 2019a, 2019b). For example, Román Dobarco et al. (2019b) found that the main sources of uncertainty for soil AWC maps were not the pedotransfer function for predicting AWC but the input maps of coarse fragments and particle size fractions.

Among the reviewed papers, none provided a complete probability distribution for each property at six depth intervals. In addition, no study provided the joint probability distribution of several soil properties, which are necessary to combine these soil properties to develop soil indicators for digital soil assessment. Thus, further efforts are necessary to fulfil the *GlobalSoilMap* specifications (marginal probability distribution and joint probability distribution), as Heuvelink (2014) stressed.

In *GlobalSoilMap* specifications (Tier 1), the uncertainty of the estimates of soil properties should be presented at 90% prediction intervals (PIs) (Arrouays et al., 2014a). An important question is whether 90% PIs are useful for decision-making or should be narrower, especially for poorly predicted soil properties. In some cases, as Helmick et al. (2014) demonstrated, the actual width of these intervals might be larger than is practically useful. For some decision makers, a PI of 75% might be adequate to support a decision. If we compared the PI with conventional map purity, which rarely exceeds 80% (Marsman and de Gruijter, 1986), even for detailed maps, we might wonder if 90% PI is practical and useful for some end-users. Models such as quantile regression forest (Meinshausen, 2006; Vaysse and Lagacherie, 2017; Kasraei et al., 2021) allow the use of different quantiles to estimate PIs. Moreover, communication on uncertainty would be more efficient if it was based not only on the primary soil attributes but also on the consequences of their uncertainties when they are used as inputs for final products delivered by end-users (Heuvelink, 1998; Richer-de-Forges et al., 2019) or modellers. For the modeller, knowing only a modal or a mean value and a PI seems useless unless hypotheses on the distribution of the target soil property can be formulated. In some cases, simple hypotheses such as lognormal or triangular distributions may be used (Odgers et al., 2014b). Heuvelink (2014) proposed the representation of uncertainty in soil property maps by using probability distribution functions.

Although the maps we reviewed were mainly shown as “pixel” or “voxel” representations, they are point-based predictions. In moving to the Tier 2 *GlobalSoilMap* product, predictions on an area (or a block) rather than a point must be made. This should have consequences on the way the uncertainties of these predictions are modelled and estimated. This task is not trivial, because most soil data used for training and validation are, at present, point observations. Moreover, in many applications, end-users are interested in obtaining information on areal units (e.g. watershed, municipality, farm, field); thus, spatially aggregating soil property prediction and uncertainty is often necessary (Vaysse et al., 2017).

5. Conclusions

In this paper, we reviewed 244 articles on the use of DSM to map *GlobalSoilMap* soil properties at a broad scale (>10,000 km^2) published from January 2003 to July 2021. This review provides the following insights.

(1) The number of DSM publications increased exponentially; however, most of the growth occurred on the application, not in the development of new DSM approaches.

(2) Many articles do not provide information on how the soil sample data were collected, such as sampling year and sampling strategy. We suggest that this information should be reported because of its high relevance to DSM quality and the design of future sampling campaigns.

(3) Most of the studies focused on mapping SOC/SOM, SOC stocks, and soil particle size fractions. Further efforts are necessary to predict other soil properties essential for soil function.

(4) Half the articles produced maps for topsoil only (<30 cm). Studies on deep soil, down to 200 cm, need to be more represented (21.7%). We also observed a decreasing model performance trend with deeper depth intervals for many soil properties. New sampling efforts should focus on the whole soil profile and finding improved covariates and new sensors for deep soil.

(5) In many cases, missing covariates constrain the prediction accuracy of the models. For example, the age factor has rarely been used, because of the lack of relevant covariates; therefore, advances in technology are necessary to better represent age.

(6) Nonlinear models (i.e. machine learning) have been increasingly used in DSM. Incorporating more pedological knowledge for reasonable modelling is necessary from scientific and technical viewpoints. Additionally, a focus should be figuring out how DSM could provide new insights into pedological processes.

(7) In addition to commonly used model performance indicators (i.e. R^2, RMSE, and ME), the SSVR and PICP are suggested for model performance evaluation so that the spatial structure and the efficiency of the uncertainty estimates can be accounted for.

(8) We call for a continual effort in legacy data rescue and new data collection for mapping historic and recent soil status, forming the basis for monitoring and/or forecasting soil dynamics to support evidence-based decision-making on soil resources. Furthermore, instead of using either a dynamic mechanistic simulation model (e.g. CENTURY, RothC) or a static empirical model (DSM) for mapping soil changes, we suggest integrating both to provide accurate estimates.

(9) Insights gained in this review indicate that the DSM community is working progressively to provide fine-resolution digital soil maps to address the global challenges related to soil resources. However, challenges remain, especially in integrating DSM efforts with other disciplines and developing functional soil maps and metrics to incorporate them in decision-making processes at multiple levels.

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.

Acknowledgements

This work is supported by LE STUDIUM Loire Valley Institute for Advanced Studies (France). Songchao Chen received support from the China Scholarship Council (grant no. 201606320211). This work benefited from collaborations established in the framework of the EU EJP Soil “Towards climate-smart sustainable management of agricultural soils” (grant no. 862695).

References

Adhikari, K., Hartemink, A.E., 2016. Linking soils to ecosystem services—A global review. Geoderma 262, 101–111.
Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B., Greve, M.H., 2013. High-resolution 3-D mapping of soil texture in Denmark. Soil Science Society of America Journal 77 (3), 860–876.
Adhikari, K., Mishra, U., Owens, P.R., Libohova, Z., Wills, S.A., Riley, W.J., Hoffman, F. M., Smith, D.R., 2020. Importance and strength of environmental controllers of soil organic carbon changes with scale. Geoderma 375, 114472. https://doi.org/10.1016/j.geoderma.2020.114472.
Adhikari, K., Owens, P.R., Libohova, Z., Miller, D.M., Wills, S.A., Nemecek, J., 2019. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. Science of the Total Environment 667, 833–845.
Akpa, S.I.C., Odeh, I.O.A., Bishop, T.F.A., Hartemink, A.E., 2014. Digital mapping of soil particle-size fractions for Nigeria. Soil Science Society of America Journal 78 (6), 1953–1966.
Akpa, S.I.C., Odeh, I.O.A., Bishop, T.F.A., Hartemink, A.E., Amapu, I.Y., 2016. Total soil organic carbon and carbon sequestration potential in Nigeria. Geoderma 271, 202–215.
Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J.W., Mooney, S., van Wesemael, B., Wander, M., Chabbi, A., 2020. Towards a global-scale soil climate mitigation strategy. Nature communications 11 (1). https://doi.org/10.1038/s41467-020-18887-7.
An, Y., Yang, L., Zhu, A.-X., Qin, C., Shi, JingJing, 2018. Identification of representative samples from existing samples for digital soil mapping. Geoderma 311, 109–119.
Angelini, M.E., Heuvelink, G.B.M., 2018. Including spatial correlation in structural equation modelling of soil properties. Spatial Statistics 25, 35–51.
Angelini, M.E., Heuvelink, G.B.M., Kempen, B., 2017. Multivariate mapping of soil with structural equation modelling. European Journal of Soil Science 68 (5), 575–591.
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., 2014a. Chapter three — Globalsoilmap: Toward a fine-resolution global grid of soil properties. Advances in Agronomy 125, 93–134.
Arrouays, Dominique, Lagacherie, Philippe, Hartemink, Alfred E., 2017a. Digital soil mapping across the globe. Geoderma Regional 9, 1–4.
Arrouays, Dominique, Leenaars, Johan G.B., Richer-de-Forges, Anne C., Adhikari, Kabindra, Ballabio, Cristiano, Greve, Mogens, Grundy, Mike, Guerrero, Eliseo, Hempel, Jon, Hengl, Tomislav, Heuvelink, Gerard, Batjes, Niels, Carvalho, Eloi, Hartemink, Alfred, Hewitt, Alan, Hong, Suk-Young, Krasilnikov, Pavel, Lagacherie, Philippe, Lelyk, Glen, Libohova, Zamir, Lilly, Allan, McBratney, Alex, McKenzie, Neil, Vasquez, Gustavo M., Mulder, Vera Leatitia, Minasny, Budiman, Montanarella, Luca, Odeh, Inakwu, Padarian, Jose, Poggio, Laura, Roudier, Pierre, Saby, Nicolas, Savin, Igor, Searle, Ross, Solbovoy, Vladimir, Thompson, James, Smith, Scott, Sulaeman, Yiyi, Vintila, Ruxandra, Rossel, Raphael Viscarra, Wilson, Peter, Zhang, Gan-Lin, Swerts, Martine, Oorts, Katrien, Karklins, Aldis, Feng, Liu, Ibelles Navarro, Alexandro R., Levin, Arkadiy, Laktionova, Tetiana, Dell'Acqua, Martin, Suvannang, Nopmanee, Ruam, Waew, Prasad, Jagdish, Patil, Nitin, Husnjak, Stjepan, Pásztor, László, Okx, Joop, Hallett, Stephen, Keay, Caroline, Farewell, Timothy, Lilja, Harri, Juilleret, Jérôme, Marx, Simone, Takata, Yusuke,

Kazuyuki, Yagi, Mansuy, Nicolas, Panagos, Panos, Van Liedekerke, Mark, Skalsky, Rastislav, Sobocka, Jaroslava, Kobza, Josef, Eftekhari, Kamran, Alavipanah, Seyed Kacem, Moussadek, Rachid, Badraoui, Mohamed, Da Silva, Mayesse, Paterson, Garry, Gonçalves, Maria da Conceição, Theocharopoulos, Sid, Yemefack, Martin, Tedou, Silatsa, Vrscaj, Borut, Grob, Urs, Kozák, Josef, Boruvka, Lubos, Dobos, Endre, Taboada, Miguel, Moretti, Lucas, Rodriguez, Dario, 2017b. Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. GeoResJ 14, 1–19.

