Geoderma 449 (2024) 117007

Contents lists available at ScienceDirect

Geoderma

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

Applications and challenges of digital soil mapping in Africa

Andree M. Nenkam [a,*], Alexandre M.J-C. Wadoux [b], Budiman Minasny [a], Francis B.T. Silatsa [c], Martin Yemefack [d], Sabastine Ugbemuna Ugbaje [e], Stephen Akpa [f], George Van Zijl [g], Abdelkrim Bouasria [h,i], Yassine Bouslihim [j], Lydia Mumbi Chabala [k], Ashenafi Ali [l], Alex B. McBratney [a]

[a] Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia
[b] LISAH, Univ. Montpellier, AgroParisTech, INRAE, IRD, L'Institut Agro, Montpellier, France
[c] ISRIC – World Soil Information, PO Box 353, 6700 AJ Wageningen, The Netherlands
[d] Sustainable Tropical Actions (STA), a non-profit organisation dedicated to the protection of the environment and fight against climate change, Yaounde, Cameroon
[e] CSIRO Agriculture and Food, Ngunnawal Country, Clunies Ross Street, Black Mountain, ACT 2601, Australia
[f] Faculty of Higher Education, Holmes Institute, Sydney NSW 2000, Australia
[g] Unit for Environmental Sciences and Management, North-West University, Potchefstroom 2520, South Africa
[h] Chouaib Doukkali University, El Jadida, Morocco
[i] Agmetrix, El Jadida, Morocco
[j] National Institute of Agricultural Research, CRRA Tadla, Morocco
[k] University of Zambia, School of Agricultural Sciences, Department of Soil Science, P.O. Box 32379, Lusaka, Zambia
[l] Department of Geography and Environmental Studies, Addis Ababa University (AAU), Addis Ababa, Ethiopia

ARTICLE INFO

Handling Editor: Morgan Cristine L.S.

Keywords:
Predictive models
Soil information
Legacy soil data
Digital soil assessment
Sampling density

ABSTRACT

The mapping of soils in Africa is at least a century old. We currently have access to various maps depicting mapping units locally and for the continent. In the past two decades, there has been a growing interest in alternatives for generating soil maps through digital soil mapping (DSM) techniques. There are, however, numerous challenges pertaining to the implementation of DSM in Africa, such as the unavailability of appropriate covariates, age and positional error in the measurements, low sampling density, and spatial clustering of the soil data used to fit and validate the models. This review aims to investigate the current state of DSM in Africa, identify challenges specific to implementing DSM in Africa and the ways it has been solved in the literature. We found that nearly half of African countries had an existing digital soil map covering either a local or national area, and that most studies were performed at a local extent. Soil carbon was the most common property under study, whereas soil hydraulic variables were seldom reported. Nearly all studies performed mapping for the topsoil up to 30 cm and calculated validation statistics using existing datasets but without collecting a post-mapping probability sample. Few studies (i.e., 11%) reported an estimate of map uncertainty. Half of the studies had in mind a downstream application (e.g., soil fertility assessment) in the map generation. We further correlated the area of study and sampling density and found a strong negative relationship. About 30% of the studies relied on legacy soil datasets and had a lack of sufficient spatial coverage of their area of study. From this review, we highlight some research opportunities and suggest improvements in the current methodologies. Future research should focus on capacity building in DSM, new data collection, and legacy data rescue. New initiatives, that should be initiated and led from within the continent, could support the long-term monitoring of soils and updating of soil information systems while ensuring their contextualised usability. This pairs with better delivery of existing DSM studies to stakeholders and the generation of a value-added proposition to governmental institutions.

1. Introduction

The mapping of soils in Africa is at least a century old. In 1923, Shantz et al. (1923) created a map that depicted vegetation and soil mapping units for the continent. Later, several maps of soil for small localities (e.g., Hornby, 1938; Trapnell and Clothier, 1937; Trapnell et al., 1948) or large geographical areas were created (e.g., Baeyens, 1938; FAO-UNESCO, 1977; Van Ranst et al., 2010). One of the most

* Correspondence to: Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, New South Wales, Australia.
E-mail address: andree.nenkam@sydney.edu.au (A.M. Nenkam).

https://doi.org/10.1016/j.geoderma.2024.117007
Received 2 February 2024; Received in revised form 1 August 2024; Accepted 12 August 2024
Available online 26 August 2024

comprehensive soil maps covering the continent is the polygon map from the FAO/UNESCO (FAO-UNESCO, 1977). The report that accompanies the FAO soil map of Africa reveals the diversity of soils spanning Africa: Yermosols in the arid zones of Northern Africa, Xerosols across South Africa and Namibia, Vertisols covering Ethiopia and South-Sudan, or Fluvisols in central Africa (Cameroon and Central African Republic). The report also gave information on the soil suitability for crops and cultivation and concluded that highly fertile soils, suitable for crop cultivation, are mostly located across tropical areas. Such soils are, for example, Luvisols in West Africa (Nigeria, Benin, Ghana) and Cambisols in North Africa (northern Morocco, northern Algeria and Tunisia). In the last decade, new initiatives have built upon the FAO/UNESCO map and have attempted to provide new insights into the distribution and variety of soil types and properties in Africa, some recent examples include the Soil Atlas of Africa (Jones et al., 2013), the harmonised soil map of Africa at the continental scale (Dewitte et al., 2013), and the recent harmonised soil property values for broad-scale modelling (WISE30sec, Batjes, 2016). A recent overview of legacy soil maps compilation for the continent can also be found in Arrouays et al. (2017). In most cases, however, soils were arranged into homogeneous or association classes using polygon map units.

While historical polygon-based or legacy soil maps offer advantages for both coarse and fine-scale applications - such as identifying areas suitable for agriculture or quantifying agricultural productivity - they exhibit several limitations (Van Ranst et al., 2010). They are usually generated using conceptual soil-landscape knowledge that may not be consistent amongst surveyors. Also, soil polygon maps provide aggregated soil information over broad areas and are not intended to directly provide detailed point-specific values of soil chemical, physical, and biological properties (Scull et al., 2003; Lagacherie et al., 2006). Finally, there is usually no quantitative estimate of the accuracy of the polygon-based soil maps, making them insufficient for practitioners and multidisciplinary environmental studies, which require an estimate of the level of confidence that we have about the map. In parallel, there is a growing demand for quantitative digital information on soil properties to inform private landowners, land managers, modellers (e.g., crop and hydrology modellers), policy and decision makers. Integrating polygon maps into current and future modelling studies that require grid-based quantitative estimates of soil properties can be challenging (Hartemink et al., 2010).

An approach which addresses the limitations of polygon-based maps is digital soil mapping (DSM), where empirical models are fitted between measurements of soil properties and a set of environmental covariates. The covariates represent the various *scorpan* factors which can be used as a basis for the spatial prediction of soil properties and classes, and in which s stands for soil, c for climate, o for organism, r for relief, p for parent material, a for age and n for geographical location (McBratney et al., 2003). The empirical relationship can be fitted with conventional statistical models (e.g., linear and non-linear regression, generalised linear models, generalised additive models), geostatistics (e.g., ordinary kriging, linear model of co-regionalisation), machine learning or deep learning algorithms (e.g., regression tree models, artificial neural network), ensemble of predictive algorithms, or a combination thereof.

Perusal of the literature shows that DSM has gained increasing attention in Africa. Examples of DSM studies are Kempen (2005), Cambule et al. (2013), Were et al. (2016), Chabala et al. (2017), Hounkpatin et al. (2018b), Leenaars et al. (2020) and Takoutsing et al. (2022) for landscape-scale (i.e., 500 to 15,000 km^2), Ugbaje and Reuter (2013), Omuto (2013), Chabala et al. (2014), Ramifehiarivo et al. (2017), Venter et al. (2021), Bahri et al. (2022) and Ali et al. (2024) and Silatsa et al. (2020) for larger areas, such as a nation, and Hengl et al. (2015), Vågen et al. (2016), Hengl et al. (2017a) and Leenaars et al. (2018) and Hengl et al. (2021b) for the whole continent. For a small geographical area, Kempen (2005) made a map of the spatial distribution of soil organic carbon (SOC) in the Nioro du Rip catchment in Senegal with a grid spacing of 30 m using a classification tree. At a national extent, Ugbaje and Reuter (2013) predicted numerous soil properties (i.e., sand, clay, bulk density, and available water content) in Nigeria with a grid spacing of 100 m using a regression tree cubist model coupled with kriging of the residuals. In addition, countrywide predictions of soil organic carbon stocks were produced for several countries, for instance, Nigeria, Madagascar, Ghana, and Cameroon (Akpa et al., 2016a; Ramifehiarivo et al., 2017; Owusu et al., 2020; Silatsa et al., 2020). At the continental scale, Leenaars et al. (2018) made maps of rootable depth and root zone plant-available water holding capacity at 250 m resolution, while Hengl et al. (2021b) made maps of several soil properties (e.g., SOC, pH, total nitrogen, and exchangeable basis) at 30 m resolution. Among other findings, these studies concluded that implementing DSM for Africa was challenging for several reasons, including the unavailability of appropriate covariates, poor soil-covariate structure, age and positional error in the measurement, and the low sampling density used to build the models.

There are, indeed, several challenges with the application of DSM in Africa. One such challenge, usually considered as a global issue in implementing DSM (Wadoux et al., 2020; Chen et al., 2022), is the availability, quality and quantity of soil databases used to calibrate and validate the DSM models. While new datasets are constantly being gathered, they are usually held by a number of institutions and may not be accessible to the public (Paterson et al., 2015). As a result, to date, several DSM studies carried out in the continent use the same source of calibration data, i.e., the freely available African Soil Profile (AfSP) database (Leenaars et al., 2014) which comprises legacy soil data for both the top and the subsoils. Nonetheless, the data are spatially clustered because they are a compendium of smaller datasets gathered from multiple sources (e.g., reconnaissance surveys, legacy soil maps, soil surveys, project reports and published documents). Further, some of these datasets may not reflect the current soil condition since they were collected more than 40 years ago (Hengl et al., 2015), or were measured from various laboratory methods, some of which may result in large measurement errors (Odeh et al., 2012). Finally, the density of the data available to calibrate and validate a DSM model is relatively small compared to studies performed in other parts of the world, such as Australia (e.g., Román Dobarco et al., 2022) or Europe (e.g., Mulder et al., 2016). Ugbaje and Reuter (2013), for instance, only had 251 bulk density measurements to produce a digital map of the available water capacity in Nigeria, resulting in approximately 1 measurement per 3000 km^2. All these, taken together, affect the quality of the digital soil maps but also the ability of the practitioner to trust the maps and use them in downstream applications.

In this paper, we review the current state of DSM in Africa, identify challenges specific to implementing DSM in Africa, highlight some research opportunities, and suggest improvements. It is the first time that such an exhaustive review has been conducted for the whole continent of Africa. The paper is divided into three sections. First, the current state of DSM in Africa is reviewed, including when DSM is used as input to a downstream application (e.g., for crop modelling). Second, we discuss the identified challenges that pertain to the application of DSM in Africa. Third, we highlight opportunities and a set of suggestions as a means to progress towards a better accounting of the specificity of DSM in the African continent.

2. Methodology

We reviewed the existing literature on DSM in Africa using two approaches: a systematic search followed by a grey literature search using a standard search engine. We describe the two approaches in the following two sections.

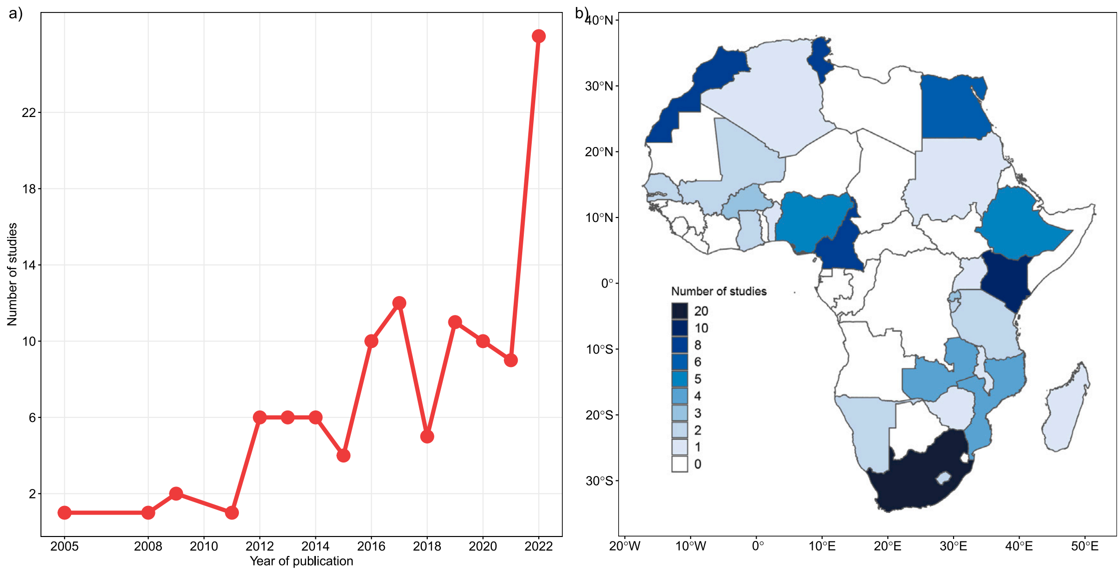

Fig. 1. (a) Annual number of DSM studies in Africa and (b) spatial distribution of studies per country. Studies at regional, continental and global scales were not considered in (b).

2.1. Systematic search

The systematic search was made in two databases: Web of Science (Core collection) and Scopus.

- In Web of Science (WoS), the search was conducted on January 30th, 2023 using the keywords "Digital soil mapping" OR "Soil mapping" OR "Predictive soil modelling" OR "High resolution mapping" in the title, abstract or keywords, for records that contained the word "Africa" or the name of any African country. We limited our search to (i) English because French is not among the search languages of WoS and (ii) to an upper time interval of 31st December 2022. The search led to 184 articles, which we narrowed down to 90 articles after manually screening through titles, keywords and abstracts and discarding records that had no direct link with DSM or soil, or whose study case was outside Africa. Only studies that conducted DSM at a local, national, regional (i.e., a large area covering many countries in Africa), continental, or global scale were retained.
- In Scopus, the search was conducted the same day, using the same search procedure as in WoS, but this time returning records for both English and French. We found 239 articles among which three were in French. There were also 154 records that were already included in the WoS search results. The articles focusing on conventional soil and vegetation mapping were discarded. After removing the duplicates between the two search engines, we followed the same refinement procedure as for the records found in WoS. This yielded 26 articles on either soil mapping or DSM in Africa. We kept articles focusing on soil mapping because, from reading the abstract, it was difficult to identify whether the work was performed using conventional or digital soil mapping.

2.2. Grey literature search

Numerous DSM projects conducted in Africa may not be published in indexed scientific journals. We therefore conducted a grey literature search to complement the systematic search. The Advanced Search of the Google Engine was used on January 31st, 2023. We limited the search to African countries and to results that had the keywords mentioned earlier in the title of web pages or documents only. Records that were already included in the database resulting from the search in WoS or Scopus were disregarded. An additional search in French using the term "Cartographie numérique des sols en Afrique" was made. The search in English yielded 36 new records from which we found 20 documents (i.e., PhD theses, scientific articles, conference presentation or posters) and 16 web links summarising DSM projects in Africa. We found 10 web links in French on soil mapping for the search with the keywords in French.

3. Current status of DSM in africa

The systematic and grey literature search yielded a total of 110 published articles, of which 83 were from WoS or Scopus, and 27 from the grey literature search. In each article, we extracted the country and the location of the study area, the purpose of the study, the spatial extent of the study area, the sampling design, the sample size from which we calculated the sampling density, as well as the modelling and validation approaches. The Pearson's correlation coefficient, r, was computed where appropriate, to investigate variables relationship. A summary of the information extracted from the articles can be found in Table 1.

3.1. Number and spatial distribution of studies

Fig. 1 shows the annual number of DSM studies between 2005 and 2022 (Fig. 1a) and the spatial distribution of studies per country (Fig. 1b) within Africa. The annual number of published DSM studies consistently increased until 2022 (see Fig. 1a). Nonetheless, the number of DSM studies varied greatly per country. South Africa was the most active country with 20 studies, followed by Kenya with 10 studies, and by Cameroon, Morocco and Tunisia with 8 studies each (Fig. 1b). Moreover, the results revealed that DSM studies have been carried out on approximately 49% of African countries (i.e., 26 countries out of 53). All the reviewed studies are summarised in Table 1.

Further classifying the number of publications by region (i.e., Central Africa, East Africa, North Africa, Southern Africa and West Africa), we found that each region had at least one DSM study. Southern Africa had the largest number of studies with a total of 35 documented research studies spanning 8 countries (i.e., Madagascar, Malawi, Mozambique, Lesotho, Namibia, South Africa, Zambia and Zimbabwe), while Central Africa, which consists of 6 countries, had the smallest number of studies, i.e., 8 documented research studies, all conducted in Cameroon. Nearly 60% of the countries in Western Africa and 85% in Central Africa had no DSM study as recorded in our literature search.

Table 1
List of the 110 records obtained from the literature search with additional information obtained from the full-text read.

Country[1]	Spatial extent[2]	Sampling design	Soil properties[3]	Data density (unit/km^2)	Number of co-variates	Predictive models[4]	Grid spacing (m)	Map quality indices[5]	Validation[6]	Purpose	Reference
Africa	Continent	Stratified random sampling	SOC, pH_{water}, SB, sand, RDR	0.00035	7	RF	250	R2, RMSE, OA	k-fold CV	Map production	Vågen et al. (2016)
Africa	Continent	Legacy soil data	SOC, pH_{water}, sand, silt, clay, BD, CEC, total N, exch. acidity, Al content, exch. bases (Ca, K, Mg, Na)	0.00092	Not provided	RF + RK, MLR	250	RMSE, ME	k-fold CV	Map production	Hengl et al. (2015)
Africa	Continent	Legacy soil data	pH_{water}, SOC, total N, Total SOC, clay, silt, sand, BD, coarse fragments, depth to bedrock, eCEC, extr. P, K, Ca, Mg, S, Na, Fe, Zn	0.00500	Not provided	RF, XGBoost, NN, cubist, NN (ensemble machine learning)	30	LCCC, RMSE	k-fold CV	Map production	Hengl et al. (2021b)
Algeria	Local	Legacy soil data	Soil groups	2.06438	12	BCT, RF, SVML, SVMR, NNET, MLP	30	OA, UA, PA, kappa	k-fold CV	Model comparison	Assami and Hamdi-Aissa (2019)
Benin	National	Legacy soil data	SOC stock, SOC, total N, pH_{water}, exch. K, avail. P, sum of bases, CEC, and base saturation	0.01504	48	QRF, Cubist	100	RMSE, R2, LCCC, MAE	Data splitting with repetition	Soil fertility assessment	Hounkpatin et al. (2022)
Burkina Faso	Local	Simple random	Sand, silt, clay, CEC, SOC, total N	1.72759	53	MLR, RF, SVM, SGB	5	RMSE, sMAPE	Independent data	Map production	Forkuor et al. (2017)
Burkina Faso	Local	Not provided	SOC stock	0.45161	23	RF, MLR	90	RMSE, ME	Data splitting	Carbon stock assessment	Hounkpatin et al. (2018a)
Burkina Faso	Local	Not provided	WRB Soil groups	0.45161	32	RF	Not provided	Kappa	Data splitting	Map production	Hounkpatin et al. (2018b)
Burundi	Local	Not provided	SOC, clay	13.46667	26	(1) geo-matching using land units obtained by spatial overlay of DEM-derived landforms and lithologic units, (2) LR, (3) GWR, (4) GAM, (5) BRT and (6) ANN (7) RK	20	ME, MAE, RMSE, SRMSE or NRMSE, Rp, SDp	Data splitting	Model comparison	Sindayihebura et al. (2017)
Burundi, Rwanda	National	Not provided	P, K, Ca, Mg, S, Cu, Zn, B, pH_{water}, Al+H, eCEC, SOM	0.036 & 0.038	100	RF	250	R2, RMSE	k-fold CV	Map production	Ruiperez Gonzalez et al. (2015)
Cameroon	Local	Nested hierarchical sampling	pH_{water}, clay, SOC	0.45584	12	RK supported by restricted maxi- mum likelihood (REML) parameter estimation	250	ME, RMSE, R2, PICP	LOOCV	Map production	Takoutsing et al. (2022)

(continued on next page)

Table 1 (*continued*).