Arrouays, Dominique, McBratney, Alex, Bouma, Johan, Libohova, Zamir, Richer-de-Forges, Anne C., Morgan, Cristine L.S., Roudier, Pierre, Poggio, Laura, Mulder, Vera Leatitia, 2020a. Impressions of digital soil maps: The good, the not so good, and making them ever better. Geoderma Regional 20, e00255. https://doi.org/10.1016/j.geodrs.2020.e00255.

Arrouays, D., McKenzie, N., Hempel, J., Richer-de Forges, A.C., McBratney, A.B., 2014b. GlobalSoilMap: Basis of the Global Spatial Soil Information System. CRC Press.

Arrouays, Dominique, Poggio, Laura, Salazar Guerrero, Osvaldo A., Mulder, Vera Laetitia, 2020b. Digital soil mapping and GlobalSoilMap. Main advances and ways forward. Geoderma Regional 21, e00265. https://doi.org/10.1016/j.geodrs.2020.e00265.

Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Derrien, D., Fekiacova, Z., Hatté, C., 2018. Atmosphere–soil carbon transfer as a function of soil depth. Nature 559 (7715), 599–602.

Banfield, Callum C., Dippold, Michaela A., Pausch, Johanna, Hoang, Duyen T.T., Kuzyakov, Yakov, 2017. Biopore history determines the microbial community composition in subsoil hotspots. Biology and Fertility of Soils 53 (5), 573–588.

Ellili Bargaoui, Yosra, Walter, Christian, Michot, Didier, Saby, Nicolas P.A., Vincent, Sébastien, Lemercier, Blandine, 2019. Validation of digital maps derived from spatial disaggregation of legacy soil maps. Geoderma 356, 113907. https://doi.org/10.1016/j.geoderma.2019.113907.

BATJES, N.H., 1996. Total carbon and nitrogen in the soils of the world. European Journal of Soil Science 47 (2), 151–163.

Baveye, P.C., Baveye, J., Gowdy, J., 2016. Soil "ecosystem" services and natural capital: Critical appraisal of research on uncertain ground. Frontiers in Environmental Science 4, 41.

Beamish, D., 2014. Peat mapping associations of airborne radiometric survey data. Remote Sensing 6 (1), 521–539.

Behrens, T., Schmidt, K., MacMillan, R.A., Viscarra Rossel, R.A., 2018. Multi-scale digital soil mapping with deep learning. Scientific Reports 8 (1), 1–9.

Bellon-Maurel, Véronique, Fernandez-Ahumada, Elvira, Palagos, Bernard, Roger, Jean-Michel, McBratney, Alex, 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. Trends in Analytical Chemistry 29 (9), 1073–1081.

Bouma, Johan, 2014. Soil science contributions towards sustainable development goals and their implementation: Linking soil functions with ecosystem services. Journal of Plant Nutrition and Soil Science 177 (2), 111–120.

Brus, D.J., 2019. Sampling for digital soil mapping: A tutorial supported by R scripts. Geoderma 338, 464–480.

Brus, D.J., De Gruijter, J.J., Van Groenigen, J.W., 2007. Designing spatial coverage samples using the k-means clustering algorithm. Developments in Soil Science 31, 183–192.

Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. European Journal of Soil Science 62 (3), 394–407.

Bui, Elisabeth N., Searle, Ross D., Wilson, Peter R., Philip, Seonaid R., Thomas, Mark, Brough, Dan, Harms, Ben, Hill, Jason V., Holmes, Karen, Smolinski, Henry J., Van Gool, Dennis, 2020. Soil surveyor knowledge in digital soil mapping and assessment in Australia. Geoderma Regional 22, e00299. https://doi.org/10.1016/j.geodrs.2020.e00299.

Carré, F., McBratney, Alex B., Minasny, B., 2007. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. Geoderma 141 (1-2), 1–14.

Castaldi, Fabio, Hueni, Andreas, Chabrillat, Sabine, Ward, Kathrin, Buttafuoco, Gabriele, Bomans, Bart, Vreys, Kristin, Brell, Maximilian, van Wesemael, Bas, 2019. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands. ISPRS Journal of Photogrammetry and Remote Sensing 147, 267–282.

Caubet, Manon, Román Dobarco, Mercedes, Arrouays, Dominique, Minasny, Budiman, Saby, Nicolas P.A., 2019. Merging country, continental and global predictions of soil texture: Lessons from ensemble modelling in France. Geoderma 337, 99–110.

Cavicchioli, Ricardo, Ripple, William J., Timmis, Kenneth N., Azam, Farooq, Bakken, Lars R., Baylis, Matthew, Behrenfeld, Michael J., Boetius, Antje, Boyd, Philip W., Classen, Aimée T., Crowther, Thomas W., Danovaro, Roberto, Foreman, Christine M., Huisman, Jef, Hutchins, David A., Jansson, Janet K., Karl, David M., Koskella, Britt, Mark Welch, David B., Martiny, Jennifer B.H., Moran, Mary Ann, Orphan, Victoria J., Reay, David S., Remais, Justin V., Rich, Virginia I., Singh, Brajesh K., Stein, Lisa Y., Stewart, Frank J., Sullivan, Matthew B., van Oppen, Madeleine J.H., Weaver, Scott C., Webb, Eric A., Webster, Nicole S., 2019. Scientists' warning to humanity: Microorganisms and climate change. Nature Reviews Microbiology 17 (9), 569–586.

Chen, Songchao, Arrouays, Dominique, Angers, Denis A., Chenu, Claire, Barré, Pierre, Martin, Manuel P., Saby, Nicolas P.A., Walter, Christian, 2019a. National estimation of soil organic carbon storage potential for arable soils: A data-driven approach coupled with carbon-landscape zones. Science of The Total Environment 666, 355–367.

Chen, Songchao, Arrouays, Dominique, Angers, Denis A., Martin, Manuel P., Walter, Christian, 2019b. Soil carbon stocks under different land uses and the applicability of the soil carbon saturation concept. Soil and Tillage Research 188, 53–58.

Chen, Songchao, Martin, Manuel P., Saby, Nicolas P.A., Walter, Christian, Angers, Denis A., Arrouays, Dominique, 2018. Fine resolution map of top- and subsoil carbon sequestration potential in France. Science of the Total Environment 630, 389–400.

Chen, Songchao, Mulder, Vera Leatitia, Martin, Manuel P., Walter, Christian, Lacoste, Marine, Richer-de-Forges, Anne C., Saby, Nicolas P.A., Loiseau, Thomas, Hu, Bifeng, Arrouays, Dominique, 2019c. Probability mapping of soil thickness by random survival forest at a national scale. Geoderma 344, 184–194.

Chen, Songchao, Richer-de-Forges, Anne C., Leatitia Mulder, Vera, Martelet, Guillaume, Loiseau, Thomas, Lehmann, Sébastien, Arrouays, Dominique, 2021a. Digital mapping of the soil thickness of loess deposits over a calcareous bedrock in central France. Catena 198, 105062. https://doi.org/10.1016/j.catena.2020.105062.

Chen, Songchao, Xu, Hanyi, Xu, Dongyun, Ji, Wenjun, Li, Shuo, Yang, Meihua, Hu, Bifeng, Zhou, Yin, Wang, Nan, Arrouays, Dominique, Shi, Zhou, 2021b. Evaluating validation strategies on the performance of soil property prediction from regional to continental spectral data. Geoderma 400, 115159. https://doi.org/10.1016/j.geoderma.2021.115159.

Chenu, Claire, Angers, Denis A., Barré, Pierre, Derrien, Delphine, Arrouays, Dominique, Balesdent, Jérôme, 2019. Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. Soil Tillage Research 188, 41–52.

Ciampalini, Rossano, Lagacherie, Philippe, Gomez, Cecile, Grünberger, Olivier, Hamrouni, Mohamed Hédi, Mekki, Insaf, Richard, Antoine, 2013. Detecting, correcting and interpreting the biases of measured soil profile data: A case study in the Cap Bon Region (Tunisia). Geoderma 192, 68–76.

de Gruijter, Jaap J., Bierkens, Marc F.P., Brus, Dick J., Knotters, Martin (Eds.), 2006. Sampling for Natural Resource Monitoring. Springer Berlin Heidelberg, Berlin, Heidelberg.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D. J., Bardgett, R.D., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 359 (6373), 320–325.

Demattê, José Alexandre Melo, Fongaro, Caio Troula, Rizzo, Rodnei, Safanelli, José Lucas, 2018. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. Remote Sensing of Environment 212, 161–175.

Dharumarajan, S., Hegde, Rajendra, Janani, N., Singh, S.K., 2019. The need for digital soil mapping in India. Geoderma Regional 16, e00204. https://doi.org/10.1016/j.geodrs.2019.e00204.

Dharumarajan, S., Kalaiselvi, B., Suputhra, Amar, Lalitha, M., Hegde, Rajendra, Singh, S. K., Lagacherie, Philippe, 2020. Digital soil mapping of key GlobalSoilMap properties in Northern Karnataka Plateau. Geoderma Regional 20, e00250. https://doi.org/10.1016/j.geodrs.2019.e00250.

Dominati, Estelle, Patterson, Murray, Mackay, Alec, 2010. A framework for classifying and quantifying the natural capital and ecosystem services of soils. Ecological Economics 69 (9), 1858–1868.

Evans, D.L., Janes-Bassett, V., Borrelli, P., Chenu, C., Ferreira, C.S., Griffiths, R.I., et al., 2021. Sustainable futures over the next decade are rooted in soil science. European Journal of Soil Science. https://doi.org/10.1111/ejss.13145.

FAO, 2011. FAO in the 21st century.

Finke, Peter A., 2012. On digital soil assessment with models and the Pedometrics agenda. Geoderma 171-172, 3–15.

Gholizadeh, Asa, Žižala, Daniel, Saberioon, Mohammadmehdi, Borůvka, Luboš, 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. Remote Sensing of Environment 218, 89–103.

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.

Goulden, M.L., Bales, R.C., 2019. California forest die-off linked to multi-year deep soil drying in 2012–2015 drought. Nature Geoscience 12 (8), 632–637.

Gray, Jonathan M., Bishop, Thomas F.A., 2016. Change in soil organic carbon stocks under 12 climate change projections over New South Wales, Australia. Soil Science Society of America Journal 80 (5), 1296–1307.

Gray, Jonathan M., Bishop, Thomas F.A., Wilford, John R., 2016. Lithology and soil relationships for soil modelling and mapping. Catena 147, 429–440.

Grundy, Michael J., Searle, Ross, Meier, Elizabeth A., Ringrose-Voase, Anthony J., Kidd, Darren, Orton, Thomas G., Triantafilis, John, Philip, Seonaid, Liddicoat, Craig, Malone, Brendan, Thomas, Mark, Gray, Jonathan, Bennett, John McLean, 2020. Digital soil assessment delivers impact across scales in Australia and the Philippines. Geoderma Regional 22, e00314. https://doi.org/10.1016/j.geodrs.2020.e00314.

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152 (3-4), 195–207.

Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: Finding solutions for global issues. Soil Science Society of America Journal 75 (4), 1201–1213.

Guevara, Mario, Olmedo, Guillermo Federico, Stell, Emma, Yigini, Yusuf, Aguilar Duarte, Yameli, Arellano Hernández, Carlos, Arévalo, Gloria E., Arroyo-Cruz, Carlos Eduardo, Bolivar, Adriana, Bunning, Sally, Bustamante Cañas, Nelson, Cruz-Gaistardo, Carlos Omar, Davila, Fabian, Dell Acqua, Martin, Encina, Arnulfo, Figueredo Tacona, Hernán, Fontes, Fernando, Hernández Herrera, José Antonio, Ibelles Navarro, Alejandro Roberto, Loayza, Veronica, Manueles, Alexandra M., Mendoza Jara, Fernando, Olivera, Carolina, Osorio Hermosilla, Rodrigo, Pereira, Gonzalo, Prieto, Pablo, Ramos, Iván Alexis, Rey Brina, Juan Carlos, Rivera, Rafael, Rodríguez-Rodríguez, Javier, Roopnarine, Ronald, Rosales Ibarra, Albán, Rosales Riveiro, Kenset Amaury, Schulz, Guillermo Andrés, Spence, Adrian, Vasques, Gustavo M., Vargas, Ronald R., Vargas, Rodrigo, 2018. No silver bullet for digital soil mapping: country-specific soil organic carbon estimates across Latin America. SOIL 4 (3), 173–193.

Helfenstein, A., Mulder, V.L., Heuvelink, G., Okx, J., 2021. BIS-3D: high resolution 3D soil maps for the Netherlands using accuracy thresholds. In EGU General Assembly Conference Abstracts, EGU21–7836.

Helmick, J.L., Nauman, T.W., Thompson, J.A., 2014. Developing and assessing prediction intervals for soil property maps derived from legacy databases. 2014. In: Arrouays D., McKenzie N.J., Hempel J., Richer-de-Forges A.C., McBratney A.B. (eds), 2014. GlobalSoilMap. Basis of the global soil information system. Taylor & Francis, CRC Press, London, p. 359–366.

Hengl, Tomislav, Mendes de Jesus, Jorge, Heuvelink, Gerard B.M., Ruiperez Gonzalez, Maria, Kilibarda, Milan, Blagotić, Aleksandar, Shangguan, Wei, Wright, Marvin N., Geng, Xiaoyuan, Bauer-Marschallinger, Bernhard, Guevara, Mario Antonio, Vargas, Rodrigo, MacMillan, Robert A., Batjes, Niels H., Leenaars, Johan G.B., Ribeiro, Eloi, Wheeler, Ichsani, Mantel, Stephan, Kempen, Bas, Bond-Lamberty, Ben, 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12 (2), e0169748. https://doi.org/10.1371/journal.pone.016974810.1371/journal.pone.0169748.g00110.1371/journal.pone.0169748.g00210.1371/journal.pone.0169748.g00310.1371/journal.pone.0169748.g00410.1371/journal.pone.0169748.g00510.1371/journal.pone.0169748.g00610.1371/journal.pone.0169748.g00710.1371/journal.pone.0169748.g00810.1371/journal.pone.0169748.g00910.1371/journal.pone.0169748.g01010.1371/journal.pone.0169748.g01110.1371/journal.pone.0169748.g01210.1371/journal.pone.0169748.g01310.1371/journal.pone.0169748.g01410.1371/journal.pone.0169748.t00110.1371/journal.pone.0169748.t00210.1371/journal.pone.0169748.t00310.1371/journal.pone.0169748.t004.

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 6, e5518.

Heuvelink, G., Angelini, M., Poggio, L., Bai, Z., Batjes, N., van den Bosch, R., Sanderman, J., 2020. Space-time machine learning for modelling soil organic carbon change. In: In EGU General Assembly Conference Abstracts, p. (p. 3621)..

Heuvelink, G.B.M., 1998. Error propagation in environmental modelling with GIS. Taylor & Francis, London.

Heuvelink, G.B.M., 2014. Uncertainty quantification of GlobalSoilMap products. In: Arrouays, D., McKenzie, N.J., Hempel, J., Richer-de-Forges, A.C., McBratney, A.B. (Eds.), GlobalSoilMap. Basis of the global soil information system. Taylor & Francis, CRC Press, London, pp. 335–340.

Hu, Bifeng, Bourennane, Hocine, Arrouays, Dominique, Denoroy, Pascal, Lemercier, Blandine, Saby, Nicolas P.A., 2021. Developing pedotransfer functions to harmonize extractable soil phosphorus content measured with different methods: A case study across the mainland of France. Geoderma 381, 114645. https://doi.org/10.1016/j.geoderma.2020.114645.

Huang, J., Hartemink, A.E., Zhang, Y., 2019. Climate and land-use change effects on soil carbon stocks over 150 years in Wisconsin, USA. Remote Sensing 11 (12), 1504.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Jansson, Janet K., Hofmockel, Kirsten S., 2020. Soil microbiomes and climate change. Nature Reviews Microbiology 18 (1), 35–46.

Jenny, H., 1941. Factors of Soil Formation: A System of Quantitative Pedology. McGraw Hill Book Company, New York, p. 281.

Jónsson, J.Ö.G., Davíðsdóttir, B., Nikolaidis, N.P., 2017. Valuation of soil ecosystem services. Advances in Agronomy 142, 353–384.

Kasraei, Babak, Heung, Brandon, Saurette, Daniel D., Schmidt, Margaret G., Bulmer, Chuck E., Bethel, William, 2021. Quantile regression as a generic approach for estimating uncertainty of digital soil maps produced from machine-learning. Environmental Modelling & Software 144, 105139. https://doi.org/10.1016/j.envsoft.2021.105139.

Keesstra, Saskia D., Bouma, Johan, Wallinga, Jakob, Tittonell, Pablo, Smith, Pete, Cerdà, Artemi, Montanarella, Luca, Quinton, John N., Pachepsky, Yakov, van der Putten, Wim H., Bardgett, Richard D., Moolenaar, Simon, Mol, Gerben, Jansen, Boris, Fresco, Louise O., 2016. The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. SOIL 2 (2), 111–128.

Kelleway, Jeffrey J., Cavanaugh, Kyle, Rogers, Kerrylee, Feller, Ilka C., Ens, Emilie, Doughty, Cheryl, Saintilan, Neil, 2017. Review of the ecosystem service implications of mangrove encroachment into salt marshes. Global Change Biology 23 (10), 3967–3983.

Kempen, Bas, Brus, Dick J., de Vries, Folkert, 2015. Operationalizing digital soil mapping for nationwide updating of the 1:50,000 soil map of the Netherlands. Geoderma 241-242, 313–329.