Cameroon	Local	Legacy soil data	SOC stock (temporal 1985 & 2017)	0.0486 & 0.0534	18	QRF	Not provided	RMSE, ME, R2	k-fold CV	Carbon stock assessment	Nguemezi et al. (2021)
Cameroon	Local	Legacy soil data	SOC, clay	0.01967	Not provided	Random forest	250	R2, RMSE	LOOCV	Covariate performance	Silatsa et al. (2017)
Cameroon	Local	Nested hierarchical sampling	pH_{water}, clay, SOC	0.45584	46	RF, RK	250	ME, R2, RMSE	LOOCV	Model comparison	Takoutsing and Heuvelink (2022)
Cameroon	Local	Legacy soil data	pH_{water}, SOC, sand, silt, clay	0.00454; 0.005116	27	RF	250	R2, RMSE , LCCC	LOOCV	Map production	Silatsa et al. (2017)
Cameroon	National	Legacy soil data	pH_{water}, EC, ESP	0.00197; 0.00061; 0.00228	Not provided	ML approach	Not provided	Not provided	Not provided	Land condition assessment	Kome et al. (2021)
Cameroon	National	Legacy soil data	SOC stock	0.00301	12	RF, GBR, RF+OK, RF+IDW, GBR+OK, GBR+IDW	100	ME, MAE, R2, RMSE , coefficient of efficiency	k-fold CV	Carbon stock assessment	Silatsa et al. (2020)
Cameroon	Local	Nested hierarchical sampling	SOC, N, clay	1.60000	0	OK	Not provided	Not provided	Not provided	Map production	Takoutsing et al. (2017)
Egypt	Local	Legacy soil data	SMUs	Map surveyed at different scale	10	MNLR	28.5 m & 90 m	McFadden pseudo R-squares, OA, kappa	Not provided	Map production	Abdel-Kader (2011)
Egypt	Local	Not provided	SMUs, pH_{water}, Carbonate, effective soil depth, gypsum, and soil salinity	Not provided	1	OK	250	Not provided	Not provided	Land capability assessment	Ismail and Yacoub (2012)
Egypt	Local	Stratified random sampling	EC, clay and SOM	0.31053	Not provided	MARS, PLSR	15	R2, RMSE, SRMSE	Data splitting & LOOCV	Covariate performance	Nawar et al. (2015)
Egypt	Local	Grid sampling	Soil electrical resistivity, pH_{water}, EC, BD	625.00000	0	OK	Not provided	Not provided	Not provided	Land condition assessment	Swileam et al. (2019)
Egypt	Local	Not provided	SMUs	Not provided	18	MNLR	28.5 & 90	McFadden pseudo R-squares; OA, kappa, UA, errors of omission and commission, an overall error measure	Not provided	Map production	Abdel-Kader (2013)
Egypt	Local	Not provided	Soil capability index	0.01194	Not provided	RVFL, SCA, AFO	Not provided	accuracy, sensitivity, specificity, and precision	Not provided	Land capability assessment	Alnaimy et al. (2022)
Ethiopia	Local	Not provided	soil groups, drainage class, pH_{water}, extr. Zn, avail. P.	0.09676	139	RF	250	OA, ME, RMSE, SRMSE	Not provided	Land condition assessment	Leenaars et al. (2020)
Ethiopia	Local	Grid sampling	Sand, silt, clay, pH_{water}, exch. acidity, exch. K, exch. Ca, exch. Mg, Na, CEC, SOC, total N, Avail. P, texture class	9.79912	0	OK	100	Not provided	Not provided	Land condition assessment	Iticha and Takele (2019)

(*continued on next page*)

Table 1 (*continued*).

Ethiopia	Local	Grid sampling	pH_{water}, SOC, total N, avail. P, exch. Mg, exch. K, exch. Ca, CEC, exch. acidity	0.17726	0	OK	25	Not provided	Not provided	Land condition assessment	Sori et al. (2021)
Ethiopia	Local	Grid sampling	SMUs	13.87868	Not provided	OK	100	Not provided	Not provided	Land condition assessment	Iticha et al. (2022)
Ethiopia	National	Legacy soil data	WRB soil groups	0.01320	27	RF	250	CM	Data splitting	Map production	Ali et al. (2022)
Ghana	Local	Unbiased hybrid stratification algorithm or Smart sampling and cLHS	WC	0.00111	19	RF, XGB (ensemble modelling)	100	RMSE, R2, MAE, LCCC	k-fold CV	Land condition assessment	Nketia et al. (2022)
Ghana	National	Legacy soil data	SOC stock	0.00312	10	RK	1000	ME , MAE, R2, RMSE	LOOCV	Carbon stock assessment	Owusu et al. (2020)
Kenya	Local	cLHS	Mean weight diameter of soil aggregates.	1.76471	15	Cubist + RK	30	R2, RMSE, MSE, ME,	Data splitting with repetition	Land condition assessment	Kamamia et al. (2021)
Kenya	Local	Stratified random sampling	SOC stock, total N stock	0.33846	20	MLR, MLRK, GWR, GWRK	30	ME, RMSE	k-fold CV	Carbon stock assessment	Were et al. (2016)
Kenya	Local	Simple random	SOC, clay	0.01436	28	OK and Step-wise MLR	90	ME, RMSE, SRMSE	Cluster CV	Map production	Mora-Vallejo et al. (2008)
Kenya	Local	Legacy soil data	SOC, clay, silt, sand	0.04432	23	iPSM, SMLR, OK	30	RMSE, R2, LCCC, ME	Independent data	Map production	Minai et al. (2021)
Kenya	Local	Not provided	WRB soil groups, effective soil depth, WC and soil drainage class.	0.01429	16	K-means clustering and SOLIM	30	OA, Kappa, ME, RMSE & R2	Data splitting	Land condition assessment	Ngunjiri et al. (2019)
Kenya	Local	Stratified random sampling	SOC stock	0.33846	16	ANFIS-EG	30	ME, RMSE, RPD	Data splitting	Carbon stock assessment	Were et al. (2017)
Kenya	Local	Stratified random sampling	SOC stock	0.33846	16	SVR, ANN, RF	30	RMSE, ME & R2	Data splitting	Carbon stock assessment	Were et al. (2015)
Kenya	National	Legacy soil data	WRB soil groups	0.00099	Not provided	GLMM + RK	250	Kappa	Data splitting	Map production	Omuto (2013)
Kenya	Local	cLHS	soil erodibility factor	1.76471	17	Cubist + RK	30	R2	Data splitting	Land condition assessment	Kamamia et al. (2022)
Kenya	Local	Grid sampling	total SOC, P, total N, exch. K and exch. Na, Ca, Mg, Mn, Zn, Cu, Fe	340.71550	0	OK	Not provided	RMSE	LOOCV	Land condition assessment	Mwendwa et al. (2022)
Lesotho	National	Legacy soil data	SOC, BD, coarse fragments, SOC stock	0.00530	28	GLM, RF, SVM, BGLM, BCART, CART, Ranger, QRF, QNR	250	ME, RMSE, R2, NSE	Data splitting	Carbon stock assessment	Ramakhanna et al. (2022)
Madagascar	National	Legacy soil data	SOC stock	0.00203	10	RF	30	RMSE	Data splitting	Carbon stock assessment	Ramifehiarivo et al. (2017)
Malawi	Local	Simple random	SOC, total N, Avail. P, exch. K, C:N ratio, C:P ratio, soil structural stability	4.46939	37	RF	10	RMSE, R2	k-fold CV	Land condition assessment	Mponela et al. (2020)

(*continued on next page*)

Table 1 (*continued*).

Mali	Local	Legacy soil data	Clay, sand, silt, pH_{water}, total N and SOC	Not provided	19	Cubist	1000	ME, RMSE, r, R2	Independent data	Model extrapolation	Nenkam et al. (2022)
Mali	Local	Stratified random sampling	clay, sand, silt, pH_{water}, SOC, total N, P, K	4.49438	0	OK	Not provided	not provided	Not provided	Land condition assessment	Dembele et al. (2016)
Morocco	Local	Systematic sampling	SOM	0.72936	7	MLR, ANN	15	RMSE, MAE, R2	Data splitting	Model comparison	Bouasria et al. (2022)
Morocco	Local	Not provided	SOM	0.84404	Not provided	DT, k-NN, ANN	15	MAE, RMSE, R2	Data splitting	Model comparison	Bouasria et al. (2020)
Morocco	Local	Stratified random sampling	Soil aggregate stability (mean weight diameter, slow wetting, fast wetting, and their means)	0.07705	15	RF	15	R2, RMSE, MAE	Data splitting	Land condition assessment	Bouslihim et al. (2021)
Morocco	Local	Systematic sampling	SOM	0.19940	19	OK, RK, MLR, RF, QRF, GPR and an ensemble model.	30	LCCC, MSE, RMSE, NRMSE, RPIQ, MSEc, RMSEc	Data splitting	Sample ratio performance	John et al. (2022c)
Morocco	Local	Grid sampling	P, K	0.0191; 0.04	20	RF	30	R2, RMSE, MAE	Data splitting	Sampling design comparison	John et al. (2022a)
Morocco	Local	Grid sampling	Soil salinity	4.45305	0	OK	100	RMSE, ME, LCCC	LOOCV	Land condition assessment	Dakak et al. (2017)
Morocco	Local	Grid sampling	pH_{water}, SOM, K, P	0.22797	17	OK, KED	30	RMSE	k-fold CV	Covariate performance	John et al. (2022b)
Morocco	Local	Stratified random sampling	Soil salinity	0.05938	19	linear, logarithmic, and polynomial degree two and four regression models	30	R2, RMSE, NSE, Pbias	Not provided	Land condition assessment	Rafik et al. (2022)
Mozambique	Local	Stratified random sampling	SOC	4.20000	9	OK, KED, linear regression.	1000	RMSE, SRMSE, LOOCV	Independent data	Model extrapolation	Cambule et al. (2013)
Mozambique	Local	cLHS	Land suitability maps	2.15; 1.72	7	SoLIM	30	OA	Data splitting	Land capability assessment	Van Zijl et al. (2014)
Mozambique	Local	Not provided	WRB Soil groups	1.10000	2	SoLIM	Not provided	OA	Independent data	Map production	Van Zijl et al. (2012)
Mozambique	Local	Stratified random sampling	SOC stock	0.00576	5	OK and UK	Not provided	RMSE	Data splitting & LOOCV	Carbon stock assessment	Cambule et al. (2014)
Namibia	Local	Directed stratified sampling design	SOC stock	0.00204	222	OK, RK (only used SoilGrids SOC stock as covariates), and RF + OK (222 cov)	250	ME, MAE, R2	k-fold CV	Carbon stock assessment	Nijbroek et al. (2018)
Namibia	National	Legacy soil data	Clay	0.00044	109	RK, FRK (filtered RK), PPR (projection pursuit regression), FPPR (filtered PPR), RF, FRF(filtered RF).	250	ME, RMSE, LCCC	k-fold CV	Measurement error impact	van der Westhuizen et al. (2022)

(continued on next page)

Table 1 (*continued*).

Nigeria	National	Not provided	PAWC, sand, silt, clay, BD, SOC, pH_{water}, CEC	0.00121	35	Cubist + RK	100	ME, relative error	k-fold CV	Land condition assessment	Ugbaje and Reuter (2013)
Nigeria	National	Legacy soil data	SOC, BD, SOC stock	0.0007697, 0.00024032	23	RF, cubist, BRT	1000	RMSE, R2 and LCCC	Data spliting with repetition	Carbon stock assessment	Akpa et al. (2016a)
Nigeria	National	Legacy soil data	SOC, BD, total N	0.000994 & 0.000272 & 0.001178	0	OK	Not provided	Only described the spatial variation of the maps	Not provided	Soil health assessment	Boluwade (2019)
Nigeria	National	Legacy soil data	Clay, silt, sand	0.00106	23	RF	1000	R2, ME , RMSE, LCCC	Data splitting	Map production	Akpa et al. (2014)
Nigeria	National	Legacy soil data	BD, eCEC	0.00028; 0.000679	18	RF	1000	MAE, RMSE, R2, LCCC	Data splitting	Map production	Akpa et al. (2016b)
Rwanda	National	Not provided	total N, P, K, soil nutrient balance	0.01728	15	Ensemble modelling from 4 ML algorthims (RF, GBM, xgbDART, SVMRadial)	250	R2, MAEm, RMSE	k-fold CV	Land condition assessment	Uwiragiye et al. (2022)
Rwanda	National	Not provided	SOC, pH_{water}	0.03462	1	RK	250	R2, MAE	Data splitting	Map production	Piikki et al. (2017)
Senegal	Local	Cluster random sampling	SOC	0.08564	Not provided	Classification Tree	30	Mean prediction error and MSPE, RMSE, R2	Independent data	Map production	Stoorvogel et al. (2009)
Senegal	Local	Cluster random sampling	SOC, silt, sand, clay, pH_{water}, BD	0.04902	Not provided	Linear regression	30	ME, MSE, RMSE	Not provided	Map production	Kempen (2005)
South Africa	Local	Smart sampling and cLHS	hydropedological soil mapping units	5.76387	Not provided	SoLIM	30	OA	Independent data	Map production	Van Zijl and Le Roux (2014)
South Africa	Local	Expert samples and cLHS	Sand, silt, clay, SOC, gravel	51.85681	31	simple linear models (Ridge regression, linear boosted models, quantile regression) and non-linear models (SVM, RF, SGB, cubist, penalised additive spines)	10	RMSE, R2	LOOCV	Map production	Flynn et al. (2019a)
South Africa	Local	Transects	hydropedological soil mapping units	Not provided	14	MNLR	30	Total evaluation point accuracy, UA, PA, kappa	Data splitting	Hydrology modelling	Van Zijl et al. (2020)
South Africa	Local	cLHS	hydropedological soil mapping units	1.18333	Not provided	MNLR	30	OA, kappa	Data splitting	Hydrology modelling	van Zijl et al. (2019)
South Africa	Local	cLHS	hydropedological soil mapping units	0.21739	4	SoLIM	30	OA, kappa	Not provided	Hydrology modelling	Smit and van Tol (2022)
South Africa	Local	cLHS	WRB Soil groups	0.72289	23 & 24	MNLR for soil classes, Cubist for soil depth	30	Total evaluation point accuracy, UA, PA, kappa, R2 for soil depth	Data splitting	Land condition assessment	Du Plessis et al. (2020)

(*continued on next page*)

Table 1 (*continued*).

South Africa	Local	Simple random	BD, WC	1378.26087	Not provided	Inverse distance weighting with 3/12 neighbouring points (IDW3, IDW12), regular spline with tension and OK	Not provided	ME, MAE	Data splitting	Land condition assessment	Dlamini and Chaplot (2012)
South Africa	Local	Expert samples and cLHS	Soil depth class	25.40984	14	MNLR applied in DSMART, RF, OLR	10	Kappa, OA, PA, UA	Independent data	Land condition assessment	Flynn et al. (2019b)
South Africa	Local	Legacy soil data	TMUs	1.57729	52	RF, DSMART	30	CM, OA, RI	Independent data	Map production	Flynn et al. (2020)
South Africa	National	Legacy soil data	SOC stock	0.00478	40	RF	30	R2, RMSE, MAE	Data spliting with repetition	Carbon stock assessment	Venter et al. (2021)
South Africa	Local	Not provided	Soil forms	Not provided	7	SoLIM	20	Not provided	Not provided	Soil database comparison	Van Zijl and Botha (2016)
South Africa	Local	Nested hierarchical sampling	Soil associations	1.72;1.21;0.69	Not provided	SoLIM	30	OA	Independent data	Map production	Van Zijl et al. (2013)
South Africa	Local	cLHS	Erosion sensitivity index	1.17298	Not provided	MLR	Not provided	ANOVA	Data splitting	Land condition assessment	Parwada and van Tol (2020)
South Africa	Local	Stratified random sampling	Soil associations	6.1 & 11.8	27	k-nearest neighbour, nearest shrunken centroid, discriminatory analysis, MNLR, linear and radial SVM, decision trees, SGBoost, RF, NN	30	Kappa & CM	Not provided	Model comparison	Flynn et al. (2019c)
South Africa	Local	Simple random	SOC stock	0.12667	3	UK	100	MSE, RMSE	Not provided	Carbon stock assessment	Wiese et al. (2016)
South Africa	Local	Simple random sampling	SOC stock	3.32000	44	RF + RK	10	R2, RMSE, SRMSE, ME	Data splitting	Carbon stock assessment	Wiese (2019)
South Africa	Local	Expert samples and cLHS	Soil associations, Texture class, soil depth class, bleached topsoil	25.40984	85	LDA, RR, SVM, NB, LogitBoost, SGB, RF, CS	10	Kappa, PA, UA	LOOCV	Model comparison	Flynn et al. (2022a)
South Africa	Local	Not provided	Land type	Not provided	17	DSMART	Not provided	R2, RMSE, PBIAS, NSE, KGE	Not provided	Hydrology modelling	Van Tol and Van Zijl (2022)
South Africa	Local	Stratified random sampling	Soc stock	0.97792	52	Cubist	30	RMSE	Data splitting	Carbon stock assessment	Flynn et al. (2022b)
South Africa	Local	cLHS + Smart sampling	hydropedological soil mapping units	2.82429	Not provided	SoLIM	Not provided	CM	Independent data	Hydrology modelling	van Zijl et al. (2016)
Sub-Saharan Africa	Regional	Legacy soil data	Root depth and root-zone PAWC	0.00056	Not provided	Linear regression & rules	1000	R2, ME, MAE, RMSE, RMdSE.	Not provided	Land condition assessment	Leenaars et al. (2018)

(*continued on next page*)

Table 1 (*continued*).