Kempen, Bas, Dalsgaard, Søren, Kaaya, Abel K., Chamuya, Nurdin, Ruipérez-González, Maria, Pekkarinen, Anssi, Walsh, Markus G., 2019. Mapping topsoil organic carbon concentrations and stocks for Tanzania. Geoderma 337, 164–180.

Kerry, Ruth, Oliver, Margaret A., 2008. Determining nugget: Sill ratios of standardized variograms from aerial photographs to krige sparse soil data. Precision Agriculture 9 (1-2), 33–56.

Keskin, Hamza, Grunwald, Sabine, Harris, Willie G., 2019. Digital mapping of soil carbon fractions with machine learning. Geoderma 339, 40–58.

Kidd, Darren, Searle, Ross, Grundy, Mike, McBratney, Alex, Robinson, Nathan, O'Brien, Lauren, Zund, Peter, Arrouays, Dominique, Thomas, Mark, Padarian, José, Jones, Edward, Bennett, John McLean, Minasny, Budiman, Holmes, Karen, Malone, Brendan P., Liddicoat, Craig, Meier, Elizabeth A., Stockmann, Uta, Wilson, Peter, Wilford, John, Payne, Jim, Ringrose-Voase, Anthony, Slater, Brian, Odgers, Nathan, Gray, Jonathan, van Gool, Dennis, Andrews, Kaitlyn, Harms, Ben, Stower, Liz, Triantafilis, John, 2020. Operationalising Digital Soil Mapping - Lessons from Australia. Geoderma Regional 23, e00335. https://doi.org/10.1016/j.geodrs.2020.e00335.

Kidd, Darren, Webb, Mathew, Malone, Brendan, Minasny, Budiman, McBratney, Alex, 2015. Digital soil assessment of agricultural suitability, versatility and capital in Tasmania. Australia. Geoderma. Regional 6, 7–21.

Koch, Andrea, McBratney, Alex, Adams, Mark, Field, Damien, Hill, Robert, Crawford, John, Minasny, Budiman, Lal, Rattan, Abbott, Lynette, O'Donnell, Anthony, Angers, Denis, Baldock, Jeffrey, Barbier, Edward, Binkley, Dan, Parton, William, Wall, Diana H., Bird, Michael, Bouma, Johan, Chenu, Claire, Flora, Cornelia Butler, Goulding, Keith, Grunwald, Sabine, Hempel, Jon, Jastrow, Julie, Lehmann, Johannes, Lorenz, Klaus, Morgan, Cristine L., Rice, Charles W., Whitehead, David, Young, Iain, Zimmermann, Michael, 2013. Soil security: Solving the global soil crisis. Global Policy 4 (4), 434–441.

Kuzyakov, Y., Blagodatskaya, E., 2015. Microbial hotspots and hot moments in soil: Concept and review. Soil Biology and Biochemistry 83, 184–199.

Laborczi, Annamária, Szatmári, Gábor, Kaposi, András Dezső, Pásztor, László, 2019. Comparison of soil texture maps synthetized from standard depth layers with directly compiled products. Geoderma 352, 360–372.

Lacoste, Marine, Lemercier, Blandine, Walter, Christian, 2011. Regional mapping of soil parent material by machine learning based on point data. Geomorphology 133 (1-2), 90–99.

Lacoste, M., Mulder, V.L., Richer-de-Forges, A.C., Martin, M.P., Arrouays, D., 2016. Evaluating large-extent spatial modeling approaches: A case study for soil depth for France. Geoderma Regional 7 (2), 137–152.

Lagacherie, P., Gomez, C., 2018. Vis-NIR-SWIR Remote Sensing Products as New Soil Data for Digital Soil Mapping. In: Pedometrics (pp. 415–437). Springer, Cham.

Lagacherie, P., Arrouays, D., Bourennane, H., Gomez, C., Martin, M., Saby, N.P., 2019. How far can the uncertainty on a Digital Soil Map be known?: A numerical experiment using pseudo values of clay content obtained from Vis-SWIR hyperspectral imagery. Geoderma, 337, 1320–1328.

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, 114503. https://doi.org/10.1016/j.geoderma.2020.114503.

Lagacherie, P., McBratney, A.B., 2006. Chapter 1 spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. Developments in Soil Science, 3–22.

Lal, Rattan, 2018. Digging deeper: A holistic perspective of factors affecting soil organic carbon sequestration in agroecosystems. Global Change Biology 24 (8), 3285–3301.

Lal, Rattan, Bouma, Johan, Brevik, Eric, Dawson, Lorna, Field, Damien J., Glaser, Bruno, Hatano, Ryusuke, Hartemink, Alfred E., Kosaki, Takashi, Lascelles, Bruce, Monger, Curtis, Muggler, Cristine, Ndzana, Georges Martial, Norra, Stefan, Pan, Xicai, Paradelo, Remigio, Reyes-Sánchez, Laura Bertha, Sandén, Taru, Singh, Bal Ram, Spiegel, Heide, Yanai, Junta, Zhang, Jiabao, 2021. Soils and sustainable development goals of the United Nations: An International Union of Soil Sciences perspective. Geoderma Regional 25, e00398. https://doi.org/10.1016/j.geodrs.2021.e00398.

Lal, Rattan, Smith, Pete, Jungkunst, Hermann F., Mitsch, William J., Lehmann, Johannes, Nair, P.K. Ramachandran, McBratney, Alex B., de Moraes Sá, João Carlos, Schneider, Julia, Zinn, Yuri L., Skorupa, Alba L.A., Zhang, Hai-Lin, Minasny, Budiman, Srinivasrao, Cherukumalli, Ravindranath, Nijavalli H., 2018. The carbon sequestration potential of terrestrial ecosystems. Journal of Soil and Water Conservation 73 (6), 145A–152A.

Lamichhane, Sushil, Kumar, Lalit, Wilson, Brian, 2019. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. Geoderma 352, 395–413.

Leenaars, Johan G.B., Claessens, Lieven, Heuvelink, Gerard B.M., Hengl, Tom, Ruiperez González, Maria, van Bussel, Lenny G.J., Guilpart, Nicolas, Yang, Haishun, Cassman, Kenneth G., 2018. Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. Geoderma 324, 18–36.

Liang, Peng, Qin, Cheng-Zhi, Zhu, A-Xing, 2021. Comparison on two case-based reasoning strategies of automatically selecting terrain covariates for digital soil mapping. Transactions in GIS 25 (5), 2419–2437. https://doi.org/10.1111/tgis.v25.510.1111/tgis.12831.

Liang, Zongzheng, Chen, Songchao, Yang, Yuanyuan, Zhou, Yue, Shi, Zhou, 2019. High-resolution three-dimensional mapping of soil organic carbon in China: Effects of SoilGrids products on national modeling. Science of The Total Environment 685, 480–489.

Libohova, Z., Seybold, C., Wills, S., Beaudette, D., Peaslee, S., Lindbo, D., Adhikari, K., Owens, P.R., 2019. The anatomy of uncertainty for soil pH measurements and predictions: Implications for modelers and practitioners. European Journal of Soil Science 70, 185–199.

Liu, Feng, Rossiter, David G., Song, Xiao-Dong, Zhang, Gan-Lin, Yang, Ren-Min, Zhao, Yu-Guo, Li, De-Cheng, Ju, Bing, 2016. A similarity-based method for three-dimensional prediction of soil organic matter concentration. Geoderma 263, 254–263.

Liu, Feng, Wu, Huayong, Zhao, Yuguo, Li, Decheng, Yang, Jin-Ling, Song, Xiaodong, Shi, Zhou, Zhu, A-Xing, Zhang, Gan-Lin, 2021. Mapping high resolution National Soil Information Grids of China. Science Bulletin. https://doi.org/10.1016/j.scib.2021.10.013.

Liu, Feng, Zhang, Gan-Lin, Song, Xiaodong, Li, Decheng, Zhao, Yuguo, Yang, Jinling, Wu, Huayong, Yang, Fei, 2020a. High-resolution and three-dimensional mapping of soil texture of China. Geoderma 361, 114061. https://doi.org/10.1016/j.geoderma.2019.114061.

Liu, J., Zhu, A.-X., Rossiter, D., Du, F., Burt, J.A., 2020b. Trustworthiness indicator to select sample points for the individual predictive soil mapping method (iPSM). Geoderma 373, 114440.

Loiseau, T., Chen, S., Mulder, V.L., Román Dobarco, M., Richer-de-Forges, A.C., Lehmann, S., Bourennane, H., Saby, N.P.A., Martin, M.P., Vaudour, E., Gomez, C., Lagacherie, P., Arrouays, D., 2019. Satellite data integration for soil clay content modelling at a national scale. International Journal of Applied Earth Observation and Geoinformation 82, 101905. https://doi.org/10.1016/j.jag.2019.101905.

Loiseau, Thomas, Richer-de-Forges, Anne C., Martelet, Guillaume, Bialkowski, Anne, Nehlig, Pierre, Arrouays, Dominique, 2020. Could airborne gamma-spectrometric data replace lithological maps as co-variates for digital soil mapping of topsoil particle-size distribution? A case study in Western France. Geoderma Regional 22, e00295. https://doi.org/10.1016/j.geodrs.2020.e00295.