Sub-Saharan Africa	Regional	Legacy soil data	pH_{water}, SOC, clay	0.00070	31	linear regression without interaction, linear regression with interaction, regression tree, RF, ANN	250	RMSE, R2	Data splitting	Map production	Zhang (2013)
Sudan, Lesotho	Local	Legacy soil data	EC, ESP, pH_{water}	Not provided	Not provided	linear regression, RF, SVM, regression trees, QRF, RK, cubist	Not provided	RMSE, R2, NSE, OA	Not provided	Land condition assessment	Omuto et al. (2022)
Tanzania	Local	Legacy soil data	Soil taxonomy subgroups	Not provided	18	RF and J48	30	OA	Not provided	Map production	Massawe et al. (2018)
Tanzania	Local	Stratified random sampling	SMUs	0.00518	18	J48 and RF	Not provided	*p*-value (t.test)	Independent data	Map production	Massawe et al. (2016)
Tunisia	National	Legacy soil data	SOC stock	0.00941	36	QRF	100	R2, RMSE, ME	k-fold CV	Carbon stock assessment	Bahri et al. (2022)
Tunisia	Local	Not provided	Clay, sand, Fe and CEC	0.38053	280 spectral bands	Co-kriging	30	R2, RMSE	LOOCV	Map production	Ciampalini et al. (2012b)
Tunisia	Local	Simple random sampling	clay, sand, silt, calcium carbonate, free iron, CEC, SOC, pH_{water}	0.43000	Not provided	PLSR	5	RMSE, R2, RPD	Not provided	Map production	Gomez et al. (2012)
Tunisia	Local	Legacy soil data	Clay, silt, sand, SOC, pH,CEC	0.03154	20	RK, OK	30	95% confidence interval & proportion of true values	LOOCV	Map production	Ciampalini et al. (2012a)
Tunisia	Local	Not provided	Clay	0.60930	31	MLP-BP	30	R2, MAE, RPD or NRMSE, RPIQ	Data splitting	Covariate performance	Gasmi et al. (2022b)
Tunisia	Local	Random sampling from a hyperspectral image	Clay	Not provided	16	QRF	5	MSE, MSE		Sample ratio performance	Lagacherie et al. (2020)
Tunisia	Local	Random sampling from a hyperspectral image	Clay	Not provided	Not provided	QRF	5	ME, MSE, SSMSE	Independent data	Uncertainty quantification	Lagacherie et al. (2019)
Tunisia	Local	Not provided	Clay	0.08733	6	MLR	30	R2, RMSE, RPD, RPIQ	Data splitting	Covariate performance	Gasmi et al. (2021)
Uganda	Local	Stratified random sampling	Landscape classes	0.08717	47	BDT, RF	20	OA, Kappa	Data splitting	Map production	Hansen et al. (2009)
World	Global	Legacy soil data	Topsoil thickness, soil depth, SOC, clay, sand	0.00001	7	Disaggregation of Complex Mapping Units: overlay the complex soil map and the elevation to create a toposequence from which the simple soil mapping units is derived.	1000	RMSE difference	Independent data	Map production	Stoorvogel et al. (2009)

(*continued on next page*)

Table 1 (*continued*).

World	Global	Legacy soil data	SOC, BD, CEC, pH_{water}, clay, silt, sand, coarse fragments, WRB soil groups, depth to bedrock	0.00029	150	MNLR, NN, RF, xgboost (ensemble modelling)	250	R2	k-fold CV	Map production	Hengl et al. (2017)
Zambia	Local	Stratified random sampling	SOC	0.14669	Not provided	OK	Not provided	RMSE , ASE, NRMSE	LOOCV	Map production	Chabala et al. (2017)
Zambia	Local	Stratified random sampling	pH_{water}	0.18770	Not provided	OK	Not provided	RMSE , ASE, NRMSE, ME	LOOCV	Map production	Chabala et al. (2014)
Zambia	National	Stratified random sampling	pH_{water}	0.00228	11	Linear mixed model, RF + OK	1000	ME, MSE, median square error	LOOCV	Model comparison	Makungwe et al. (2021)
Zambia	National	Simple random sample	pH, SOC, P	0.00228	Not provided	OK, IDW	Not provided	RMSE , ASE, NRMSE	LOOCV	Map production	Chapoto et al. (2016)
Zimbabwe	Local	Stratified random sampling	SOC	Not provided	Not provided	SMLR	Not provided	ME, MAE, RMSE, RmedSE	LOOCV	Map production	Van Apeldoorn et al. (2014)

[1] Refers to the study area for which a digital soil map was made.

[2] Refers to the scale of the study area. Local:within the country; national:country scale; regional: covering multiple countries; continental and global.

[3] SB: sum of exchangeable bases; pH_{water}: pH using water; Al: aluminium; Avail. P: available phosphorus; BD: bulk density; EC: electrical conductivity; exch. Ca: calcium; CEC: cation exchange capacity; eCEC: efficient cation exchange capacity; Exch. K: exchangeable potassium; extr. Zn: extractable zinc; extr. P: extractable phosphorus; S: sulfure; B: Boron; Fe: iron; exch. Mg: magnesium; Mn: manganese; N: nitrogen; Na: sodium; PAWC: plant available water content; SOC: soil organic carbon; SOM: soil organic matter; ESP: exchangeable sodium percentage; RDR: root depth restriction; SMUs: soil mapping units; WC: soil moisture storage capacity; TMUs: terrain morphological units.

[4] MNLR: Multinomial logistic regression; SoLIM: Soil land inference model; PLSR: Partial Least Square regression; BRT: Boosted regression tree; RF: random forest; SVM: support vector machines; SVML: linear support vector machines; SVMR: radial-basis support vector machines; NNET: single hidden-layer neural networks; MLP: multilayer- perceptron neural network; GLM: generalised linear model; BGLM: Boosting generalised linear model; CART: classification and regression tree; BCART: classification and regression tree with bagging; Ranger: Ranger Random Forest; QRF: quantile regression forest; QNR: quantitative neural network; IDW3: Inverse distance weighting with 3 neighbouring points; IDW12: Inverse distance weighting with 12 neighbouring points; TSF: regular spline with tension; iPSM: individual predictive soil mapping; ANN: artificial neural network; GWR: geographically weighted regression; GWRK: geographically weighted regression-kriging; GAM: Generalised additive model; DSMART: disaggregating and harmonising of soil map units through resampled classification trees; REML: restricted maximum likelihood; ANFIS-EG: adaptive neuro-fuzzy inference system; MARS: multivariate adaptive regression splines; MLP-BP: Multilayer Perceptron with backpropagation learning algorithm; DT: decision trees; k-NN: k-Nearest Neighbours; OK: ordinary kriging, RK: regression kriging; GPR: Gaussian process regression; GBM: gradient boosting; XGBDART/XGBoost: extreme Gradient Boosting; GLMM: generalised linear mixed-effects model; NSC: nearest shrunken centroid; LDA: linear discriminatory analysis; RR: ridge regression; NB: Naïve Bayes classification; LogitBoost: Boosted logistics tree; Stochastic gradient boosting: SGB: Stochastic gradient boosting; C5: C5.0 Decision tree; RVFL: traditional random vector functional link; SCA: sine cosine algorithm; AFO: aptenodytes forsteri optimisation algorithm; LBM: linear boosted model; QR: quantile regression; PAS: penalised additive spines; LR: least squares linear regression.

[5] RMSE: root mean square error; R^2/MEC: coefficient of determination/model efficiency coefficient; LCCC: Lin concordance coefficient; MAE: mean absolute error; ME: mean prediction error; MSE: mean square prediction error; PBias: percent bias; RMedSE: root median square error; McR^2: McFadden pseudo R-squares; RPD: ratio of performance to deviation; Rp: the mean rank; SDp: the standard deviation of ranks; CI: confidence interval; MSEc: mean residual variance; RMSEc: root mean residual variance; RPIQ: ratio of performance to interquartile distance; SSMSE: mean square error skill score; r: correlation coefficient; MSPE: mean squared prediction error; ASE: average standard error; KGE: Klinf-Gupta efficiency; OA: overall accuracy; UA: user's accuracy; PA: producer's accuracy; RI: relative index; CM: confusion matrix.

[6] CV: cross-validation; LOOCV: leave-one-out CV.

3.2. Spatial extent

Fig. 2 presents the spatial extent over which the DSM studies in our literature review were conducted, either local (0.06–450,000 km^2), national (26,000–1.22 × 10^6 km^2), regional (1.8 × 10^7 km^2 for Sub-Saharan Africa, the only regional study), continental (3 × 10^7 km^2) and global. Most of the DSM studies (74.5%, i.e., 79 studies) were carried out at a local extent.

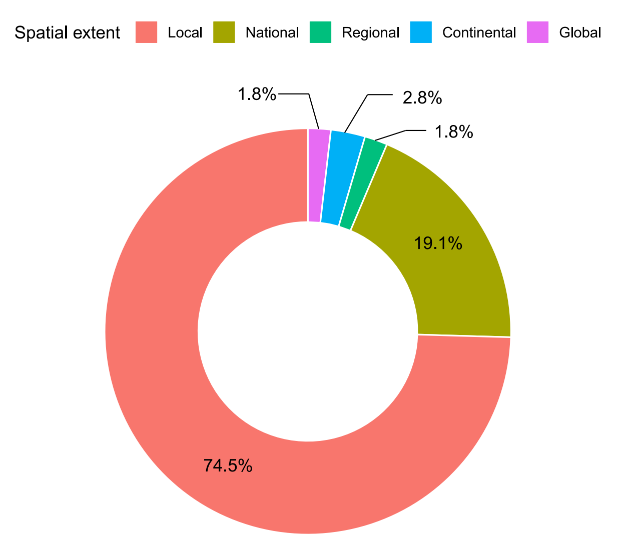

Fig. 2. Relative proportion of studies based on the spatial extent at which DSM was implemented, for local (0.06–450,000 km^2), national (26,000–1.22 × 10^6 km^2), regional (1.8 × 10^7 km^2), and continental (3 × 10^7 km^2) extents.

3.3. Soil properties, sampling designs and maximum soil depth

The quantitative and categorical soil variables mapped are shown in Fig. 3. More than two-third (72%) of the studies mapped quantitative properties, in which soil organic carbon (SOC) received the greatest attention in 31% of studies, followed by clay (27%) and pH_{water} (22%). Phosphorus (P) was the most frequently mapped nutrient (12%), seconded by total *N* (11%). In Fig. 3, "Other properties" refers to soil properties mapped by a single study. Example of such soil properties are the root erosion sensitivity index, the C:N or C:P ratios, copper and sulfur. On the other hand, fewer DSM studies (28%) mapped categorical soil variables. Soil classes were the most common mapped variable (15%), either through World Reference Base (WRB) taxonomic units, soil associations or conceptual soil mapping units. Fewer studies

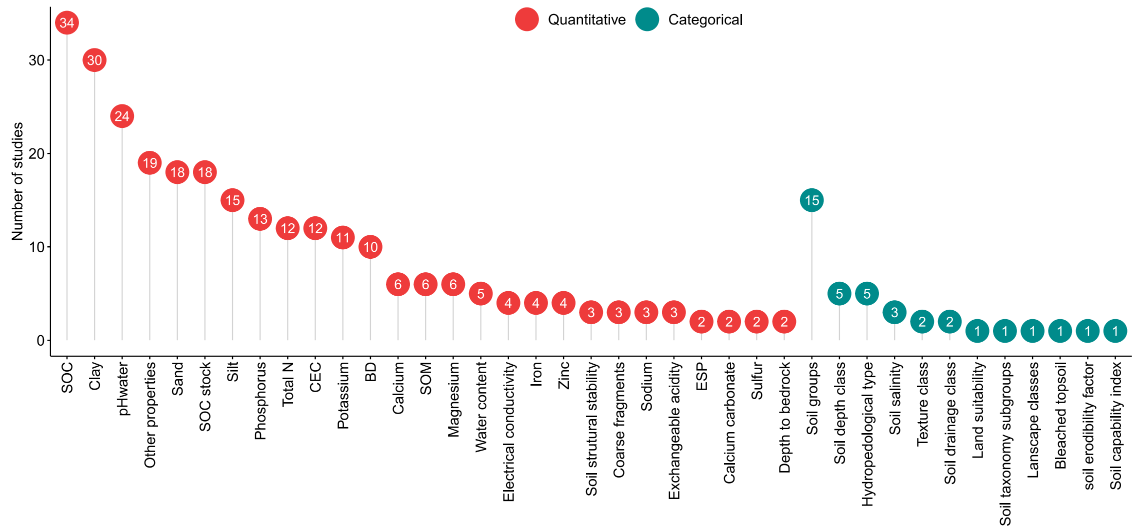

Fig. 3. Frequency of studies that mapped quantitative and categorical soil variables.

mapped soil depth classes (4.5%) and hydropedological soil types (5%). Hydropedological soil types were mapped only in South Africa.

The sampling designs used to collect the soil samples are shown in Fig. 4a. Approximately, 17% of studies did not provide information on the sampling design. Out of those who did report the design, nearly 36% of studies used probability sampling with stratified random sampling (≈16%) being the most common. Fig. 4a also shows that a larger number of studies (≈46%) used non-probability sampling strategies, with 29% of studies relying on legacy soil data. Legacy soil data was classified as non-probability sampling design because the studies did not explicitly provide information on the sampling design, suggesting that legacy data are an assemblage of several datasets. Non-probability sampling also included a smaller number of studies (≈17%) in which cLHS, nested hierarchical sampling, directed stratified and transects were used. In a nested hierarchical sampling, a site is subdivided into a given number of clusters, and within each cluster a limited number of geographical locations are selected for sampling (further details in Vågen et al., 2010), while in a directed stratified sampling design, an existing soil map is used to increase the sampling density in areas with high heterogeneity and decrease sampling density in areas which are relatively homogeneous (Nijbroek et al., 2018).

Fig. 4b shows the maximum soil depth at which the soil properties were mapped. Most of the studies mapped soil properties down to 30 cm (18%) and 20 cm (17%). Few studies (18%) focused on soil depth between 31 cm and 100 cm, and only 16% mapped soil properties for soil depth below 100 cm with the majority mapping down to 200 cm depth. Studies mapping categorical soil variables did not usually provide the soil depth information and were classified as "Not provided".

3.4. Sample size, sampling density

Fig. 5 shows scatter plots of the area of the study cases against the sample size and the sampling density. The sample size varied between 100 and 100,000 units and strongly increased ($r = 0.74$) with the area of the study case (Fig. 5a). The sampling density, conversely, strongly decreased ($r = -0.95$) for larger areas (Fig. 5b). The average sample size of studies carried out at a local extent was 300 units (Fig. 5a), which increased to an average of 1470 units for studies at a national extent and to an average of 62,290 units at a continental extent.

Fig. 5b shows that the sampling density ranges from 1378 units/km^2 for a 0.23 km^2 area reported in Dlamini and Chaplot (2012), through a density of 0.0047 units/km^2 at a national scale of 1,220,000 km^2 area in a study reported in Venter et al. (2021) to a density of

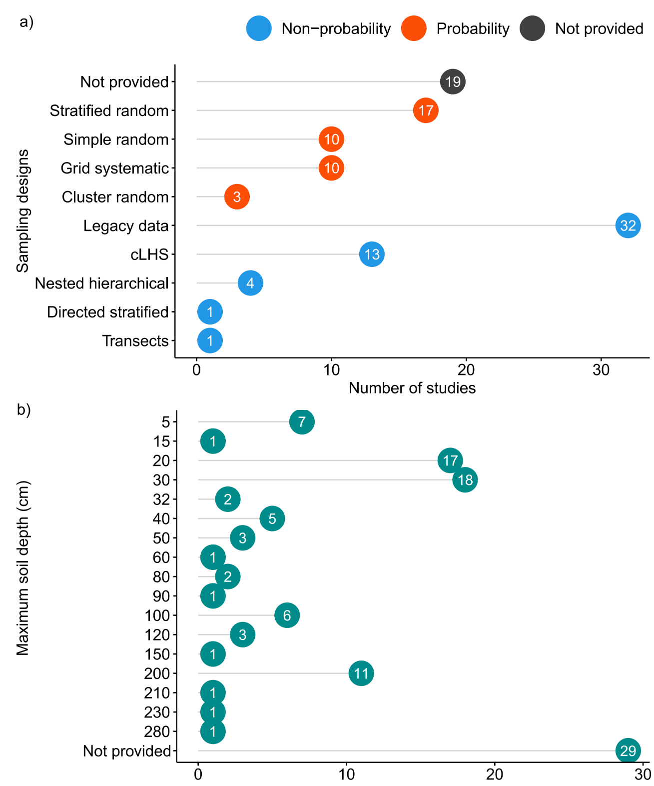

Fig. 4. Bar plots with (a) the sampling designs used in the studies classified as probability and non-probability designs, and (b) maximum soil depth at which the soil properties were mapped.

0.00092 units/km^2 for the study of Hengl et al. (2015) at a continental scale. For studies at global scales, we found a density of 0.000012 km^2. See Table 1 for more examples pertaining to different spatial extent.

3.5. Factors of soil formation and environmental covariates

Fig. 6 shows the source and number of the covariates used in DSM studies in Africa, along with their categorisation into the *scorpan* factors of soil formation and their frequency.

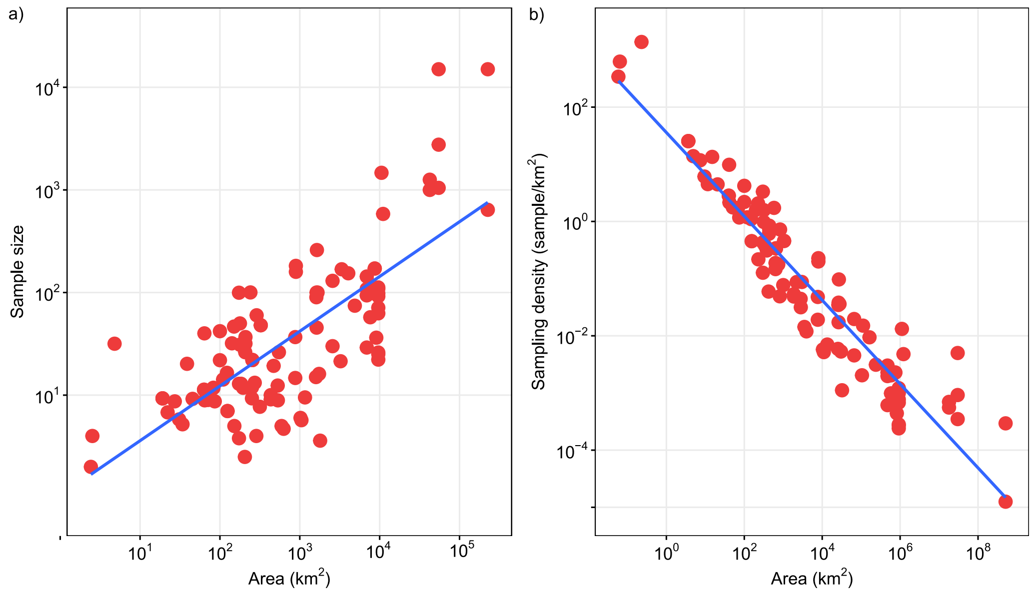

Fig. 5. Scatter plots of (a) the relation between area of the study case and sample size, and (b) area and sampling density. The axis are in a logarithmic scale. The lines are added for visualisation purposes and show a linear regression line fitted with ordinary least-squared.