Louis, B.P., Saby, N.P.A., Orton, T.G., Lacarce, E., Boulonne, L., Jolivet, C., Ratié, C., Arrouays, D., 2014. Statistical sampling design impact on predictive quality of harmonization functions between soil monitoring networks. Geoderma 213, 133–143.

Luo, Yiqi, Ahlström, Anders, Allison, Steven D., Batjes, Niels H., Brovkin, Victor, Carvalhais, Nuno, Chappell, Adrian, Ciais, Philippe, Davidson, Eric A., Finzi, Adien, Georgiou, Katerina, Guenet, Bertrand, Hararuk, Oleksandra, Harden, Jennifer W., He, Yujie, Hopkins, Francesca, Jiang, Lifen, Koven, Charlie, Jackson, Robert B., Jones, Chris D., Lara, Mark J., Liang, Junyi, McGuire, A. David, Parton, William, Peng, Changhui, Randerson, James T., Salazar, Alejandro, Sierra, Carlos A., Smith, Matthew J., Tian, Hanqin, Todd-Brown, Katherine E.O., Torn, Margaret, Groenigen, Kees Jan, Wang, Ying Ping, West, Tristram O., Wei, Yaxing, Wieder, William R., Xia, Jianyang, Xu, Xia, Xu, Xiaofeng, Zhou, Tao, 2016. Toward more realistic projections of soil carbon dynamics by Earth system models. Global Biogeochemical Cycles 30 (1), 40–56.

Ma, Tianwu, Brus, Dick J., Zhu, A-Xing, Zhang, Lei, Scholten, Thomas, 2020. Comparison of conditioned Latin hypercube and feature space coverage sampling for predicting soil classes using simulation from soil maps. Geoderma 370, 114366. https://doi.org/10.1016/j.geoderma.2020.114366.

Ma, Wanzhu, Zhan, Yu, Chen, Songchao, Ren, Zhouqiao, Chen, Xiaojia, Qin, Fangjin, Lu, Ruohui, Lv, Xiaonan, Deng, Xunfei, 2021a. Organic carbon storage potential of cropland topsoils in East China: Indispensable roles of cropping systems and soil managements. Soil and Tillage Research 211, 105052. https://doi.org/10.1016/j.still.2021.105052.

Ma, Yuxin, Minasny, Budiman, Malone, Brendan P., Mcbratney, Alex B., 2019a. Pedology and digital soil mapping (DSM). European Journal of Soil Science 70 (2), 216–235.

Ma, Yuxin, Minasny, Budiman, McBratney, Alex, Poggio, Laura, Fajardo, Mario, 2021b. Predicting soil properties in 3D: Should depth be a covariate? Geoderma 383, 114794. https://doi.org/10.1016/j.geoderma.2020.114794.

Ma, Y., Minasny, B., Welivitiya, W.D.D.P., Malone, B.P., Willgoose, G.R., McBratney, A.B., 2019b. The feasibility of predicting the spatial pattern of soil particle-size distribution using a pedogenesis model. Geoderma 341, 195–205.

Malone, Brendan, Searle, Ross, 2021a. Updating the Australian digital soil texture mapping (Part 1*): re-calibration of field soil texture class centroids and description of a field soil texture conversion algorithm. Soil Research 59 (5), 419. https://doi.org/10.1071/SR20283.

Malone, Brendan, Searle, Ross, 2021b. Updating the Australian digital soil texture mapping (Part 2*): spatial modelling of merged field and lab measurements. Soil Research 59 (5), 435. https://doi.org/10.1071/SR20284.

Malone, Brendan P., Jha, Sanjeev K., Minasny, Budiman, McBratney, Alex B., 2016. Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. Geoderma 262, 243–253.

Malone, Brendan, Searle, Ross, 2020. Improvements to the Australian national soil thickness map using an integrated data mining approach. Geoderma 377, 114579. https://doi.org/10.1016/j.geoderma.2020.114579.

Marchant, B.P., Lark, R.M., 2007. The Matérn variogram model: Implications for uncertainty propagation and sampling in geostatistical surveys. Geoderma 140 (4), 337–345.

Marsman, B.A., de Gruijter, J.J., 1986. Quality of soil maps. A comparison of survey methods in a sandy area. Soil Survey Papers, No. 15. Netherlands Soil Survey Institute, Wageningen, 103.

Martin, Manuel P., Dimassi, Bassem, Román Dobarco, Mercedes, Guenet, Bertrand, Arrouays, Dominique, Angers, Denis A., Blache, Fabrice, Huard, Frédéric, Soussana, Jean-François, Pellerin, Sylvain, 2021. Feasibility of the 4 per 1000 aspirational target for soil carbon: A case study for France. Global Change Biology 27 (11), 2458–2477.

McBratney, Alex, Field, Damien J., Koch, Andrea, 2014. The dimensions of soil security. Geoderma 213, 203–213.

McBratney, A.B, Mendonça Santos, M.L, Minasny, B, 2003. On digital soil mapping. Geoderma 117 (1-2), 3–52.

McKenzie, N.J., Austin, M.P., 1993. A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. Geoderma 57 (4), 329–355.

McNally, Sam R., Beare, Mike H., Curtin, Denis, Meenken, Esther D., Kelliher, Francis M., Calvelo Pereira, Roberto, Shen, Qinhua, Baldock, Jeff, 2017. Soil carbon sequestration potential of permanent pasture and continuous cropping soils in New Zealand. Global Change Biology 23 (11), 4544–4555.

Meersmans, J., Arrouays, D., Van Rompaey, A.J., Pagé, C., De Baets, S., Quine, T.A., 2016. Future C loss in mid-latitude mineral soils: Climate change exceeds land use mitigation potential in France. Scientific Reports 6, 35798.

Meersmans, J., van Wesemael, B., De Ridder, F., Fallas Dotti, M., De Baets, S., Van Molle, M., 2009a. Changes in organic carbon distribution with depth in agricultural soils in northern Belgium, 1960–2006. Global Change Biology 15 (11), 2739–2750.

Meersmans, J., van Wesemael, B., De Ridder, F., Van Molle, M., 2009b. Modelling the three-dimensional spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 152 (1-2), 43–52.

Meersmans, J., van Wesemael, B., Goidts, E., Van Molle, M., De Baets, S., De Ridder, F., 2011. Spatial analysis of soil organic carbon evolution in Belgian croplands and grasslands, 1960–2006. Global Change Biology 17 (1), 466–479.

Meinshausen, N., 2006. Quantile Regression Forests. Journal of Machine Learning Research 7, 983–999.

Meyer, Hanna, Reudenbach, Christoph, Hengl, Tomislav, Katurji, Marwan, Nauss, Thomas, 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environmental Modelling & Software 101, 1–9.

Miller, B.A., Juilleret, J., 2020. The colluvium and alluvium problem: Historical review and current state of definitions. Earth-Science Reviews, 103316.

Minasny, Budiman, Berglund, Örjan, Connolly, John, Hedley, Carolyn, de Vries, Folkert, Gimona, Alessandro, Kempen, Bas, Kidd, Darren, Lilja, Harry, Malone, Brendan, McBratney, Alex, Roudier, Pierre, O'Rourke, Sharon, Rudiyanto, Padarian, José, Poggio, Laura, ten Caten, Alexandre, Thompson, Daniel, Tuve, Clint, Widyatmanti, Wirastuti, 2019. Digital mapping of peatlands–A critical review. Earth-Science Reviews 196, 102870. https://doi.org/10.1016/j.earscirev.2019.05.014.

Minasny, Budiman, Hong, Suk Young, Hartemink, Alfred E., Kim, Yoo Hak, Kang, Seong Soo, 2016. Soil pH increase under paddy in South Korea between 2000 and 2012. Agriculture, Ecosystems & Environment 221, 205–213.

Minasny, B., Malone, B.P., McBratney, A.B., 2012. Digital Soil Assessments and Beyond. Proceedings of the 5th Global Workshop on Digital Soil Mapping, Sydney Australia, CRC Press, London, 472.

Minasny, Budiman, Malone, Brendan P., McBratney, Alex B., Angers, Denis A., Arrouays, Dominique, Chambers, Adam, Chaplot, Vincent, Chen, Zueng-Sang, Cheng, Kun, Das, Bhabani S., Field, Damien J., Gimona, Alessandro, Hedley, Carolyn B., Hong, Suk Young, Mandal, Biswapati, Marchant, Ben P., Martin, Manuel, McConkey, Brian G., Mulder, Vera Leatitia, O'Rourke, Sharon, Richer-de-Forges, Anne C., Odeh, Inakwu, Padarian, José, Paustian, Keith, Pan, Genxing, Poggio, Laura, Savin, Igor, Stolbovoy, Vladimir, Stockmann, Uta, Sulaeman, Yiyi, Tsui, Chun-Chih, Vågen, Tor-Gunnar, van Wesemael, Bas, Winowiecki, Leigh, 2017. Soil carbon 4 per mille. Geoderma 292, 59–86.

Minasny, Budiman, McBratney, Alex B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences 32 (9), 1378–1388.

Minasny, B., McBratney, A.B., 2013. Why you don't need to use RPD. Pedometron 33, 13.

Minasny, Budiman, McBratney, Alex.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311.

Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital mapping of soil carbon. Advances in Agronomy 118, 1–47.