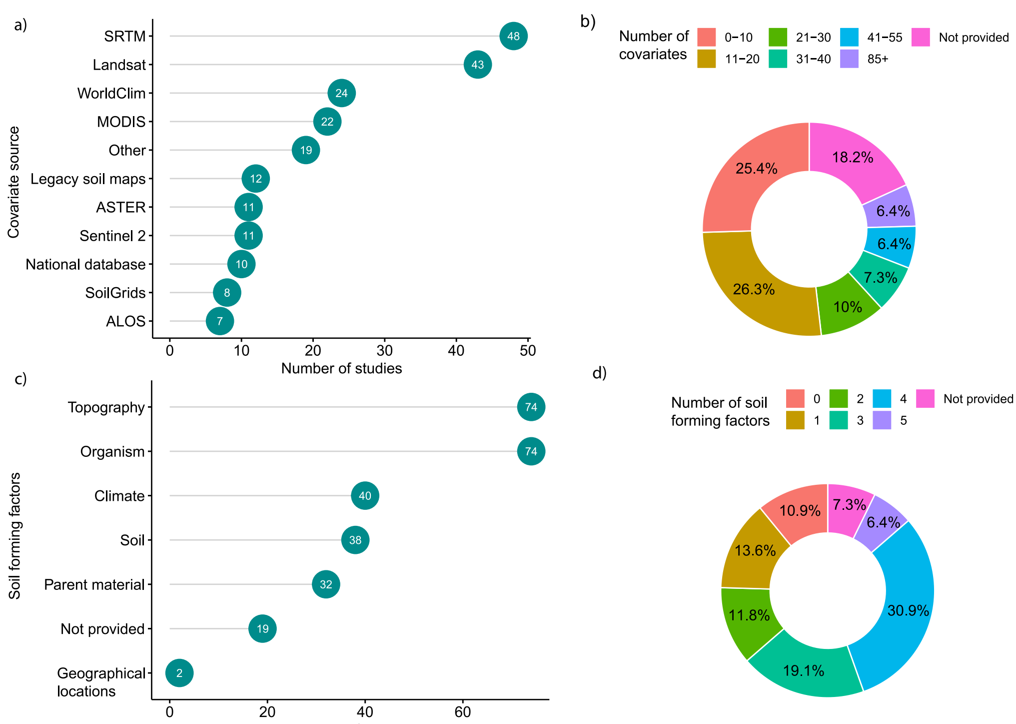

Fig. 6. Bar plots and pi-charts showing the frequency (a) of the source of the covariates, (b) of the number of covariates used, (c) of the soil forming factors to which the covariates were categorised and (d) of the number of soil forming factors. Note that, on the bar plots the frequency is given in numbers, while on the pi-charts it is given in percentages.

Nearly 80% of studies provided the source of their covariates (See Fig. 6a). The most common covariates came from remote-sensing imagery, for instance, the Shuttle Radar Topography Mission (SRTM) by 44% studies, Landsat (39%), and MODIS (20%), while fewer covariates were sourced from legacy maps (11%, either soil or agricultural maps) or from national databases (9%, e.g., national meteorological data). Covariates, such as the multi-spectral RapidEye bands or Light Detection And Ranging (LiDAR), just to mention afew, were used by less than 5% of the studies and were therefore categorised as “Other” in Fig. 6a. The number of covariates used varied between 1 to above 85 (Fig. 6b) with 26% of the studies using between 11 and 20 covariates, while 6% of studies used either between 31 and 40 or above 85 covariates. No clear relationship was found between the number of covariates and the area of the study cases.

The number of covariates were grouped into scorpan factors and Fig. 6d shows that most studies (31%) used 4 factors. The most commonly used scorpan factor was topography (67%, see Fig. 6c) with elevation, slope, topographical position index (TPI) and multi-resolution valley bottom flatness (MRVBF) being the four most used topographical covariates. The factor organism/vegetation (67%) was equally used.

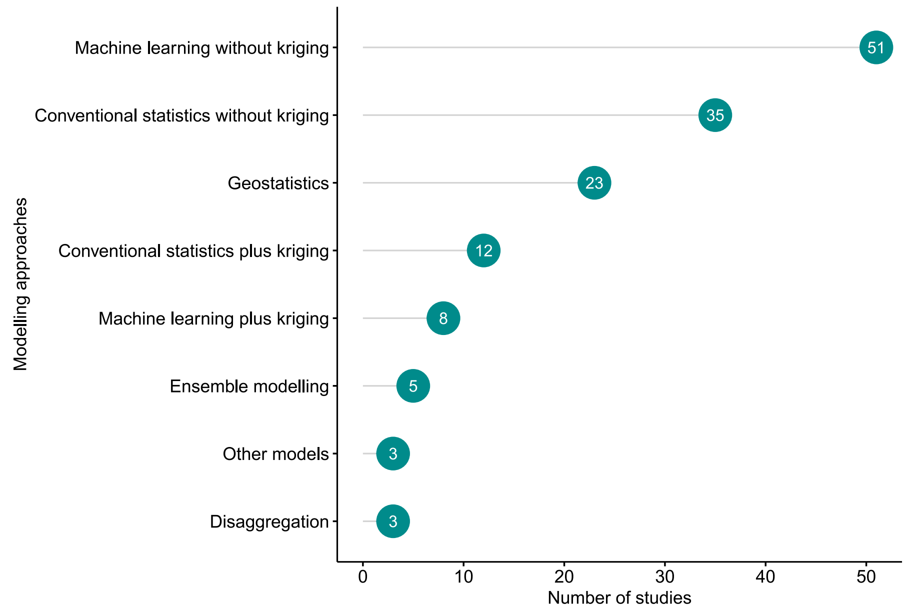

Fig. 7. The different typologies of modelling approaches used in the DSM studies covering the African continent along with their frequency of use.

Fewer studies relied on the factors climate (36.4%), soil (34.5%) and parent material (29%). Only 2% of studies used a direct estimate of the spatial location factor using, for example, the geographical coordinates as covariates. No study explicitly used the age factor which reflects the amount of rock weathering that has occurred over time. The environmental covariates of 17% of the studies could not be classified into factors of soil formation, either because none were used (e.g., studies that used OK as a predictive model), or no further information on the covariates was provided aside from the number used.

3.6. Models and model optimisation

The modelling algorithms used to predict quantitative or categorical variables were classified into 7 groups shown in Fig. 7. This classification was adapted from Chen et al. (2022).

Approximately half of the studies (46%) used machine-learning (ML) algorithms (without kriging). The most popular ML algorithms reported were RF, cubist, XGBoost, ANN, SVM, BDT, and SGB. Linear regression and landscape rules algorithms were classified as conventional statistics without kriging (31%). Examples are MLR, SoLIM, PLSR, and MNLR. Few studies (21%) used geostatistics models (e.g., OK, UK and co-kriging). Moreover, geostatistical modelling was combined with ML and conventional statistics models in 8% and 12% of the study cases, respectively. This combination is referred to as regression kriging. A smaller number of studies (4%) used ensemble modelling in which the predictions of multiple algorithms are aggregated to make the final predictions. In Fig. 7, "Other models" included deterministic interpolation such as inverse distance weighting or a combination of the latter with ML. A relatively small part of studies produced maps of soil classes using a disaggregation algorithm.

The models built were optimised through parameter tuning and covariates selection. Approximately, 16% of the studies optimised the hyper-parameters of the soil predictive models. Common optimisation approaches used were grid-search and sequential model-based algorithms. About 47% of the studies implemented a covariates selection either as a pre-processing step or embedded in the modelling step through the soil-covariates relationship. The most common pre-processing approach was the variation inflation factor, the step-wise selection using the Akaike Information Criterion on linear regression models, and the recursive feature elimination used on ML algorithms.

3.7. Validation approaches, validation statistics and uncertainty quantification

The validation statistics used to evaluate the predicted soil maps are shown in Fig. 8a. Quantitative soil maps were evaluated using statistical indices such as RMSE (58%), which was usually associated with an estimate of the deviation to the 1:1 line, such as the MEC or R^2 (42%). Approximately one-third of the studies reported bias estimates (e.g., ME). Indices, like RPIQ and r were used by less than 5% of the studies and were categorised as "Other indices". Common validation statistics for categorical variables were overall accuracy (13%) and kappa (12%), while fewer studies quantified the confusion matrix.

The validation statistics were estimated using various validation approaches (Fig. 8b). These included data splitting (34%), cross-validation (i.e., k-fold (18%), leave-one-out (12%), cluster (2%)). About 12% of the studies implemented a validation approach where a probability sample was used to assess the quality of the map, referred to as independent data validation. Approximately 21% of studies did not report the validation approach used to evaluate the quality of their predictions.

Fig. 8c shows the relationship between sampling density and RMSE for studies mapping SOC (%). This was done to illustrate the relationship between the magnitude of the mapping error and sampling density. SOC was selected as it had received the most attention in the reviewed DSM studies (see Fig. 3). Overall, the RMSE decreased as the sampling density increased. However, the correlation was $r = 0.3$, indicating a weak relationship.

Quantifying the uncertainty associated with model predictions is critical in DSM. Only 11% of the reviewed studies computed the uncertainty associated with their predicted maps, through the estimation of prediction intervals. Further, few (i.e., nearly 4%) went a step further to assess the quality of that estimation by computing the prediction interval coverage probability (PICP). PICP determines the proportion of times the prediction interval actually contains the true value of the target variable through a cross-validation procedure.

3.8. Map resolution

The resolution or grid spacing at which the digital soil maps were produced are shown in Fig. 9a. Nearly 19% of the studies provided no information on the map resolution. The most common spatial resolution

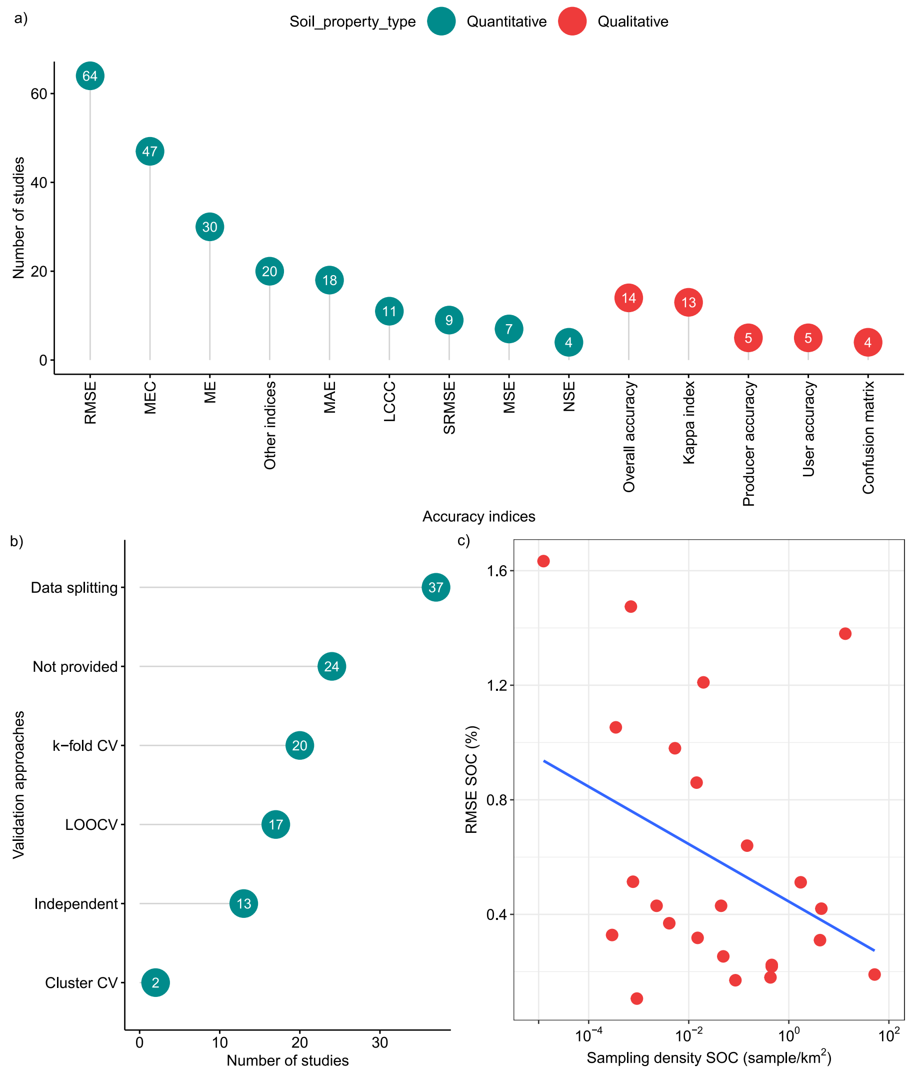

Fig. 8. Frequency (a) of validation indices found in the literature and (b) of validation approaches. The relationship between RMSE and sampling density for SOC in reported in (c).

was 30 m, and this can be associated with the increased use of SRTM and Landsat (30 m). The second most used resolution was 250 m, which also relates to the resolution of MODIS (250 m), the 2nd most used organism-related covariate. So, we can highlight that the covariates are the most significant factor that controls the final prediction map resolution. We further see a moderately strong and positive association ($r = 0.65$) between the logarithm area of the study case and logarithm resolution (see Fig. 9b).

3.9. Participating institutions

Fig. 10 shows the location of the institutions that performed DSM studies within the continent of Africa. Approximately 37% of DSM studies were conducted solely by institutions located within Africa. However, there are regional disparities in these statistics. For example, nearly 90% of DSM studies in South Africa were led by institutions located within Africa, of which 80% were conducted solely by South African institutions. Fewer studies (i.e., 10%) were conducted by institutions located outside Africa and 52% of DSM studies in Africa were a collaborative result of various institutions located both within and outside of Africa. We also evaluated the first authors and found that 76% originated from Africa.

3.10. Objectives of the DSM studies

DSM studies were carried out for a wide range of purposes (Fig. 11). The most popular objective was map production (≈35%) in which soil property (e.g., OC, pH, etc.) maps were generated to understand the spatial variation of these properties and increase the availability and access to soil information. The second most common objective was soil and land condition assessment by nearly 22% of the studies, including but not limited to soil fertility assessment and water holding capacity evaluation. Fewer studies (≈15%) generated maps to assess carbon stock and the soil's potential to sequester carbon while other studies focused on model comparison (7%), covariate performance (≈5%) and hydrology modelling (≈5%). Less than 2% of the studies generated maps to assess land capability, model extrapolation and sample ratio performance.

4. Challenges and opportunities

Based on our literature review, we identified challenges which we describe below with an outline for research opportunities.

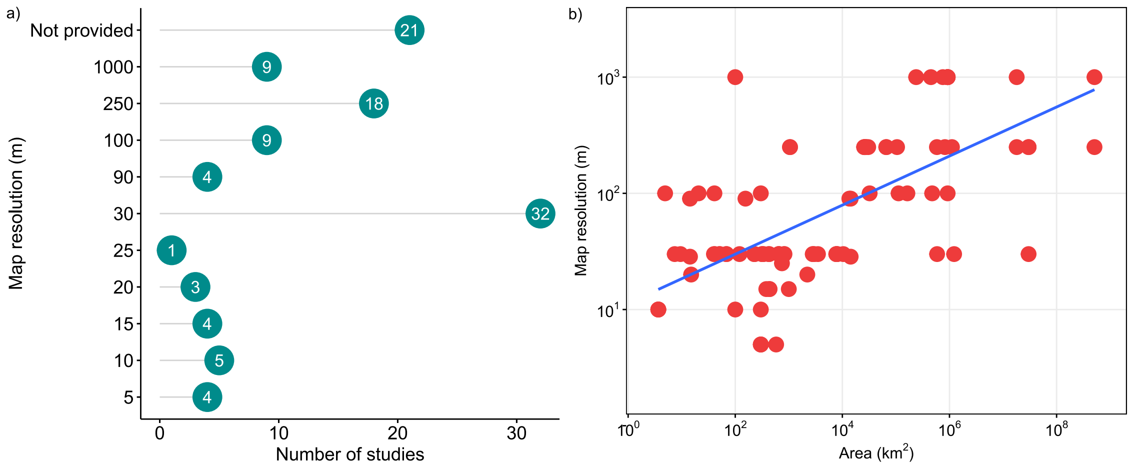

Fig. 9. Bar and scatter plots with (a) the frequency of studies based for various map resolutions and (b) the study case area plotted against the map resolution. The blue line in (b) is added for visualisation purposes and shows a linear regression line fitted with ordinary least-squares.

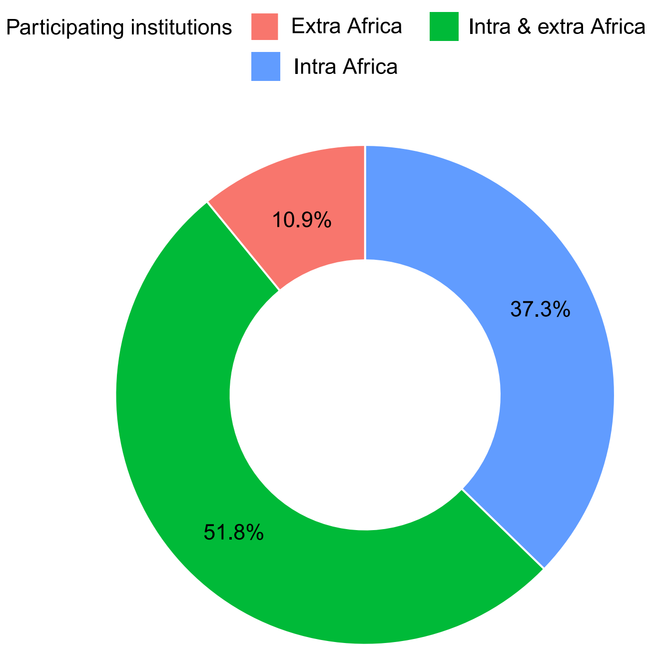

Fig. 10. Origin of the participating institutions in studies on digital soil mapping in Africa.

4.1. Why do many african countries have no intra-national or national DSM studies?

Our literature review has revealed that about 51% of (i.e., 27) African countries had no recorded DSM study. This covers approximately 17 M km^2 of land area. It implies that to obtain soil information, these countries need to source soil data from legacy digitised soil maps (e.g., the digital soil map of the world, FAO-UNESCO, 1977; Batjes, 2016; Nachtergaele et al., 2023) or global or continental-scale digital soil maps (e.g., SoilGrids250 m or iSDA, Poggio et al., 2021; Hengl et al., 2021b). The soil information from such maps, however, is usually coarse with some debate on whether it accurately reflects local soil geography (Buenemann et al., 2023). The information can also be outdated, as is the case when using legacy digitised maps to understand the pattern of dynamic properties. Four main factors may explain the lack of published DSM studies in these African countries, these are:

- The lack of data to build DSM models. DSM models are data-driven, and their predictive ability is commensurate with the quantity and spatial distribution of the available soil data (Wadoux et al., 2020). Many African countries, however, still suffer from limited access to soil data (e.g., low geographic coverage of the observations) or from poor-quality data. Fig. 12 shows as an example the spatial distribution of the observations available in the African Soil Profile database (AfSP, Leenaars et al., 2014), which is a compendium of soil observations compiled from disparate sources and covering the continent. It reveals that many countries have a low sampling density, resulting in little or no available soil data. These include Guinea, the Democratic Republic of Congo, Chad, Libya and, Liberia, among other countries. Without a sufficient spatial coverage, performing DSM is a challenging task that relies on soil model extrapolation (Nenkam et al., 2022) or disaggregation of existing coarse-resolution maps (Flynn et al., 2019b).
- The lack of capacity to perform digital soil mapping. Carrying out a DSM study usually requires digital skills (e.g., for spatial data management and programming), which national soil experts do not always have. Capacity building on DSM was identified as the primary action to improve soil data availability within the continent by the African Soil Partnership launched a decade ago under the umbrella of FAO's Global Soil Partnership (GSP) (FAO, 2015). The GSP performed numerous trainings on DSM between 2014 and 2022 (a list of the countries that received the trainings can be found in the Supplementary Material). Some of the trainings involved 30 countries among which Benin, Bostwana, Gabon, Ghana, and Tanzania to mention a few. ISRIC and the University of Sydney also performed trainings of African soil experts in DSM. We found a correspondence between the countries involved, the date of such trainings and the generation of DSM studies; for instance, Cameroon, Morocco, Nigeria, Ethiopia, South Africa and Zambia performed DSM studies shortly after the trainings. Angola, conversely, to our knowledge did not undergo a training. This might explain why it has not yet implemented a DSM study despite having soil data on over 1000 soil profiles (Fig. 12), which could be appropriate to generate baseline digital soil maps. It appears from our non-exhaustive list of trainings conducted in Africa and linked to publications, that capacity building is a clear actionable step (see also Paterson et al., 2015). Besides, our review indicated that only approximately one-third of the DSM studies published were conducted solely by researchers whose institutions were based in Africa. This further suggests a need to organise more training courses for African experts, to reinforce the capacity of local universities, to provide training in DSM or

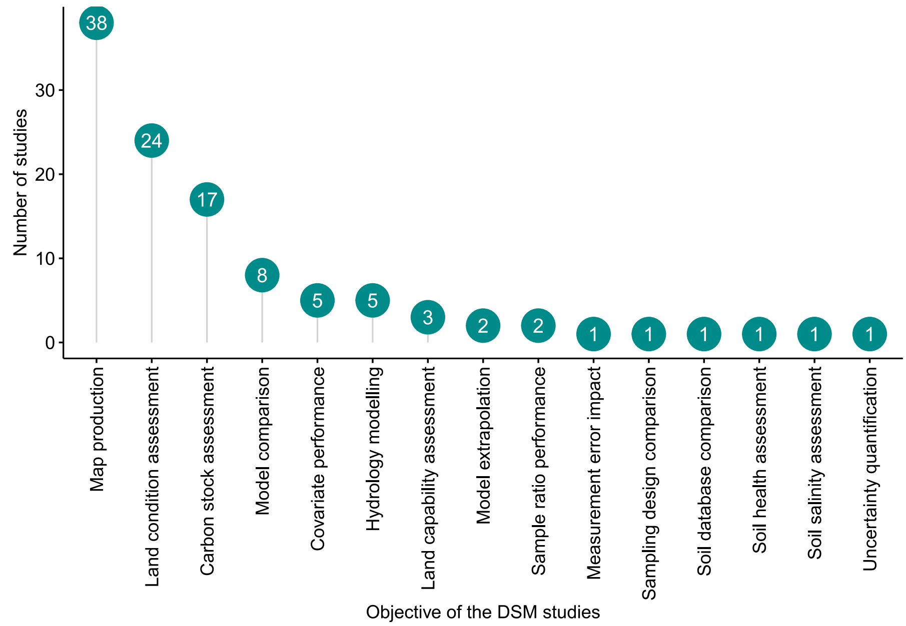

Fig. 11. Frequency of objectives of DSM studies found in this literature review.

to build DSM centres of excellence in Africa, which will in turn support capacity building within the continent.

- It may not be an immediate necessity to generate digital soil maps. Countries and national institutions, consequently, may invest little financial efforts to update existing soil information and soil research at large. We found in our review that DSM studies in several countries were financed by either private industries or through a collaboration with international institutions. In South Africa, for example, 6 out of 10 DSM studies mentioned in van Zijl et al. (2019) were solicited and financed by private industries, while in Cameroon most of the DSM studies were financed by international organisations and achieved through scientific collaborations with international institutions (e.g., Takoutsing et al., 2017; Silatsa et al., 2017; Nguemezi et al., 2021). In Morocco, the national initiative Fertimap (Moroccan Government, 2016), partly funded by a private organisation, led to the production of soil fertility maps covering agricultural areas in 2013. Another example in Morocco is the project Al-Moutmir initiative, founded by a private organisation, which is contributing new data gathering and the provision of digital soil maps (e.g., Gasmi et al., 2022a; Mamassi et al., 2022). We noted however very few national initiatives supporting soil data collection or digital soil mapping implementation in the long term. EthioSIS in Ethiopia (Hof, 2014; Gebretsadik, 2014), for example, had a continuous production of nutrient maps to support fertiliser recommendations from 2012 to 2019, while Rwanda recently launched the Rwanda Soil Information Services (RwaSIS) (Centre for Agriculture and Biosciences International, 2020), over 1000 soil profiles was added to the Ghana soil information system (GhaSIS) (Leenaars et al., 2017), and Zambia, Kenya, Nigerian and Kenya are currently planning to create SIS. South Africa has, in parallel, outlined a growing demand for soil information and identified actionable steps to satisfy this demand (Paterson et al., 2015). One such action was the requirement to survey hydropedological classes to support environmental impact assessment (van Tol, 2020). We did not find, regrettably, a mention of a priority concerning soil data collection to support DSM by national governments in the reviewed literature.
- Conventional mapping methods are preferred over DSM approaches to generate the soil information. Soil scientists in national institutions who are responsible for the generation of conventional soil maps may view DSM approaches as a threat; they may fear that DSM may jeopardise their knowledge and skills and that conventional polygon-based maps become obsolete. A similar observation was made in Arrouays et al. (2020) with examples in the Netherlands, France, Australia and the US, but this applies equally to Africa.

4.2. Covariates

Creating local covariates could be beneficial to the DSM community in Africa. Our review revealed that, there is almost no covariate specific to a country within or to the continent of Africa. One exception is South Africa which created its own high resolution (2 to 5 m) topographical data, the Stellenbosch University Digital Elevation Model (SUDEM, Van Niekerk, 2014) which was useful in numerous DSM studies (e.g., van Zijl et al., 2016; Flynn et al., 2019a). Most of the studies relied on global covariates databases, which usually ignored local factors of soil formation (e.g., agricultural management and land use often seldom measured) and consequently led to poor model performance (e.g., Nijbroek et al., 2018; Stoorvogel et al., 2009). Such local maps could be created by downscaling or statistical analysis of existing national agricultural statistic maps, as done previously in other parts of the world (e.g., Liu et al., 2020). Numerous studies reviewed have commented on the need for more localised covariates (e.g., Nijbroek et al., 2018; Kamamia et al., 2021). Therefore, more research efforts are needed on how to derive high-resolution environmental variables specific to either an individual country or the continent.

Gamma radiometry data provide much more detailed information compared to the commonly used optical remote sensing data (Reinhardt and Herrmann, 2019), however it was not used by any study in our review. Gamma radiometry data commonly comprise soil signatures on potassium (K), uranium (U) and thorium (Th) among other elements (Reinhardt and Herrmann, 2019). These signatures could be sensed from at least 15 to 30 cm down the soil profile despite the presence of vegetation cover (McBratney et al., 2003). Gamma radiometry has been used to accurately measure soil properties (e.g., Oliveira et al., 1997) and classes (e.g., Schuler et al., 2011) and therefore could have a great potential to improve model predictions. Several countries in Africa have gamma radiometry data (Eberle and Paasche, 2012; Bokar

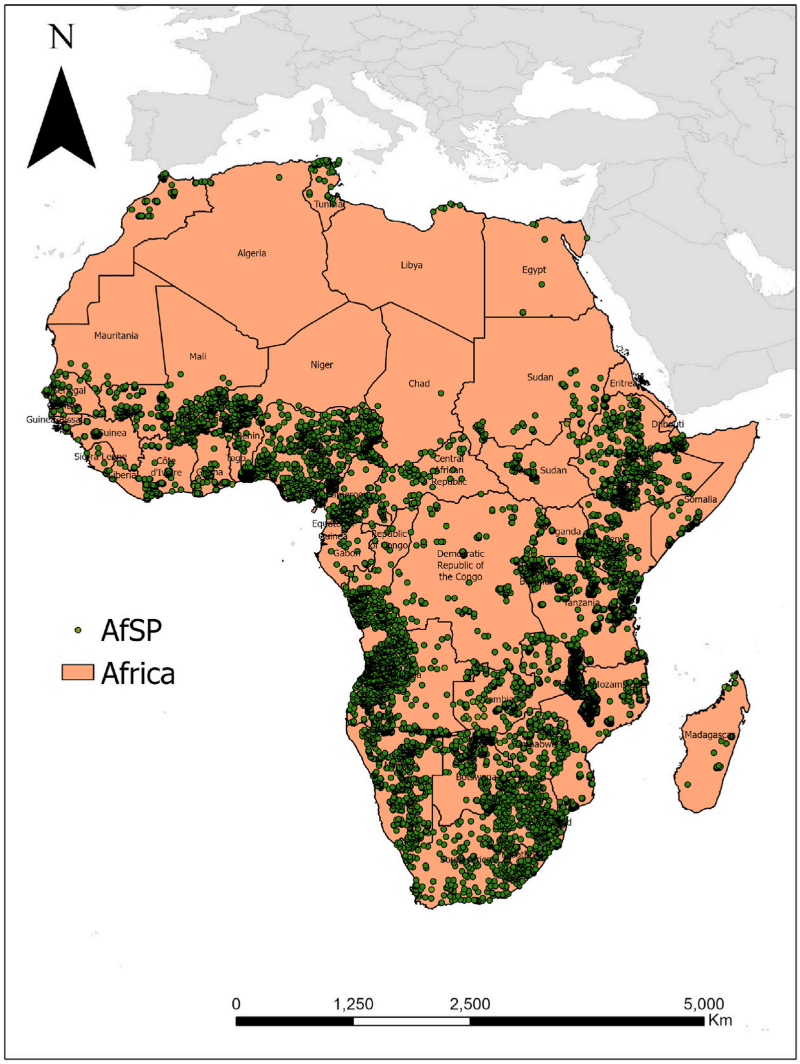

Fig. 12. Spatial distribution of the observations available in the African Soil Profile database (AfSP).

et al., 2020), but they are usually privately owned and only used in geophysics for the mining industries. In addition, they are spatially clustered across the countries (see Bokar et al. (2020) for Mali). To solve this issue, one could rely on the homosoils concept (Mallavan et al., 2010; Nenkam et al., 2023) and use quantitative extrapolation to increase the spatial coverage of the radiometry data as done by Malone et al. (2016) for New South Wales in Australia. Governments or national research institutes could collaborate with mining industries to build a database of radiometry data for their countries and Africa at large, and thus increase public access and usability. Overall, more effort is required to create covariates specific to local conditions within the continent so that various fit-for purpose maps can be created (Atkinson et al., 2017).

4.3. Target soil properties

We found that all the soil properties suggested in the GlobalSoilMap project specification (Arrouays et al., 2014) were mapped at least once within the continent. However, some relevant soil characteristics did not receive much attention as expected. These include rootable soil depth or hydraulic properties like water content at field capacity (FC) and permanent wilting point (PWP). These properties are crucial for soil capacity assessments, soil management, hydrology and crop modelling. The limited focus on mapping hydraulic properties could be attributed to the high cost and time-consuming nature of the measurement process, whether in the lab or in the field (Minasny and Hartemink, 2011). Indeed, less than 10% of the soil profiles in the AfSP database had measurements on volumetric water content at FC and PWP (Leenaars et al., 2018), which when shared among countries, cannot be enough to build DSM models for map production. As an alternative, pedotransfer functions (PTFs, McBratney et al., 2002) could be used as in Ugbaje and Reuter (2013) for Nigeria and Leenaars et al. (2018) for Sub-Saharan Africa. However, both studies used PTFs developed elsewhere, that is in areas with probably completely different environmental conditions as Nigeria or Sub-Saharan Africa, thus adding more uncertainty in the predictions and consequently limiting their use for decision making. Therefore, a cost effective way for mapping hydraulic properties could be to invest on New data collection for the creation of PTFs that are specific to African environmental conditions, whether at a national, regional or continental scale.

We identified a unique soil type prediction, namely hydropedological map units, which was mapped in South Africa only, but accounted for nearly 5% of the DSM studies reviewed. Hydropedology, perhaps better-termed pedohydrology, is an interconnected branch of soil science and hydrology that studies the movement of water through the soil (Lin, 2003). Hydropedology relies on the spatial and vertical distribution of soil characteristics, including the soil bio-chemical content to explain hydrological processes at various scales (Van Zijl et al., 2020). There has been increased interest in hydropdedology studies, across the world, for the past 20 years (Lin, 2003; D'Amore et al., 2012; Al-Maktoumi et al., 2016; Pinto et al., 2017). However, to our knowledge, mapping hydropdedology map units using a cost-effective method like DSM has only been done in South Africa. See Van Zijl and Le Roux (2014), Van Tol et al. (2015), Mamera and van Tol (2018) and Julich et al. (2022) for examples and applications. Hydropedology is critical for hydrology research and consequently for water resource management. It is useful to enhance the ability of hydrology models (Van Tol et al., 2015). It could also be useful to estimate quantitative soil properties using PTFs (Pachepsky et al., 2006), or DSM for wetland soils (Pennock et al., 2014). However, more research is needed to evaluate the ability of hydropedological maps in mapping other soil properties and nutrients, as well as soil functions such as crop production, nutrient cycling and biodiversity conservation.

4.4. Challenges with using legacy soil data

The AfSP database (AfSP Leenaars et al., 2014) was the main source of soil data in numerous DSM studies of our literature review. While using legacy soil data was deemed a necessity, being the only available comprehensive soil database (for both topsoil and subsoil) for the continent, it triggered several operational challenges regarding their use in DSM.

- *Sampling date.* Legacy datasets in Africa were particularly old. For example, in the AfSiS dataset the soil sampling dates back to 1940 and most of the data come from sampling that was carried out between 1980 and 2000 (see Leenaars et al., 2014). While this may not affect DSM for stable and slowly changing soil properties such as texture and depth, it would become critical for dynamic properties that could change rapidly following changes in land use and climate. This is the case, for example, for OC, pH, and micro-nutrients important for agro-environmental management in the African context of agricultural expansion. While most studies disregarded this challenge, DSM methodologies are also lacking to better account for the date which soils were sampled. More research on the influence of the sampling date on the prediction outcome could potentially be useful. Although some solutions have been found, one might provide weights to the soil data depending on their obsolescence to answer the problem at hand (e.g., mapping of carbon). Another obvious solution is to use recent data only (which are sometimes fragmented) and to engage in new data collection.
- *Spatial clustering.* The soil datasets from legacy surveys were usually spatially clustered (Leenaars et al., 2014; Akpa et al., 2016a; Nenkam et al., 2022). This is a common issue in many parts of the world where datasets are obtained from the gathering of surveys designed initially for various purposes (e.g., agricultural experiments). Clustering of soil datasets in the geographical space might lead to unrealistic estimates of the soil properties or classes over an area because of the over-representation of specific areas in the geographic or feature space within the model. Clustering might also provide biased estimates of the map accuracy. Declustering techniques can be applied at the modelling or validation stage, by giving more weights to samples in sparsely sampled areas. A recent example application is Nenkam et al. (2022). Alternatively, one may consider collecting data so they cover geographical or feature spaces that are unseen in the existing legacy database.
- *Positional errors.* DSM studies take as input georeferenced soil data, but the locations themselves may have a substantial error. This is especially true for samples collected before the advent of the GPS available to civilians in the 2000s. Leenaars et al. (2014) have reported that 90% of the compiled soil profiles in AfSP are georeferenced with an accuracy between 1 and 700 m. This issue is also common in global soil databases. Few studies investigated this aspect of the measurement error elsewhere. One of the few is Samsonova et al. (2018), in which they concluded that positional error had a major effect on model predictions. Additional research in this direction is required to find ways of estimating and accounting for this error in subsequent DSM studies.
- *Measurement errors.* Measurement refers to errors occurring during field sampling, sample handling and management and during the laboratory estimate (van Leeuwen et al., 2022). Using legacy data measurement error also refers to errors arising because of laboratory methods and sampling protocols that vary over time and by regional and national surveys. While some of these discrepancies are accounted for in large-scale compilation of datasets (for example, in Leenaars et al., 2014), the quantification of this error is usually beyond the researcher's control because one would need (i) information on quality control standards for each observation and (ii) replicates of the measurement in different laboratories. This might explain why no study in our literature review did account for measurement error. This is also a frequent challenge in DSM studies, with few examples in the literature only (e.g., Takoutsing et al., 2022). We need to develop methodologies to quantify this error and propagate it in subsequent modelling steps.
- *No private data.*The database does not include privately owned data that were not made publicly available. In Mali for example, the independent data used in Nenkam et al. (2022) to validate the soil maps were collected from disparate sources and not included in the AfSP database. While in Morocco, though Fig. 12 shows sparsely distributed data, there has recently been data collection under the Fertimap and Al-Moutmir initiatives but none are made publicly available. Such data may complement and reduce the spatial clustered nature of existing legacy soil database at a national level, and consequently enhance DSM related research for the provision of accurate soil information. There is need to develop win-win collaboration strategies between national governments and private institutions for the public release of soil data.