Montagne, David, Cornu, S., 2010. Do we need to include soil evolution module in models for prediction of future climate change? Climatic Change 98 (1-2), 75–86.

Montanarella, Luca, Pennock, Daniel Jon, McKenzie, Neil, Badraoui, Mohamed, Chude, Victor, Baptista, Isaurinda, Mamo, Tekalign, Yemefack, Martin, Singh Aulakh, Mikha, Yagi, Kazuyuki, Young Hong, Suk, Vijarnsorn, Pisoot, Zhang, Gan-Lin, Arrouays, Dominique, Black, Helaina, Krasilnikov, Pavel, Sobocká, Jaroslava, Alegre, Julio, Henriquez, Carlos Roberto, de Lourdes Mendonça-Santos, Maria, Taboada, Miguel, Espinosa-Victoria, David, AlShankiti, Abdullah, AlaviPanah, Sayed Kazem, Elsheikh, Elsiddig Ahmed El Mustafa, Hempel, Jon, Camps Arbestain, Marta, Nachtergaele, Freddy, Vargas, Ronald, 2016. World's soils are under threat. SOIL 2 (1), 79–82.

Mulder, V.L., de Bruin, S., Weyermann, J., Kokaly, R.F., Schaepman, M.E., 2013. Characterizing regional soil mineral composition using spectroscopy and geostatistics. Remote Sensing of Environment 139, 415–429.

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 Environment 573, 1352–1369.

Mulder, V.L., van Eck, C.M., Friedlingstein, P., Arrouays, D., Regnier, P., 2019. Controlling factors for land productivity under extreme climatic events in continental Europe and the Mediterranean Basin. Catena 182, 104124. https://doi.org/10.1016/j.catena.2019.104124.

Nauman, Travis W., Duniway, Michael C., 2019. Relative prediction intervals reveal larger uncertainty in 3D approaches to predictive digital soil mapping of soil properties with legacy data. Geoderma 347, 170–184.

Odgers, N.P., McBratney, A.B., Minasny, B., 2014a. Digital soil property mapping and uncertainty estimation using soil class probability rasters. In: Arrouays, D., McKenzie, N.J., Hempel, J., Richer-de-Forges, A.C., McBratney, A.B. (Eds.), GlobalSoilMap. Basis of the global soil information system. Taylor & Francis, CRC Press, London, pp. 341–346.

Odgers, Nathan P., Sun, Wei, McBratney, Alex B., Minasny, Budiman, Clifford, David, 2014b. Disaggregating and harmonising soil map units through resampled classification trees. Geoderma 214-215, 91–100.

Omuto, Christian T., Vargas, Ronald R., 2015. Re-tooling of regression kriging in R for improved digital mapping of soil properties. Geosciences Journal 19 (1), 157–165.

O'Rourke, Sharon M., Angers, Denis A., Holden, Nicholas M., McBratney, Alex B., 2015. Soil organic carbon across scales. Global change biology 21 (10), 3561–3574.

Ottoy, Sam, Van Meerbeek, Koenraad, Sindayihebura, Anicet, Hermy, Martin, Van Orshoven, Jos, 2017. Assessing top- and subsoil organic carbon stocks of Low-Input High-Diversity systems using soil and vegetation characteristics. Science of The Total Environment 589, 153–164.

Padarian, J., Minasny, B., McBratney, A.B., 2017. Chile and the Chilean soil grid: a contribution to GlobalSoilMap. Geoderma Regional 9, 17–28.