4.5. On the use of legacy soil maps

Users also recognise the value of legacy soil maps to provide information that is readily understandable by farmers and land planners. In areas of deficit in soil information and when digital soil maps are not readily available, it is usual to base decisions from the information obtained from legacy polygon-based conventional soil maps. It comes with challenges, which are described thoroughly in Van Ranst et al. (2010). There are, for example, inconsistencies between taxonomies used over time or between different localities. This poses problems for small-scale soil inventories and the transfer of information and knowledge between places, but probably not within the same country. Fig. 13 shows some of the small and large-scale maps generated for Mali during colonial times, using two or more taxonomic systems. Another challenge is the coverage at which legacy maps were generated. One example is Mali where only 50% of the country is covered by a map at the scale of 1:500,000 - ±100,000 (Van Ranst et al., 2010). Another example is Morocco, where legacy soil maps generated during the last century cover only 30% of the country (Badraoui and Stitou, 2001). In a recent study, Mukumbuta et al. (2022) attempted to access legacy soil survey information in Zambia, and found that only 22% of memoir and map sheets of previous surveys realised in colonial and post-colonial times could be traced. In contrast, several countries, including Benin, Burkina Faso, Rwanda, and Swaziland Van Ranst et al. (2010), Togo, Angola, Mozambique and Central African Republic, have completed detailed and nationally comprehensive soil mapping efforts during the same period. Most of the legacy soil maps are available in digital scanned format in the ISRIC Data Hub. In Rwanda, the legacy

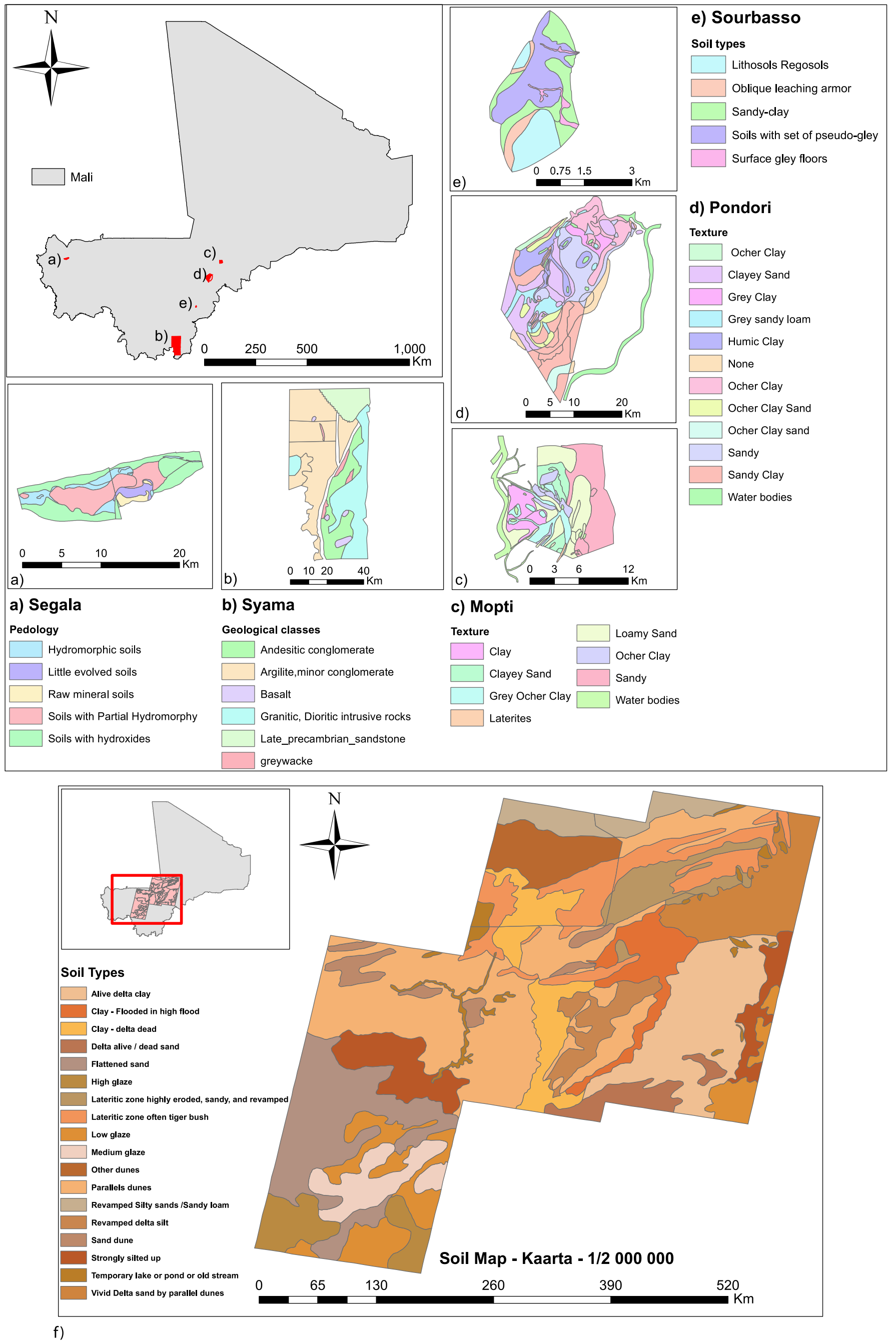

Fig. 13. Some legacy soil maps at small and large geographic parts of Mali. The maps were sourced from Panagos et al. (2011) and digitised as part of this research.

maps were used at a broader scale, for settlement planning, disaster-preparedness planning, and agricultural input application (Van Ranst et al., 2010). However, using them for decision-making at a farm scale is not recommended due to high uncertainty (Buenemann et al., 2023). Nonetheless, pursuing the data rescue of legacy soil maps, as emphasised by Arrouays et al. (2017), and making them accessible in a digital format remains critical.

Legacy soil maps in many areas served as unique and key soil information for DSM. The map may be used to design new soil surveys in DSM studies, for example using the soil map units as strata to conduct a stratified random sampling design (e.g., Massawe et al., 2016). Another use of legacy maps is by adding them as covariates in the model, as was regularly identified in our review (see Fig. 6d). Leenaars et al. (2020) concluded that integrating legacy soil maps as covariates significantly enhanced the predictive ability of the empirical models when producing a drainage soil class map in highlands in Ethiopia. In a nutshell, some countries with good legacy soil maps can benefit from this information in the creation of digital soil maps. Another approach relies on using the soil map as such to support the creation of spatially exhaustive digital soil information. This was done in Mukumbuta et al. (2021), in which it compared small-scale soil maps of Zambia against soil property measurements. Their results revealed that the maps could explain between 40% to 50% of the variation of bulk density and sand. Another approach consists in disaggregating the legacy maps using environmental covariates. In this review, disaggregation of legacy soil maps using techniques such as SoLIM (Zhu, 1997) or DSMART (Odgers et al., 2014) were intensively carried in South Africa (e.g., Van Zijl et al., 2013; Flynn et al., 2019b).

4.6. Purpose and potentials of the DSM the studies

In our literature review, nearly half of the studies focused on prediction of soil properties and classes or methodology development. This means that the other half of the studies had a purpose other than map production, such as digital soil assessment and mapping of functional properties. Some examples include Hounkpatin et al. (2022), which produced maps of several soil attributes (e.g., SOC stock, CEC, total N) and then used them to create a soil fertility map for the country of Benin in West Africa, or Leenaars et al. (2020) who predicted soil drainage class among other soil properties in several districts in Ethiopia to evaluate water infiltration and movement in the soil. Du Plessis et al. (2020) produced soil association maps to assess potential gully erosion areas and suggesting soil management strategies to prevent erosion in a catchment in Eastern Cape, South Africa. Our review suggests that many studies have made used of DSM tools to provide soil information that better matches the end-user's demands. This was recognised by many (Adhikari and Hartemink, 2016; Bouma et al., 2019; Evangelista et al., 2023) as an objective to better integrate soil science and DSM in interdisciplinary studies on soil security, ecosystem services assessment and sustainable development. This is also an opportunity to map variables that are less common in DSM studies, such as compaction and electrical conductivity, which are relevant to end users but require models that do more justice to the underlying processes and landscape evolution. Moving from DSM to soil assessment, however, requires proper quantification of the uncertainty. End users, indeed, need to be informed on the limitation of the prediction with uncertainty estimates. Our review has shown, however, that DSM studies are still lagging behind in terms of uncertainty quantification (i.e., 11% of studies reported an estimate of uncertainty). Therefore, the difficulty in measuring uncertainty and presenting it at a scale that is relevant to end-users continues to be a challenge. This could be achieved using uncertainty aggregation techniques (e.g., Courteille et al., 2024).

DSM approaches have existed for more than two decades and numerous DSM studies have looked at how to provide accurate predictions with acceptable uncertainty, as suggested by Finke (2012). Now, with the exponential increase in demand for soil information, it is time to focus more on delivering DSM products that are useful, usable and capable of appropriately informing decision-making by various stakeholders, such as precision agriculture (e.g., whether fertiliser should be applied on a field or not, or whether lime amendments is required on field with acidic soils). Note that agriculture contributes to over 30% to 60% of several African countries GDP. Information from DSM could also be used in other sectors of the national economy to make critical decisions. For example, lender and insurance institutions could use information from DSM products to preliminary access the viability and risk of a farming enterprise before committing to financing. Therefore, ready to use digital soil information would help tackle global environmental issues such as food, water and energy security, climate change mitigation, maintenance of ecosystem services, land use allocation and soil security.

Beyond the move from DSM to soil assessment, there is also room to consider DSM as input to other models whose outputs are the main interest. Integrating DSM products with crop modelling, for example, to assess crop production in the face of climate change and food security could be a powerful tool. This is in line with the broader soil science and agronomy literature (see, for example, Lagacherie et al., 2022; Guilpart et al., 2017; Claessens et al., 2015). Crop models are location-specific mechanistic models that rely on several input data, soil data being one of them, to estimate crop growth and development and thus carry out agricultural and climate change risk assessments (Keating et al., 2003). There is a promising line for future research in the integration of high-resolution DSM with crop modelling. This could solve the data paucity at a farm scale to provide spatially exhaustive estimates of the soil input parameters compared to using the measured soil observations. This would also give the opportunity to quantify the uncertainty of the crop model prediction using the quantified uncertainty of the DSM product. This approach was to date conducted at a global scale using gridded soil data at a coarse resolution (e.g., 10 km, Rosenzweig et al., 2014; Han et al., 2019). The challenge is that, these crop maps are inefficient in the African context, where the agricultural production still rely on smallholder systems with small farm size and high soil spatial variability.

4.7. Country's development

DSM products could contribute to African country's economic growth. DSM aims at providing soil related technology and data that are faster, cheaper and continuously updated (McBratney et al., 2003) to enlighten the decision-making process. Soil is a key natural capital for any nation, particularly in African countries which covers nearly 60% of the world's available arable land (Africa Union, 2015). DSM, thus, remains critical for soil's resource inventory and quantification at a spatial scale. DSM could produce detailed soil maps at various spatial scales which could positively impact the GDP of a country. Hartemink and McBratney (2008) showed that investing in soil related research may have a positive relationship with the GDP of a country. Their results suggested that the national coverage of detailed soil maps increased with increasing GDP in developed countries, while the national coverage of coarse soil maps decreased with increasing GDP for most developing countries, among which African countries. This suggests that implementing DSM studies could provide valuable and actionable soil information to various stakeholders, thereby positively impacting the development of several African countries. However, it would likely require significant investments. This is also suggested by Cook et al. (2008) who concluded that a major effort to promote DSM application will likely contribute to Africa's development. For Cook et al. (2008), DSM studies need to pass tests on *significance*, *novelty*, *actionability* and *delivery* for them to contribute to Africa's development. Our review revealed, DSM applications in Africa are yet to satisfy these milestones.

4.8. National soil information systems & policy making

Around the world, countries are putting together efforts to develop soil information systems (SISs) as a support system tool for policy. Australia, for example, had launched in 2021 a National Soil Strategy (Australian Government, 2021). The policy describes how the country manage and improve soil health in the coming 20 years. This led to the creation of the Australian National Soil Information System (ANSIS, Australian Government, 2023), an initiative to provide consistent and accessible soil data and information standards. Similar national and continental policy initiatives are taking place in the European Union and the United States.

In Africa, a few countries are developing national soil information systems (NSISs). Ethiopia has launched the Ethiopian soil information system (EthioSIS, Ethiopian Agricultural Transformation Agency, 2017) in 2017 to tackle various threats to soil, in particular soil erosion and land degradation, which significantly affected crop productivity and the country's food security. EthioSIS focused on generating digital soil fertility maps for agricultural areas at the levels of the woredas (i.e., administrative unit) using DSM techniques. From these maps, fertiliser recommendations were calculated and shared with various stakeholders of the agricultural production chain to optimise fertiliser production, supply, and usage (Wedajo Abdi, 2019; Hordofa, 2020). Regrettably, the project was discontinued in 2022 after 10 years (Ethiopian Government, 2022) and though EthioSIS's system would be enhanced with legacy soil data (See Table 1 for further information), there is to date no information on how it would be updated with new data in the upcoming future. Long term NSIS are needed for the sustainable management of soils and the mitigation of soil degradation and erosion. This importance is usually called upon by national institutions with no decision-making power (e.g., Universities in South Africa (Paterson et al., 2015)) or by international institutions (e.g., FAO's global soil partnership (Montanarella, 2015)). In most African countries, however, the general tendency has been a morose consideration of further investments for the production of NSISs using national budgets. One reason might be that several countries do not sufficiently recognise the importance of soils (Bouma et al., 2012) on the national economy and on the population welfare. In addition, in several African countries, there is not a real political vision related to soils. Concrete actions in terms of accurate soil information to understand the soil condition, at a national level, are currently either non-existent or very old.

In countries where a national database exists, the production of an up-to-date national soil information system may be constrained by (i) the quality of the data, for instance the Cameroon soil database Camsodat01 (Silatsa et al., 2017) consists mainly of legacy soil data with issues of reliability (see Section 4.4 for further considerations on this aspect) and (ii) data availability and access. The Fertimap database in Morocco, for instance, is not made publicly accessible. Another constraint is (iii) infrastructure for laboratory analysis and technical capabilities. Despite the growing need for soil information, developing a sustainable strategy to set up an NSIS in several African countries will require tackling the aforementioned constraints. An action which is urgently needed to enlighten policies on soil security, agricultural production, land degradation, water security and environmental protection, just to mention a few.

4.9. Capacity building on DSM

There is a need for a new generation of soil scientists as formulated by Hartemink and McBratney (2008). In Africa, this would enable to tackle challenges specific to the African continent. Such challenges are the inclusion of local knowledge and realities in the DSM workflow to improve the actionability and delivery of DSM products. Future capacity building needs to embrace indigenous and scientific knowledge (Snapp, 2022), which can be achieved with transdisciplinary knowledge. The transdisciplinary approach is essential to address the complexity of soils in the land system and will be crucial in the future to allow soil science to participate in studies on sustainable development (Hou et al., 2020). It is also relevant when including DSM in the soil sciences curricula of higher education institutions in Africa. One approach to support this is to allow students from other fields of research such as mathematics, data science, or computer science to integrate a major in soil science. Breaking such disciplinary barriers could significantly contribute to the advancement of DSM in Africa and boost local DSM capacity. A good balance, however, is also needed for DSM skills to coexist with knowledge in conventional soil surveys and avoid the risk stressed by Biggs et al. (2022), highlighting that DSM methods might lead to a perception that soil scientists are not needed and "modelling expert" are sufficient to make digital soil maps. Finally, we stress the importance of ongoing collaborations with institutions within and outside of Africa. Efforts should be made to enhance and improve the existing ones as highlighted by Minasny et al. (2020) who emphasised the importance of global soil science research collaboration. It could take the form of more collaborations between developed and less-developed countries. However international research should avoid helicopter-type research such as generating digital soil maps of Africa from a wealthy institution without empowering soil scientists from countries in Africa with DSM techniques. Table 1 in the Supplementary Material outlines training initiatives that various institutions have conducted for over a decade to address this challenge.

4.10. Suggestions for the way forward

DSM in Africa currently relies heavily on legacy soil data. Collection for new soil data is limited due to the high cost involved. There has been top-down initiatives such as the Africa Soil Information Service (AfSIS, Vågen et al., 2010) which collected new data at 60 sites across the continent. While bottom-up approaches are used to update the AfSP database through the GlobalSoilMap initiative (Arrouays et al., 2017), the new data are limited only to the topsoil (i.e., within the 0–30 cm depth layer), making it challenging to support land use allocation, soil fertility management and, soil health monitoring. There are also continental maps created using these new data coupled with legacy soil data (Hengl et al., 2017a, 2021b). However such maps could only provide a baseline status of the soil due to the nature of the data used (see Section 4.4). A recent initiative, namely Soils4Africa (European Commission, 2020), aims at collecting soil data on approximately 20,000 profiles distributed across the continent. However, the data would be collected only down to 50 cm soil depth. Such data are appropriate to provide information on the surface condition of the soils, but knowing the condition of deeper soils is equally essential for long-term assessment of soil health, crop productivity, and land suitability. The limitation to sampling only topsoil is often due to budget constraints, as deeper soil sampling requires more resources in terms of labour, equipment, and time, making it more expensive. However, subsoil data remains critical (de Oliveira and Bell, 2022) and if carefully planned by individual countries, it is possible to ensure the collection of soil data at deeper layers over a longer period (for example, within 20 years), consequently providing sustainable and locally relevant soil information. It is important to emphasise that, both AfSIS and Soils4Africa are useful initiatives funded and executed by international institutions outside of Africa, which means they only sample based on the project's financial capacity to achieve their objectives. Therefore, to sustainably create accurate soil information that is reliable, useful, and usable by various stakeholders within the continent, we suggest that initiatives for soil data collection and the creation of digital soil maps originate from within the continent and thus compliment external efforts while ensuring that projects are tailored to local needs, national needs and finally continental needs. An example of how this could occur is from South Africa, where hydropedological assessments are now required as part of the environmental impact assessment process for water security (van Tol, 2020), and this generates new soil data which could be incorporated into the creation of a national digital soil map.

We thus suggest the following to enhance the production of digital soil maps and move a step further with digital soil assessment.

- Promote the importance of soils. This consists of generating value-added propositions for the soil information to political institutions to emphasise its positive effects such as food security, reversing of soil degradation, environmental sustainability and consequently political stability and national economic development.
- Create long-term soil related policies. This entails national services and stakeholders to engage in lobbying activities to drive the creation of policies that favour the creation of new monitoring surveys, and support the creation of NSISs.
- Collect soil samples at deeper depths (e.g., including subsoil up to 2 m). Sample across the whole soil profile at the national level using approaches adapted to the circumstance of the country or regions. It requires approaches that are both environmental friendly and financially efficient.
- Capacity building of soil scientists at the national level on the whole process of DSM to prepare the next generation of digital soil mappers.
- More engagement among stakeholders currently championing DSM and soil information in the continent. There is a need for African soil scientists to cooperate, this can happen when there are more opportunities for networking at regional and continental level so that the scientists understand each other's work culture, develop mutual trust and have the confidence to initiate and implement joint work.

- Go beyond providing DSM products and support digital soil and land assessments (e.g., crop area allocation, fertiliser application, land use management, erosion control and mitigation, soil health management strategies, carbon sequestration, soil microbial maintenance, etc.) that can lead to actionable decisions and therefore contribute to the countries development.

5. Conclusion

The literature review on DSM presented in this article is the first of its kind for the continent of Africa. The following presents a summary of the main conclusions.

- Digital soil mapping has received increasing attention within the continent of Africa for the last two decades.
- Nearly 46% of the countries within the continent had at least one DSM study. Nearly half of the DSM studies focused on soil property or class map production or methodology development. The other half performed studies in which map production was not the main interest, but instead focused on quantifying soil functions and digital soil assessment.
- Several soil variables were mapped. The most mapped variable was soil organic carbon (SOC), whereas soil hydraulic variables were seldom reported.
- Most studies used machine learning algorithms and produced digital soil maps for the topsoil up to 30 cm depth. The majority of studies relied on legacy soil data as the source of the model calibration data and calculated validation statistics using existing datasets but without collecting a post-mapping probability sample. In addition, the area of the study cases was strongly negatively correlated with the sample density. Only a few studies estimated the prediction uncertainty.
- Due to the quality of the legacy soil data, most studies called for new data collection, which is often limited due to financial constraints. Initiatives for soil data collection and the creation of digital soil maps should be initiated and led from within the continent to support the long-term maintenance and development of monitoring networks and information systems.
- The covariates used in most studies were usually generated at a global scale, thus may most likely present bias at reflecting the local patterns of the landscape. Therefore, more research efforts are required on how to derive high-resolution environmental variables specific to either an individual country or the continent.
- Lack of digital skills was identified as one of the primary reasons, the other half of the countries in Africa had no DSM study. The creation of centres of excellence, in Africa, to train soil scientists with state-of-the-art DSM approaches is highly recommended to ensure the sustainability of DSM within the continent.
- Fewer studies focused on digital soil assessment to enhance decision-making and thus contribute to countries' development. The limited use of DSM for actionable decision-making may explain the nearly Nil investments in soil-related research. Therefore, digital soil mappers should generate value-added propositions to political institutions and engage in lobbying activities to drive the creation of soil information.
- Designing soil monitoring networks with the support of decision-makers, both locally and nationally, could support the generation of DSM products and their use in the long term for soil health, land suitability, and crop productivity risk assessment.
- There is a need to rebuild and maintain a data environment that is conducive for novel soil-related research and that aims at tackling environmental issues faced in Africa.

CRediT authorship contribution statement

Andree M. Nenkam: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alexandre M.J-C. Wadoux:** Writing – review & editing, Writing – original draft, Supervision. **Budiman Minasny:** Writing – review & editing, Supervision, Conceptualization. **Francis B.T. Silatsa:** Writing – review & editing, Writing – original draft. **Martin Yemefack:** Writing – review & editing, Writing – original draft. **Sabastine Ugbemuna Ugbaje:** Writing – review & editing, Writing – original draft. **Stephen Akpa:** Writing – review & editing, Writing – original draft. **George Van Zijl:** Writing – review & editing, Writing – original draft. **Abdelkrim Bouasria:** Writing – review & editing, Writing – original draft. **Yassine Bouslihim:** Writing – review & editing, Writing – original draft. **Lydia Mumbi Chabala:** Writing – review & editing, Writing – original draft. **Ashenafi Ali:** Writing – review & editing, Writing – original draft. **Alex B. McBratney:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

The data were included in the manuscript.

Acknowledgements

This research was inspired by the first author's PhD thesis on "High Resolution Digital Soil Mapping for Multiscale Prediction of Agricultural Production in sub-Saharan Africa" at the University of Sydney. We also acknowledge all those who have answered the survey shared by Dr. Yemefack. Their answers were useful for the completion of the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geoderma.2024.117007.

References

Abdel-Kader, F.H., 2011. Digital soil mapping at pilot sites in the northwest coast of Egypt: A multinomial logistic regression approach. Egypt. J. Remote Sens. Space Sci. 14 (1), 29–40.

Abdel-Kader, F.H., 2013. Digital soil mapping using spectral and terrain parameters and statistical modelling integrated into GIS-northwestern coastal region of Egypt. In: Developments in Soil Classification, Land Use Planning and Policy Implications. Springer, Cham, pp. 353–371.

Adhikari, K., Hartemink, A.E., 2016. Linking soils to ecosystem services—A global review. Geoderma 262, 101–111.

Africa Union, 2015. Food Security. https://au.int/en/auc/priorities/food-security. (Online; Accessed 28 January 2024).

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 Sci. Soc. America J. 78 (6), 1953–1966.

Akpa, S.I.C., Odeh, I.O.A., Bishop, T.F.A., Hartemink, A.E., Amapu, I.Y., 2016a. Total soil organic carbon and carbon sequestration potential in Nigeria. Geoderma 271, 202–215.

Akpa, S.I.C., Ugbaje, S.U., Bishop, T.F.A., Odeh, I.O.A., 2016b. Enhancing pedotransfer functions with environmental data for estimating bulk density and effective cation exchange capacity in a data-sparse situation. Soil Use Manag. 32 (4), 644–658.

Al-Maktoumi, A., Al-Ismaily, S., Kacimov, A., 2016. Based learning for undergraduate students in soil and water sciences: A case study of hydropedology in an arid-zone environment. J. Geography Higher Educ. 40 (3), 321–339.

Ali, A., Erkossa, T., Gudeta, K., Abera, W., Mesfin, E., Mekete, T., Haile, M., Haile, W., Abegaz, A., Tafesse, D., Belay, G., Getahun, M., Beyene, S., Assen, M., Regassa, A., Selassie, Y.G., Tadesse, S., Abebe, D., Walde, Y., Hussien, N., Yirdaw, A., Mera, A., Admas, T., Wakoya, F., Legesse, A., Tessema, N., A Abebe, A., Gebremariam, S., Aregaw, Y., Abebaw, B., Bekele, D., Zewdie, E., Schulz, S., Tamene, L., Elias, E., 2022. Reference soil groups map of Ethiopia based on legacy data and machine learning technique: EthioSoilGrids 1.0. EGUsphere 301, 1–40.
Ali, A., Erkossa, T., Gudeta, K., Abera, W., Mesfin, E., Mekete, T., Haile, M., Haile, W., Abegaz, A., Tafesse, D., et al., 2024. Reference soil groups map of Ethiopia based on legacy data and machine learning-technique: EthioSoilGrids 1.0. SOIL 10 (1), 189–209.
Alnaimy, M.A., Shahin, S.A., Afifi, A.A., Ewees, A.A., Junakova, N., Balintova, M., Abd Elaziz, M., 2022. Spatio prediction of soil capability modeled with modified RVFL using Aptenodytes Forsteri Optimization and digital soil assessment technique. Sustainability 14 (22), 14996.
Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B., 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., Sanchez, P.A., Thompson, J.A., Zhang, G.L., 2014. GlobalSoilMap: Toward a fine-resolution global grid of soil properties. Adv. Agronomy 125, 93–134.
Arrouays, D., Leenaars, J.G., Richer-de Forges, A.C., Adhikari, K., Ballabio, C., Greve, M., Grundy, M., Guerrero, E., Hempel, J., Hengl, T.e.a., 2017. Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. GeoResJ 14, 1–19.
Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de Forges, A.C., Morgan, C.L., Roudier, P., Poggio, L., Mulder, V.L., 2020. Impressions of digital soil maps: The good, the not so good, and making them ever better. Geod. Reg. 20, e00255.
Assami, T., Hamdi-Aissa, B., 2019. Digital mapping of soil classes in Algeria–A comparison of methods. Geod. Reg. 16, e00215.
Atkinson, J.T., Rozanov, A.B., De Clercq, W.P., 2017. Evaluating the effects of generalisation approaches and DEM resolution on the extraction of terrain indices in KwaZulu Natal, South Africa. South African J. Geomat. 6 (2), 245–261.
Australian Government, 2021. National Soil Strategy. Technical Report, Department of Agriculture, Water and the Environment, Canberra.
Australian Government, 2023. Australian National Soil Information System (ANSIS). https://www.csiro.au/en/research/natural-environment/land/soil/ansis. (Online; Accessed 28 January 2024).
Badraoui, M., Stitou, M., 2001. Status of soil survey and soil information systems in Morocco. Options Mediterr. Ser. B 34, 193–204.
Baeyens, J., 1938. Les sols de l'Afrique Centrale. Tome I: Le Bas-Congo.
Bahri, H., Raclot, D., Barbouchi, M., Lagacherie, P., Annabi, M., 2022. Mapping soil organic carbon stocks in Tunisian topsoils. Geod. Reg. 30, e00561.
Batjes, N.H., 2016. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. Geoderma 269, 61–68.
Biggs, A.J., Crawford, M., Burgess, J., Smith, D., Andrews, K., Sugars, M., 2022. Digital soil mapping in Australia. Can it achieve its goals? Soil Res. 61 (1), 1–8.
Bokar, H., Traoré, A., Mariko, A., Diallo, T., Traoré, A., Sy, A., Soumaré, O., Dolo, A., Bamba, F., Sacko, M., Touré, O., 2020. Geogenic influence and impact of mining activities on water soil and plants in surrounding areas of Morila Mine, Mali. J. Geochem. Explor. 209, 106429.
Boluwade, A., 2019. Regionalization and partitioning of soil health indicators for Nigeria using spatially contiguous clustering for economic and social-cultural developments. ISPRS Int. J. Geo-Inf. 8 (10), 458.
Bouasria, A., Ibno Namr, K., Rahimi, A., Ettachfini, E.M., Rerhou, B., 2022. Evaluation of Landsat 8 image pansharpening in estimating soil organic matter using multiple linear regression and artificial neural networks. Geo-Spatial Inf. Sci. 25, 353–364.
Bouasria, A., Namr, K.I., Rahimi, A., Ettachfini, E.M., 2020. Soil organic matter estimation by using Landsat-8 pansharpened image and machine learning. In: 2020 Fourth International Conference on Intelligent Computing in Data Sciences. ICDS, IEEE, pp. 1–8.
Bouma, J., Broll, G., Crane, T.A., Dewitte, O., Gardi, C., Schulte, R.P.O., Towers, W., 2012. Soil information in support of policy making and awareness raising. Curr. Opin. Environ. Sustain. 4 (5), 552–558.
Bouma, J., Montanarella, L., Evanylo, G., 2019. The challenge for the soil science community to contribute to the implementation of the UN sustainable development goals. Soil Use Manag. 35 (4), 538–546.
Bouslihim, Y., Rochdi, A., Aboutayeb, R., El Amrani-Paaza, N., Miftah, A., Hssaini, L., 2021. Soil aggregate stability mapping using remote sensing and GIS-based machine learning technique. Front. Earth Sci. 9, 748859.
Buenemann, M., Coetzee, M.E., Kutuahupira, J., Maynard, J.J., Herrick, J.E., 2023. Errors in soil maps: The need for better on-site estimates and soil map predictions. PLoS One 18 (1), e0270176.
Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., 2013. A methodology for digital soil mapping in poorly-accessible areas. Geoderma 192, 341–353.
Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A., 2014. Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty. Geoderma 213, 46–56.
Centre for Agriculture and Biosciences International, 2020. Rwanda Soil Information Services (RwaSIS). https://www.cabi.org/projects/rwanda-soil-information-services-rwandasis/. (Online; Accessed 29 March 2024).
Chabala, L.M., Mulolwa, A., Lungu, O., 2014. Mapping the spatial variability of soil acidity in Zambia. Agronomy 4 (4), 452–461.
Chabala, L.M., Mulolwa, A., Lungu, O., 2017. Application of ordinary kriging in mapping soil organic carbon in Zambia. Pedosphere 27 (2), 338–343.
Chapoto, A., Chabala, L.M., Lungu, O.N., 2016. A long history of low productivity in Zambia: Is it time to do away with blanket recommendations? Zambia Soc. Sci. J. 6 (2), 6.
Chen, S., Arrouays, D., Mulder, V.L., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de Forges, A.C., Walter, C., Meersmans, J., Richer-de Forges, C., 2022. Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. Geoderma 409, 115567.
Ciampalini, R., Lagacherie, P., Hamrouni, H., 2012a. Documenting GlobalSoilMap.net grid cells from legacy measured soil profile and global available covariates in Northern Tunisia. In: McBratney, A.B. (Ed.), In: Proceedings of the 5th Global Workshop on Digital Soil Mapping 2012, Sydney, Australia, vol. 1, (no. 1), CRC Press, United Kingdom, pp. 439–444.
Ciampalini, R., Lagacherie, P., Monestiez, P., Walker, E., Gomez, C., 2012b. Cokriging of soil properties with Vis-NIR hyperspectral covariates in the Cap Bon region (Tunisia). In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and beyond. CRC Press, London, pp. 393–398.
Claessens, L., Cassman, K., van Ittersum, M., Leenaars, J., van Bussel, L., Wolf, J., van Wart, J., Grassini, P., Yang, H., Boogaard, H., et al., 2015. The global yield gap atlas for targeting sustainable intensification options for smallholders in Sub-Saharan Africa. In: Wageningen Soil Conference 2015: Soil Science in a Changing World. pp. 43–43.
Cook, S.E., Jarvis, A., González, J.P., 2008. A new global demand for digital soil information. In: Hartemink, A.E., Mcbratney, A., Mendon̦ca Santos, M.L. (Eds.), In: Digital Soil Mapping with Limited Data, vol. 148, (no. 2), Springer Science & Business Media, pp. 31–41.
Courteille, L., Lagacherie, P., Boukhelifa, N., Lutton, E., Tardieu, L., 2024. Using spatial aggregation of soil multifunctionality maps to support uncertainty-aware planning decisions. Eur. J. Soil Sci. 75 (4), e13523.
Dakak, H., Huang, J., Zouahri, A., Douaik, A., Triantafilis, J., 2017. Mapping soil salinity in 3-dimensions using an EM38 and EM4Soil inversion modelling at the reconnaissance scale in central Morocco. Soil Use Manag. 33 (4), 553–567.
D'Amore, D.V., Fellman, J.B., Edwards, R.T., Hood, E., Ping, C.-L., 2012. Hydropedology of the North American Coastal Temperate Rainforest, vol. 351, Amsterdam, Netherlands, Academic Press.
de Oliveira, T.S., Bell, R.W., 2022. Introduction to subsoil constraints for crop production. In: Subsoil Constraints for Crop Production. Springer, pp. 1–10.
Dembele, D., Traore, K., Quansh, C., Jnr, E., B A B, B.M., 2016. Optimizing soil fertility management decision in Mali by remote sensing and GIS. Donnis J. Agric. Res. 3 (4), 22–34.
Dewitte, O., Jones, A., Spaargaren, O., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Michéli, E., Montanarella, L., Thiombiano, L., Van Ranst, E., Yemefack, M., Zougmore, R., 2013. Harmonisation of the soil map of Africa at the continental scale. Geoderma 211, 138–153.
Dlamini, P., Chaplot, V., 2012. On the interpolation of volumetric water content in research catchments. Phys. Chem. Earth, Parts A/B/C 50, 165–174.
Du Plessis, C., Van Zijl, G., Van Tol, J., Manyevere, A., 2020. Machine learning digital soil mapping to inform gully erosion mitigation measures in the Eastern Cape, South Africa. Geoderma 368, 114287.
Eberle, D.G., Paasche, H., 2012. Integrated data analysis for mineral exploration: A case study of clustering satellite imagery, airborne gamma-ray, and regional geochemical data suites. Geophysics 77 (4), B167–B176.
Ethiopian Agricultural Transformation Agency, 2017. The Ethiopian Soil Information System (EthioSIS). Technical Report.
Ethiopian Government, 2022. Ethiopian soil information system. https://nsis.moa.gov.et/. (Online; Accessed 28 January 2024).
European Commission, 2020. Soils4Africa. https://www.soils4africa-h2020.eu/. (Online; Accessed 15 January 2024).
Evangelista, S.J., Field, D.J., McBratney, A.B., Minasny, B., Ng, W., Padarian, J., Dobarco, M.R., Wadoux, A.M.-C., 2023. A proposal for the assessment of soil security: Soil functions, soil services and threats to soil. Soil Secur. 10, 100086.
FAO, 2015. Western and Central African countries gain expertise in digital soil mapping. Press Release, https://www.fao.org/africa/news/detail-news/en/c/280840/. consulted January 12th, 2024.
FAO-UNESCO, 1977. Soil Map of the World. Technical Report, 1:5 M, Volume VI, Africa.
Finke, P.A., 2012. On digital soil assessment with models and the Pedometrics agenda. Geoderma 171, 3–15.
Flynn, T., De Clercq, W., Rozanov, A., Clarke, C., 2019a. High-resolution digital soil mapping of multiple soil properties: An alternative to the traditional field survey? South African J. Plant Soil 36 (4), 237–247.
Flynn, T., Rozanov, A., Clarke, C., 2020. Input map and feature selection for soil legacy data. Geoderma 375, 114452.
Flynn, T., Rozanov, A., de Clercq, W., Warr, B., Clarke, C., 2019b. Semi-automatic disaggregation of a national resource inventory into a farm-scale soil depth class map. Geoderma 337, 1136–1145.

Flynn, T., Rozanov, A., Ellis, F., de Clercq, W., Clarke, C., 2022a. Farm-scale digital soil mapping of soil classes in South Africa. South Afr. J. Plant Soil 39 (3), 175–186.

Flynn, T., Van Zijl, G., Van Tol, J., Botha, C., Rozanov, A., Warr, B., Clarke, C., 2019c. Comparing algorithms to disaggregate complex soil polygons in contrasting environments. Geoderma 352, 171–180.

Flynn, T., Wiese, L., Rozanov, A., 2022b. Soil carbon stock assessment using depth and spatial models on afforested arable lands. South Afr. J. Plant Soil 39 (4), 235–247.

Forkuor, G., Hounkpatin, O.K.L., Welp, G., Thiel, M., 2017. High resolution mapping of soil properties using remote sensing variables in south-western Burkina Faso: A comparison of machine learning and multiple linear regression models. PLoS One 12 (1), e0170478.

Gasmi, A., Gomez, C., Chehbouni, A., Dhiba, D., El Gharous, M., 2022a. Using PRISMA hyperspectral satellite imagery and GIS approaches for soil fertility mapping (FertiMap) in northern Morocco. Remote Sens. 14 (16), 4080.

Gasmi, A., Gomez, C., Chehbouni, A., Dhiba, D., Elfil, H., 2022b. Satellite multi-sensor data fusion for soil clay mapping based on the spectral index and spectral bands approaches. Remote Sens. 14 (5), 1103.

Gasmi, A., Gomez, C., Lagacherie, P., Zouari, H., Laamrani, A., Chehbouni, A., 2021. Mean spectral reflectance from bare soil pixels along a Landsat-TM time series to increase both the prediction accuracy of soil clay content and mapping coverage. Geoderma 388, 114864.

Gebretsadik, M., 2014. Soil Moisture Prediction in an Agricultural Field of Gumara-Maksegnit Watershed, North Gonder, Ethiopia

Gomez, C., Lagacherie, P., Bacha, S., 2012. Using Vis-NIR hyperspectral data to map topsoil properties over bare soils in the Cap Bon region, Tunisia. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and beyond. CRC Press, London, pp. 387–392.

Guilpart, N., Grassini, P., Van Wart, J., Yang, H., Van Ittersum, M.K., Van Bussel, L.G., Wolf, J., Claessens, L., Leenaars, J.G., Cassman, K.G., 2017. Rooting for food security in Sub-Saharan Africa. Environ. Res. Lett. 12, 114036.

Han, E., Ines, A.V., Koo, J., 2019. Development of a 10-km resolution global soil profile dataset for crop modeling applications. Environ. Model. Software 119, 70–83.