Padarian, J., Minasny, B., McBratney, A.B., 2019. Using deep learning for digital soil mapping. SOIL 5 (1), 79–89.
Padarian, José, Minasny, Budiman, McBratney, Alex B., 2020. Machine learning and soil sciences: A review aided by machine learning tools. SOIL 6 (1), 35–52.
Pfeiffer, Marco, Padarian, José, Osorio, Rodrigo, Bustamante, Nelson, Olmedo, Guillermo Federico, Guevara, Mario, Aburto, Felipe, Albornoz, Francisco, Antilén, Monica, Araya, Elías, Arellano, Eduardo, Barret, Maialen, Barrera, Juan, Boeckx, Pascal, Briceño, Margarita, Bunning, Sally, Cabrol, Lea, Casanova, Manuel, Cornejo, Pablo, Corradini, Fabio, Curaqueo, Gustavo, Doetterl, Sebastian, Duran, Paola, Escudey, Mauricio, Espinoza, Angelina, Francke, Samuel, Fuentes, Juan Pablo, Fuentes, Marcel, Gajardo, Gonzalo, García, Rafael, Gallaud, Audrey, Galleguillos, Mauricio, Gomez, Andrés, Hidalgo, Marcela, Ivelic-Sáez, Jorge, Mashalaba, Lwando, Matus, Francisco, Meza, Francisco, Mora, Maria de la Luz, Mora, Jorge, Muñoz, Cristina, Norambuena, Pablo, Olivera, Carolina, Ovalle, Carlos, Panichini, Marcelo, Pauchard, Aníbal, Pérez-Quezada, Jorge F., Radic, Sergio, Ramirez, José, Riveras, Nicolás, Ruiz, Germán, Salazar, Osvaldo, Salgado, Iván, Seguel, Oscar, Sepúlveda, Maria, Sierra, Carlos, Tapia, Yasna, Tapia, Francisco, Toledo, Balfredo, Torrico, José Miguel, Valle, Susana, Vargas, Ronald, Wolff, Michael, Zagal, Erick, 2020. CHLSOC: the Chilean Soil Organic Carbon database, a multi-institutional collaborative effort. Earth System Science Data 12 (1), 457–468.
Phillips, H.R., Guerra, C.A., Bartz, M.L., Briones, M.J., Brown, G., Crowther, T.W., Orgiazzi, A., 2019. Global distribution of earthworm diversity. Science 366 (6464), 480–485.
Piikki, Kristin, Wetterlind, Johanna, Söderström, Mats, Stenberg, Bo, 2021. Perspectives on validation in digital soil mapping of continuous attributes—A review. Soil Use and Management 37 (1), 7–21.
Ploton, Pierre, Mortier, Frédéric, Réjou-Méchain, Maxime, Barbier, Nicolas, Picard, Nicolas, Rossi, Vivien, Dormann, Carsten, Cornu, Guillaume, Viennois, Gaëlle, Bayol, Nicolas, Lyapustin, Alexei, Gourlet-Fleury, Sylvie, Pélissier, Raphaël, 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nature communications 11 (1). https://doi.org/10.1038/s41467-020-18321-y.
Poggio, Laura, de Sousa, Luis M., Batjes, Niels H., Heuvelink, Gerard B.M., Kempen, Bas, Ribeiro, Eloi, Rossiter, David, 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7 (1), 217–240.
Poggio, Laura, Gimona, Alessandro, 2014. National scale 3D modelling of soil organic carbon stocks with uncertainty propagation—An example from Scotland. Geoderma 232-234, 284–299.
Poggio, Laura, Gimona, Alessandro, 2017. 3D mapping of soil texture in Scotland. Geoderma Regional 9, 5–16.
Poggio, L., Gimona, A., 2018. Towards GlobalSoilMap products in Scotland. In: GlobalSoilMap-Digital Soil Mapping from Country to Globe. CRC Press, pp. 7–12.
Pries, C.E.H., Castanha, C., Porras, R.C., Torn, M.S., 2017. The whole-soil carbon flux in response to warming. Science 355 (6332), 1420–1423.
Pahlavan Rad, Mohammad Reza, Toomanian, Norair, Khormali, Farhad, Brungard, Colby W., Komaki, Chooghi Bayram, Bogaert, Patrick, 2014. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. Geoderma 232-234, 97–106.
Ramcharan, A., Hengl, T., Nauman, T., Brungard, C., Waltman, S., Wills, S., Thompson, J., 2018. Soil property and class maps of the conterminous US at 100 meter spatial resolution based on a compilation of national soil point observations and machine learning. Soil Science Society of America Journal 82, 186–201.
Reddy, Nagarjuna N., Chakraborty, Poulamee, Roy, Sourav, Singh, Kanika, Minasny, Budiman, McBratney, Alex B., Biswas, Asim, Das, Bhabani S., 2021. Legacy data-based national-scale digital mapping of key soil properties in India. Geoderma 381, 114684. https://doi.org/10.1016/j.geoderma.2020.114684.
Rentschler, Tobias, Gries, Philipp, Behrens, Thorsten, Bruelheide, Helge, Kühn, Peter, Seitz, Steffen, Shi, Xuezheng, Trogisch, Stefan, Scholten, Thomas, Schmidt, Karsten, Minasny, Budiman, 2019. Comparison of catchment scale 3D and 2.5 D modelling of soil organic carbon stocks in Jiangxi Province, PR China. PLoS One 14 (8), e0220881. https://doi.org/10.1371/journal.pone.022088110.1371/journal.pone.0220881.g00110.1371/journal.pone.0220881.g00210.1371/journal.pone.0220881.g00310.1371/journal.pone.0220881.g00410.1371/journal.pone.0220881.g00510.1371/journal.pone.0220881.g00610.1371/journal.pone.0220881.t00110.1371/journal.pone.0220881.t00210.1371/journal.pone.0220881.t00310.1371/journal.pone.0220881.t00410.1371/journal.pone.0220881.t005.
Reyes Rojas, Luis A., Adhikari, Kabindra, Ventura, Stephen J., 2018. Projecting soil organic carbon distribution in central Chile under future climate scenarios. Journal of Environmental Quality 47 (4), 735–745.
Rial, Marcela, Cortizas, Antonio Martínez, Rodríguez-Lado, Luis, 2016. Mapping soil organic carbon content using spectroscopic and environmental data: A case study in acidic soils from NW Spain. Science of the Total Environment 539, 26–35.
Richer-de-Forges, Anne C., Arrouays, Dominique, Bardy, Marion, Bispo, Antonio, Lagacherie, Philippe, Laroche, Bertrand, Lemercier, Blandine, Sauter, Joëlle, Voltz, Marc, 2019. Mapping of Soils and Land-Related Environmental attributes in France: analysis of end-users' needs. Sustainability 11 (10), 2940. https://doi.org/10.3390/su11102940.
Richer-de-Forges, Anne C., Saby, Nicolas P.A., Mulder, Vera L., Laroche, Bertrand, Arrouays, Dominique, 2017. Probability mapping of iron pan presence in sandy podzols in South-West France, using digital soil mapping. Geoderma Regional 9, 39–46.
Robinson, D.A., Fraser, I., Dominati, E.J., Davíðsdóttir, B., Jónsson, J.O.G., Jones, L., Jones, S.B., Tuller, M., Lebron, I., Bristow, K.L., Souza, D.M., Banwart, S., Clothier, B. E., 2014. On the Value of Soil Resources in the Context of Natural Capital and Ecosystem Service Delivery. Soil Science Society of America Journal 78 (3), 685–700.
Román Dobarco, Mercedes, Arrouays, Dominique, Lagacherie, Philippe, Ciampalini, Rossano, Saby, Nicolas P.A., 2017. Prediction of topsoil texture for Region Centre (France) applying model ensemble methods. Geoderma 298, 67–77.
Román Dobarco, M., Bourennane, Hocine, Arrouays, Dominique, Saby, Nicolas P.A., Cousin, Isabelle, Martin, Manuel P., 2019a. Uncertainty assessment of GlobalSoilMap soil available water capacity products: A French case study. Geoderma 344, 14–30.
Román Dobarco, Mercedes, Cousin, Isabelle, Le Bas, Christine, Martin, Manuel P., 2019b. Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty. Geoderma 336, 81–95.
Rossiter, David G., Liu, Jing, Carlisle, Steve, Zhu, A.-Xing, 2015. Can citizen science assist digital soil mapping? Geoderma 259-260, 71–80.
Roudier, P., Burge, O.R., Richardson, S.J., McCarthy, J.K., Grealish, G.J., Ausseil, A.G., 2020. National Scale 3D Mapping of Soil pH Using a Data Augmentation Approach. Remote Sensing 12 (18), 2872.
Rumpel, Cornelia, Amiraslani, Farshad, Koutika, Lydie-Stella, Smith, Pete, Whitehead, David, Wollenberg, Eva, 2018. Put more carbon in soils to meet Paris climate pledges. Nature 564 (7734), 32–34.
Sanchez, Pedro A., Ahamed, Sonya, Carré, Florence, Hartemink, Alfred E., Hempel, Jonathan, Huising, Jeroen, Lagacherie, Philippe, McBratney, Alex B., McKenzie, Neil J., Mendonça-Santos, Maria de Lourdes, Minasny, Budiman, Montanarella, Luca, Okoth, Peter, Palm, Cheryl A., Sachs, Jeffrey D., Shepherd, Keith D., Vågen, Tor-Gunnar, Vanlauwe, Bernard, Walsh, Markus G., Winowiecki, Leigh A., Zhang, Gan-Lin, 2009. Digital soil map of the world. Science 325 (5941), 680–681.
Scharlemann, Jörn P.W., Tanner, Edmund V.J., Hiederer, Roland, Kapos, Valerie, 2014. Global soil carbon: Understanding and managing the largest terrestrial carbon pool. Carbon Management 5 (1), 81–91.
Schillaci, Calogero, Acutis, Marco, Lombardo, Luigi, Lipani, Aldo, Fantappiè, Maria, Märker, Michael, Saia, Sergio, 2017. Spatiotemporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. Science of the Total Environment 601-602, 821–832.
Schlaepfer, Daniel R., Bradford, John B., Lauenroth, William K., Munson, Seth M., Tietjen, Britta, Hall, Sonia A., Wilson, Scott D., Duniway, Michael C., Jia, Gensuo, Pyke, David A., Lkhagva, Ariuntsetseg, Jamiyansharav, Khishigbayar, 2017. Climate change reduces extent of temperate drylands and intensifies drought in deep soils. Nature Communications 8 (1). https://doi.org/10.1038/ncomms14196.
Scott, M.S., Edwards, S., Dayan, S., Nguyen, T., Cragle, J., 2016. GIS story maps: A tool to engage stakeholders in planning sustainable places. Final Report, MATS-UTC, Charlottesville, VA, USA.
Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: a review. Progress in Physical Geography 27 (2), 171–197.
Sekulić, Aleksandar, Kilibarda, Milan, Heuvelink, Gerard B.M., Nikolić, Mladen, Bajat, Branislav, 2020. Random Forest Spatial Interpolation. Remote Sensing 12 (10), 1687. https://doi.org/10.3390/rs12101687.
Seybold, Cathy A., Ferguson, Rich, Wysocki, Doug, Bailey, Scarlett, Anderson, Joe, Nester, Brian, Schoeneberger, Phil, Wills, Skye, Libohova, Zamir, Hoover, Dave, Thomas, Pam, 2019. Application of mid-infrared spectroscopy in soil survey. Soil Science Society of America Journal 83 (6), 1746–1759.
Shangguan, Wei, Hengl, Tomislav, Mendes de Jesus, Jorge, Yuan, Hua, Dai, Yongjiu, 2017. Mapping the global depth to bedrock for land surface modeling. Journal of Advances in Modeling Earth Systems 9 (1), 65–88.
Shi, Z., Ji, W., Viscarra Rossel, R.A., Chen, S., Zhou, Y., 2015. Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis–NIR spectral library. European Journal of Soil Science 66 (4), 679–687.
Silatsa, Francis B.T., Yemefack, Martin, Tabi, Fritz O., Heuvelink, Gerard B.M., Leenaars, Johan G.B., 2020. Assessing countrywide soil organic carbon stock using hybrid machine learning modelling and legacy soil data in Cameroon. Geoderma 367, 114260. https://doi.org/10.1016/j.geoderma.2020.114260.
Simon, Alois, Wilhelmy, Marcus, Klosterhuber, Ralf, Cocuzza, Elena, Geitner, Clemens, Katzensteiner, Klaus, 2021. A system for classifying subsolum geological substrates as a basis for describing soil formation. Catena 198, 105026. https://doi.org/10.1016/j.catena.2020.105026.
Soil Survey Staff, 2014. Kellogg Soil Survey Laboratory Methods Manual. Soil Survey Investigations Report No. 42, Version 5.0. R. Burt and Soil Survey Staff (eds.). U.S. Department of Agriculture, Natural Resources Conservation Service.
Song, Xiao-Dong, Yang, Fan, Ju, Bing, Li, De-Cheng, Zhao, Yu-Guo, Yang, Jin-Ling, Zhang, Gan-Lin, 2018. The influence of the conversion of grassland to cropland on changes in soil organic carbon and total nitrogen stocks in the Songnen Plain of Northeast China. Catena 171, 588–601.
Stockmann, Uta, Padarian, José, McBratney, Alex, Minasny, Budiman, de Brogniez, Delphine, Montanarella, Luca, Hong, Suk Young, Rawlins, Barry G., Field, Damien J., 2015. Global soil organic carbon assessment. Global Food Security 6, 9–16.
Sun, Xiao-Lin, Minasny, Budiman, Wang, Hui-Li, Zhao, Yu-Guo, Zhang, Gan-Lin, Wu, Yun-Jin, 2021. Spatiotemporal modelling of soil organic matter changes in Jiangsu, China between 1980 and 2006 using INLA-SPDE. Geoderma 384, 114808. https://doi.org/10.1016/j.geoderma.2020.114808.
Sun, Xiao-Lin, Wang, Yidong, Wang, Hui-Li, Zhang, Chaosheng, Wang, Zhong-Liang, 2019. Digital soil mapping based on empirical mode decomposition components of environmental covariates. European Journal of Soil Science 70 (6), 1109–1127.

Sun, X.L., Zhao, Y.G., Wu, Y.J., Zhao, M.S., Wang, H.L., Zhang, G.L., 2012. Spatio-temporal change of soil organic matter content of Jiangsu Province, China, based on digital soil maps. Soil Use and Management 28 (3), 318–328.
Taghizadeh-Mehrjardi, R., Mahdianpari, M., Mohammadimanesh, F., Behrens, T., Toomanian, N., Scholten, T., Schmidt, K., 2020a. Multi-task convolutional neural networks outperformed random forest for mapping soil particle size fractions in central Iran. Geoderma 376, 114552. https://doi.org/10.1016/j.geoderma.2020.114552.
Taghizadeh-Mehrjardi, Ruhollah, Schmidt, Karsten, Eftekhari, Kamran, Behrens, Thorsten, Jamshidi, Mohammad, Davatgar, Naser, Toomanian, Norair, Scholten, Thomas, 2020b. Synthetic resampling strategies and machine learning for digital soil mapping in Iran. European Journal of Soil Science 71 (3), 352–368.
Taghizadeh-Mehrjardi, Ruhollah, Schmidt, Karsten, Toomanian, Norair, Heung, Brandon, Behrens, Thorsten, Mosavi, Amirhosein, S. Band, Shahab, Amirian-Chakan, Alireza, Fathabadi, Aboalhasan, Scholten, Thomas, 2021. Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models. Geoderma 383, 114793. https://doi.org/10.1016/j.geoderma.2020.114793.
Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., Smith, M. E., 2014. Global diversity and geography of soil fungi. Science 346 (6213), 1256688.
Thomas, M., Clifford, D., Bartley, R., Philip, S., Brough, D., Gregory, L., Willis, R., Glover, M., 2015. Putting regional digital soil mapping into practice in Tropical Northern Australia. Geoderma 241-242, 145–157.
Turner, Katrine Grace, Anderson, Sharolyn, Gonzales-Chang, Mauricio, Costanza, Robert, Courville, Sasha, Dalgaard, Tommy, Dominati, Estelle, Kubiszewski, Ida, Ogilvy, Sue, Porfirio, Luciana, Ratna, Nazmun, Sandhu, Harpinder, Sutton, Paul C., Svenning, Jens-Christian, Turner, Graham Mark, Varennes, Yann-David, Voinov, Alexey, Wratten, Stephen, 2016. A review of methods, data, and models to assess changes in the value of ecosystem services from land degradation and restoration. Ecological Modelling 319, 190–207.
van den Hoogen, J., Geisen, S., Routh, D., Ferris, H., Traunspurger, W., Wardle, D.A., Bardgett, R.D., 2019. Soil nematode abundance and functional group composition at a global scale. Nature 572 (7768), 194–198.
Vaudour, E., Gomez, C., Fouad, Y., Lagacherie, P., 2019. Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. Remote Sensing of Environment 223, 21–33.
Vaysse, K., Heuvelink, G.B.M., Lagacherie, P., Goss, Michael, 2017. Spatial aggregation of soil property predictions in support of local land management. Soil Use and Management 33 (2), 299–310.
Vaysse, Kévin, Lagacherie, Philippe, 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. Geoderma 291, 55–64.
Viscarra Rossel, R.A., Adamchuk, V.I., Sudduth, K.A., McKenzie, N.J., Lobsey, C., 2011. Proximal soil sensing: An effective approach for soil measurements in space and time. Advances in Agronomy 113, 243–291.
Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Böttcher, K., Brodský, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C.B., Knadel, M., Morrás, H.J.M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E.M. Rufasto, Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews 155, 198–230.
Viscarra Rossel, R.A., Chen, C., Grundy, M.J., Searle, R., Clifford, D., Campbell, P.H., 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. Soil Research 53 (8), 845–864.
Viscarra Rossel, Raphael A., Lobsey, Craig R., Sharman, Chris, Flick, Paul, McLachlan, Gordon, 2017. Novel proximal sensing for monitoring soil organic C stocks and condition. Environmental Science & Technology 51 (10), 5630–5641.
Wadoux, Alexandre M.J.-C., 2019. Using deep learning for multivariate mapping of soil with quantified uncertainty. Geoderma 351, 59–70.
Wadoux, Alexandre M.J-C., Brus, Dick J., Heuvelink, Gerard B.M., 2019a. Sampling design optimization for soil mapping with random forest. Geoderma 355, 113913. https://doi.org/10.1016/j.geoderma.2019.113913.
Wadoux, A.M.C., Marchant, B.P., Lark, R.M., 2019b. Efficient sampling for geostatistical surveys. European Journal of Soil Science 70, 975–989.
Wadoux, Alexandre M.J.-C., Minasny, Budiman, McBratney, Alex B., 2020. Machine learning for digital soil mapping: applications, challenges and suggested solutions. Earth-Science Reviews 210, 103359. https://doi.org/10.1016/j.earscirev.2020.103359.
Wadoux, A.M.C., Padarian, J., Minasny, B., 2019c. Multi-source data integration for soil mapping using deep learning. SOIL 5 (1), 107–119.
Walter, C., Lagacherie, P., Follain, S., 2006. Integrating pedological knowledge into digital soil mapping. Developments in Soil Science 31, 281–615.
Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Computers & Geosciences 36 (10), 1261–1267.
Wieder, William R., Allison, Steven D., Davidson, Eric A., Georgiou, Katerina, Hararuk, Oleksandra, He, Yujie, Hopkins, Francesca, Luo, Yiqi, Smith, Matthew J., Sulman, Benjamin, Todd-Brown, Katherine, Wang, Ying-Ping, Xia, Jianyang, Xu, Xiaofeng, 2015. Explicitly representing soil microbial processes in Earth system models. Global Biogeochemical Cycles 29 (10), 1782–1800.
Wiesmeier, Martin, Hübner, Rico, Spörlein, Peter, Geuß, Uwe, Hangen, Edzard, Reischl, Arthur, Schilling, Bernd, von Lützow, Margit, Kögel-Knabner, Ingrid, 2014. Carbon sequestration potential of soils in southeast Germany derived from stable soil organic carbon saturation. Global Change Biology 20 (2), 653–665.
Wiesmeier, Martin, Mayer, Stefanie, Burmeister, Johannes, Hübner, Rico, Kögel-Knabner, Ingrid, 2020. Feasibility of the 4 per 1000 initiative in Bavaria: A reality check of agricultural soil management and carbon sequestration scenarios. Geoderma 369, 114333. https://doi.org/10.1016/j.geoderma.2020.114333.
Wiesmeier, Martin, Urbanski, Livia, Hobley, Eleanor, Lang, Birgit, von Lützow, Margit, Marin-Spiotta, Erika, van Wesemael, Bas, Rabot, Eva, Ließ, Mareike, Garcia-Franco, Noelia, Wollschläger, Ute, Vogel, Hans-Jörg, Kögel-Knabner, Ingrid, 2019. Soil organic carbon storage as a key function of soils—A review of drivers and indicators at various scales. Geoderma 333, 149–162.
Yang, Lin, Cai, Yanyan, Zhang, Lei, Guo, Mao, Li, Anqi, Zhou, Chenghu, 2021. A deep learning method to predict soil organic carbon content at a regional scale using satellite-based phenology variables. International Journal of Applied Earth Observation and Geoinformation 102, 102428. https://doi.org/10.1016/j.jag.2021.102428.
Yang, Lin, Qi, Feng, Zhu, A-Xing, Shi, Jingjing, An, Yiming, 2016. Evaluation of Integrative Hierarchical Stepwise Sampling for Digital Soil Mapping. Soil Science Society of America Journal 80 (3), 637–651.
Yigini, Yusuf, Panagos, Panos, 2016. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. Science of the Total Environment 557-558, 838–850.
Zeng, Canying, Qi, Feng, Zhu, A-Xing, Liu, Feng, 2020. Construction of land surface dynamic feedback for digital soil mapping considering the spatial heterogeneity of rainfall magnitude. Catena 191, 104576. https://doi.org/10.1016/j.catena.2020.104576.
Zeraatpisheh, Mojtaba, Jafari, Azam, Bagheri Bodaghabadi, Mohsen, Ayoubi, Shamsollah, Taghizadeh-Mehrjardi, Ruhollah, Toomanian, Norair, Kerry, Ruth, Xu, Ming, 2020. Conventional and digital soil mapping in Iran: Past, present, and future. Catena 188, 104424. https://doi.org/10.1016/j.catena.2019.104424.
ZHANG, Gan-lin, LIU, Feng, SONG, Xiao-dong, 2017. Recent progress and future prospect of digital soil mapping: A review. Journal of Integrative Agriculture 16 (12), 2871–2885.
Zhou, Tao, Geng, Yajun, Chen, Jie, Pan, Jianjun, Haase, Dagmar, Lausch, Angela, 2020. High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. Science of The Total Environment 729, 138244. https://doi.org/10.1016/j.scitotenv.2020.138244.
Zhou, Tao, Geng, Yajun, Ji, Cheng, Xu, Xiangrui, Wang, Hong, Pan, Jianjun, Bumberger, Jan, Haase, Dagmar, Lausch, Angela, 2021. Prediction of soil organic carbon and the C: N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. Science of The Total Environment 755, 142661. https://doi.org/10.1016/j.scitotenv.2020.142661.
Zhou, Yin, Biswas, Asim, Ma, Zhiqiang, Lu, Yanli, Chen, Qiuxiao, Shi, Zhou, 2016. Revealing the scale-specific controls of soil organic matter at large scale in Northeast and North China Plain. Geoderma 271, 71–79.
Zhou, Yin, Hartemink, Alfred E., Shi, Zhou, Liang, Zongzheng, Lu, Yanli, 2019. Land use and climate change effects on soil organic carbon in North and Northeast China. Science of the Total Environment 647, 1230–1238.
Zomer, R.J., Bossio, D.A., Sommer, R., Verchot, L.V., 2017. Global sequestration potential of increased organic carbon in cropland soils. Scientific Reports 7 (1), 15554.