Hansen, M.K., Brown, D.J., Dennison, P.E., Graves, S.A., Bricklemyer, R.S., 2009. Inductively mapping expert-derived soil-landscape units within dambo wetland catenae using multispectral and topographic data. Geoderma 150 (1–2), 72–84.

Hartemink, A.E., Hempel, J., Lagacherie, P., McBratney, A.B., McKenzie, N., MacMillan, R.A., Minasny, B., Montanarella, L., de Mendonça Santos, M.L., Sanchez, P., Walsh, M., Zhang, G.-L., 2010. GlobalSoilMap.net–A new digital soil map of the world. In: Boettinger, J.L., Howell, D.W., Moore, A.C., Hartemink, A.E., Kienast-Brown, S. (Eds.), Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. Springer Science & Business Media, London, pp. 423–428.

Hartemink, A.E., McBratney, A., 2008. A soil science renaissance. Geoderma 148 (2), 123–129.

Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., Mendes de Jesus, J., Tamene, L., Tondoh, J.E., 2015. Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. PLoS One 10 (6), e0125814.

Hengl, T., Leenaars, J.G., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Mendonca-Santos, M.D.L., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A., Zhang, G.L., 2017a. Soil nutrient maps of sub-saharan Africa: Assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosyst. 109, 77–102.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12 (2), e0169748.

Hengl, T., Miller, M.A.E., Križan, J., Shepherd, K.D., Sila, A., Kilibarda, M., Antonijević, O., Glušica, L., Dobermann, A., Haefele, S.M., McGrath, S.P., Acquah, G.E., Collinson, J., Parente, L., Sheykhmousa, M., Saito, K., Johnson, J.M., Chamberlin, J., Silatsa, F.B.T., Yemefack, M., Wendt, J., MacMillan, R.A., Wheeler, I., Crouch, J., 2021b. African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. Sci. Rep. 11 (1), 1–18.

Hof, S., 2014. Mapping Soil Variability with a Decision Tree Modelling Approach in the Northern Highlands of Ethiopia. Wageningen University, The Netherland.

Hordofa, M., 2020. Blended fertilizers validation based on ethio-SIS soil fertility map at Halaba, Southern Ethiopia. Int. J. Innov. Agric. Biol. Res. 8 (3), 1–6.

Hornby, A.J.W., 1938. Soil map of central nyasaland.

Hou, D., Bolan, N.S., Tsang, D.C.W., Kirkham, M.B., O'Connor, D., 2020. Sustainable soil use and management: An interdisciplinary and systematic approach. Sci. Total Environ. 729, 138961.

Hounkpatin, K.O.L., Bossa, A.Y., Yira, Y., Igue, M.A., Sinsin, B.A., 2022. Assessment of the soil fertility status in Benin (West Africa)–Digital soil mapping using machine learning. Geod. Reg. 28, e00444.

Hounkpatin, O.K.L., de Hipt, F.O.p., Bossa, A.Y., Welp, G., Amelung, W., 2018a. Soil organic carbon stocks and their determining factors in the Dano catchment (Southwest Burkina Faso). CATENA 166, 298–309.

Hounkpatin, K.O.L., Schmidt, K., Stumpf, F., Forkuor, G., Behrens, T., Scholten, T., Amelung, W., Welp, G., 2018b. Predicting reference soil groups using legacy data: A data pruning and random forest approach for tropical environment (Dano catchment, Burkina Faso). Sci. Rep. 8 (1), 1–16.

Ismail, M., Yacoub, R.K., 2012. Digital soil map using the capability of new technology in Sugar Beet area, Nubariya, Egypt. Egypt. J. Remote Sens. Space Sci. 15 (2), 113–124.

Iticha, B., Kamran, M., Yan, R., Siuta, D., Al-Hashimi, A., Takele, C., Olana, F., Kukfisz, B., Iqbal, S., Elshikh, M.S., 2022. The role of digital soil information in assisting precision soil management. Sustainability 14 (18), 11710.

Iticha, B., Takele, C., 2019. Digital soil mapping for site-specific management of soils. Geoderma 351, 85–91.

John, K., Bouslihim, Y., Bouasria, A., Razouk, R., Hssaini, L., Isong, I.A., Ait M'barek, S., Ayito, E.O., Ambrose-Igho, G., 2022a. Assessing the impact of sampling strategy in random forest-based predicting of soil nutrients: A study case from northern Morocco. Geocarto Int. 1–14.

John, K., Bouslihim, Y., Isong, I.A., Hssaini, L., Razouk, R., Kebonye, N.M., Agyeman, P.C., Penížek, V., Zádorová, T., 2022b. Mapping soil nutrients via different covariates combinations: Theory and an example from Morocco. Ecol. Processes 11 (1), 1–17.

John, K., Bouslihim, Y., Ofem, K.I., Hssaini, L., Razouk, R., Okon, P.B., Isong, I.A., Agyeman, P.C., Kebonye, N.M., Qin, C., 2022c. Do model choice and sample ratios separately or simultaneously influence soil organic matter prediction? Int. Soil Water Conserv. Res. 10 (3), 470–486.

Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Micheli, E., Montanarella, L., Spaargaren, O., Thiombiano, L., Van Ranst, E., Yemefack, M., Zougmoré, R., 2013. Soil Atlas of Africa. Technical Report, European Commission, Publications Office of the European Union, Luxembourg.

Julich, S., Moorcroft, M.-A., Feger, K., van Tol, J., 2022. The impact of overgrazing on water fluxes in a semi-arid watershed–the suitability of watershed-scale modeling in a data scarce area. J. Hydrol.: Reg. Stud. 43, 101178.

Kamamia, A.W., Vogel, C., Mwangi, H.M., Feger, K.-H., Sang, J., Julich, S., 2021. Mapping soil aggregate stability using digital soil mapping: A case study of Ruiru reservoir catchment, Kenya. Geod. Reg. 24, e00355.

Kamamia, A.W., Vogel, C., Mwangi, H.M., Feger, K.-h., Sang, J., Julich, S., 2022. Using soil erosion as an indicator for integrated water resources management: A case study of Ruiru drinking water reservoir, Kenya. Environ. Earth Sci. 81 (21), 502.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18 (3–4), 267–288.

Kempen, B., 2005. Digital Soil Mapping in the Nioro du Rip Area, Senegal (Ph.D. thesis). Wageningen University.

Kome, G., Silatsa, F., Yemefack, M., 2021. Status of salt-affected soils in Cameroon. In: Global Symposium on Salt-Affected Soils, Rome. FAO's Global Soil Partnership (GSP), pp. 67–68.

Lagacherie, P., Arrouays, D., Bourennane, H., Gomez, C., Martin, M., Saby, N.P.A., 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.

Lagacherie, P., Buis, S., Constantin, J., Dharumarajan, S., Ruiz, L., Sekhar, M., 2022. Evaluating the impact of using digital soil mapping products as input for spatializing a crop model: The case of drainage and maize yield simulated by STICS in the Berambadi catchment (India). Geoderma 406, 115503.

Lagacherie, P., McBratney, A.B., Voltz, M., 2006. Digital Soil Mapping: An Introductory Perspective. Elsevier, Boston.

Leenaars, J.G.B., Claessens, L., Heuvelink, G.B.M., Hengl, T., González, M.R., van Bussel, L.G.J., Guilpart, N., Yang, H., Cassman, K.G., 2018. Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. Geoderma 324, 18–36.

Leenaars, J.G.B., Elias, E., Wösten, J.H.M., Ruiperez-González, M., Kempen, B., 2020. Mapping the major soil-landscape resources of the Ethiopian Highlands using random forest. Geoderma 361, 114067.

Leenaars, J.G.B., Kwabena, N.A., Silatsa, F.B.T., Fening, J.O., Yemefack, M., Hengl, T., Heuvelink, G., 2017. World soil information developing from global, continental and national initiatives. GlobalSoilMap-Digital Soil Mapping from Country to Globe. July 4-6, 2017, Moscow, Russia.

Leenaars, J.G.B., van Oostrum, A.J.M., Ruiperez Gonzalez, M., 2014. Africa Soil Profiles Database: A compilation of georeferenced and standardised legacy soil profile data for Sub-Saharan Africa. Technical Report, ISRIC-World Soil Information, Version 1.2.

Lin, H., 2003. Hydropedology: Bridging disciplines, scales, and data. Vadose Zone J. 2 (1), 1–11.

Liu, Y., Heuvelink, G.B., Bai, Z., He, P., Xu, X., Ma, J., Masiliūnas, D., 2020. Space-time statistical analysis and modelling of nitrogen use efficiency indicators at provincial scale in China. Eur. J. Agron. 115, 126032.

Makungwe, M., Chabala, L.M., Chishala, B.H., Lark, R.M., 2021. Performance of linear mixed models and random forests for spatial prediction of soil pH. Geoderma 397, 115079.

Mallavan, B.P., Minasny, B., McBratney, A.B., 2010. Homosoil, a methodology for quantitative extrapolation of soil information across the globe. Digit. Soil Mapping: Bridging Res., Environ. Appl. Oper. 137–150.

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

Mamassi, A., Marrou, H., El Gharous, M., Wellens, J., Jabbour, F.-E., Zeroual, Y., Hamma, A., Tychon, B., 2022. Relevance of soil fertility spatial databases for parameterizing APSIM-wheat crop model in Moroccan rainfed areas. Agron. Sustain. Dev. 42 (5), 83.

Mamera, M., van Tol, J.J., 2018. Application of hydropedological information to conceptualize pollution migration from dry sanitation systems in the Ntabelanga Catchment Area, South Africa. Air, Soil Water Res. 11, 1178622118795485.

Massawe, B.H.J., Slater, B.K., Subburayalu, S.K., Kaaya, A.K., Winowiecki, L., 2016. Updating legacy soil maps for climate resilient agriculture: A case of Kilombero Valley, Tanzania. In: Climate Change and Multi-Dimensional Sustainability in African Agriculture. Springer, Cham, pp. 345–364.

Massawe, B.H.J., Subburayalu, S.K., Kaaya, A.K., Winowiecki, L., Slater, B.K., 2018. Mapping numerically classified soil taxa in Kilombero Valley, Tanzania using machine learning. Geoderma 311, 143–148.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer functions to soil inference systems. Geoderma 109 (1–2), 41–73.

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

Minai, J.O., Libohova, Z., Schulze, D.G., 2021. Spatial prediction of soil properties for the Busia area, Kenya using legacy soil data. Geod. Reg. 25, e00366.

Minasny, B., Fiantis, D., Mulyanto, B., Sulaeman, Y., Widyatmanti, W., 2020. Global soil science research collaboration in the 21st century: Time to end helicopter research. Geoderma 373, 114299.

Minasny, B., Hartemink, A.E., 2011. Predicting soil properties in the tropics. Earth-Sci. Rev. 106 (1–2), 52–62.

Montanarella, L., 2015. The global soil partnership. In: IOP Conference Series: Earth and Environmental Science, vol. 25, (no. 1), IOP Publishing Ltd, 012001.

Mora-Vallejo, A., Claessens, L., Stoorvogel, J., Heuvelink, G.B.M., 2008. Small scale digital soil mapping in Southeastern Kenya. CATENA 76 (1), 44–53.

Moroccan Government, 2016. Soil fertility maps of Morocco. http://www.fertimap.ma/en/presentation.html. (Online; Accessed 28 January 2024).

Mponela, P., Snapp, S., Villamor, G., Tamene, L., Le, Q.B., Borgemeister, C., 2020. Digital soil mapping of nitrogen, phosphorus, potassium, organic carbon and their crop response thresholds in smallholder managed escarpments of Malawi. Appl. Geogr. 124, 102299.

Mukumbuta, I., Chabala, L., Sichinga, S., Lark, R., 2022. Accessing and assessing legacy soil information, an example from two provinces of Zambia. Geoderma 420, 115874.

Mukumbuta, I., Chabala, L., Sichinga, S., Miti, C., Lark, R., 2021. A comparison between three legacy soil maps of Zambia at national scale: The spatial patterns of legend units and their relation to soil properties. Geoderma 402, 115193.

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. Sci. Total Environ. 573, 1352–1369.

Mwendwa, S.M., Mbuvi, J.P., Kironchi, G., Gachene, C.K.K., 2022. Assessing spatial variability of selected soil properties in Upper Kabete Campus coffee farm, University of Nairobi, Kenya. Heliyon 8 (8), e10190.

Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Aksoy, E., Batjes, N., Boateng, E., et al., 2023. Harmonized World Soil Database Version 2.0. Food and Agriculture Organization of the United Nations.

Nawar, S., Buddenbaum, H., Hill, J., 2015. Digital mapping of soil properties using multivariate statistical analysis and ASTER data in an arid region. Remote Sens. 7 (2), 1181–1205.

Nenkam, A.M., Wadoux, A.M.J.-C., Minasny, B., McBratney, A.B., Traore, P.C.S., Falconnier, G.N., Whitbread, A.M., 2022. Using homosoils for quantitative extrapolation of soil mapping models. Eur. J. Soil Sci. 73 (5), e13285.

Nenkam, A.M., Wadoux, A.M.J.C., Minasny, B., McBratney, A.B., Traore, P.C., Whitbread, A.M., 2023. Using homosoils to enrich sparse soil data infrastructure: An example from Mali. CATENA 223, 106862.

Nguemezi, C., Tematio, P., Silatsa, F.B., Yemefack, M., 2021. Spatial variation and temporal decline (1985–2017) of soil organic carbon stocks (SOCS) in relation to land use types in Tombel area, South-West Cameroon. Soil Tillage Res. 213, 105114.

Ngunjiri, M.W., Libohova, Z., Minai, J.O., Serrem, C., Owens, P.R., Schulze, D.G., 2019. Predicting soil types and soil properties with limited data in the Uasin Gishu Plateau, Kenya. Geod. Reg. 16, e00210.

Nijbroek, R., Piikki, K., Söderström, M., Kempen, B., Turner, K.G., Hengari, S., Mutua, J., 2018. Soil organic carbon baselines for land degradation neutrality: Map accuracy and cost tradeoffs with respect to complexity in Otjozondjupa, Namibia. Sustainability 10 (5), 1610.

Nketia, K.A., Asabere, S.B., Ramcharan, A., Herbold, S., Erasmi, S., Sauer, D., 2022. Spatio-temporal mapping of soil water storage in a semi-arid landscape of northern Ghana–A multi-tasked ensemble machine-learning approach. Geoderma 410, 115691.

Odeh, I.O.A., Leenaars, J., Hartemink, A., Amapu, I., 2012. The challenges of collating legacy data for digital mapping of Nigerian soils. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and beyond. CRC Press, London, pp. 453–458.

Odgers, N.P., McBratney, A.B., Minasny, B., Sun, W., Clifford, D., 2014. DSMART: An algorithm to spatially disaggregate soil map units. In: Arrouays, D., McKenzie, N., Hempel, J., Richer-de Forges, A.C., McBratney, A.B. (Eds.), GlobalSoilMap: Basis of the Global Spatial Soil Information System. Taylor & Francis, London, pp. 261–266.

Oliveira, J.C.M., Reichardt, K., Vaz, C.M.P., Swartzendruber, D., 1997. Improved soil particle-size analysis by gamma-ray attenuation. Soil Sci. Soc. America J. 61 (1), 23–26.

Omuto, C.T., 2013. Major soil and data types in Kenya. In: Paron, P., Ochieng, O.D., Omuto, C.T. (Eds.), In: Kenya: A Natural Outlook, vol. 16, Elsevier, pp. 123–132.

Omuto, C.T., Vargas, R.R., Elmobarak, A.A., Mapeshoane, B.E., Koetlisi, K.A., Ahmadzai, H., Abdalla Mohamed, N., 2022. Digital soil assessment in support of a soil information system for monitoring salinization and sodification in agricultural areas. Land Degrad. Dev. 33 (8), 1204–1218.

Owusu, S., Yigini, Y., Olmedo, G.F., Omuto, C.T., 2020. Spatial prediction of soil organic carbon stocks in Ghana using legacy data. Geoderma 360, 114008.

Pachepsky, Y.A., Rawls, W.J., Lin, H.S., 2006. Hydropedology and pedotransfer functions. Geoderma 131 (3–4), 308–316.

Panagos, P., Jones, A., Bosco, C., Kumar, P.S., 2011. European digital archive on soil maps (EuDASM): preserving important soil data for public free access. Int. J. Digit. Earth 4 (5), 434–443.

Parwada, C., van Tol, J., 2020. Mapping Soil Erosion Sensitive Areas in Organic Matter Amended Soil Associations in the Ntabelanga area, Eastern Cape Province, South Africa. J. Appl. Sci. Environ. Manag. 24 (9), 1693–1702.

Paterson, G., Turner, D., Wiese, L., Van Zijl, G., Clarke, C., Van Tol, J., 2015. Spatial soil information in South Africa: Situational analysis, limitations and challenges. South Afr. J. Sci. 111 (5–6), 1–7.

Pennock, D., Bedard-Haughn, A., Kiss, J., van der Kamp, G., 2014. Application of hydropedology to predictive mapping of wetland soils in the Canadian Prairie Pothole Region. Geoderma 235, 199–211.

Piikki, K., Söderström, M., Cordingley, J., 2017. Improved usefulness of continental soil databases for agricultural management through local adaptation. South Afr. J. Plant Soil 34 (1), 35–45.

Pinto, L.C., Mello, C.R.d., Norton, L.D., Silva, S.H.G., Taveira, L.R.S., Curi, N., 2017. Land-use effect on hydropedology in a mountainous region of Southeastern Brazil. Ciência e Agrotecnologia 41, 413–427.

Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240.

Rafik, A., Ibouh, H., El Alaoui El Fels, A., Eddahby, L., Mezzane, D., Bousfoul, M., Amazirh, A., Ouhamdouch, S., Bahir, M., Gourfi, A., Dhiba, D., Chehbouni, A., 2022. Soil salinity detection and mapping in an environment under water stress between 1984 and 2018 (Case of the Largest Oasis in Africa-Morocco). Remote Sens. 14 (7), 1606.

Ramakhanna, S.J., Mapeshoane, B.E., Omuto, C.T., 2022. Carbon sequestration potential in croplands in Lesotho. Ecol. Model. 471, 110052.

Ramifehiarivo, N., Brossard, M., Grinand, C., Andriamananjara, A., Razafimbelo, T., Rasolohery, A., Razafimahatratra, H., Seyler, F., Ranaivoson, N., Rabenarivo, M., Albrecht, A., Razafindrabe, F., Razakamanarivo, H., 2017. Mapping soil organic carbon on a national scale: Towards an improved and updated map of madagascar. Geod. Reg. 9, 29–38.

Reinhardt, N., Herrmann, L., 2019. Gamma-ray spectrometry as versatile tool in soil science: A critical review. J. Plant Nutr. Soil Sci. 182 (1), 9–27.

Román Dobarco, M., Wadoux, A.M.J., Malone, B., Minasny, B., McBratney, A.B., Searle, R., 2022. Mapping soil organic carbon fractions for Australia, their stocks and uncertainty. Biogeosci. Discuss. 2022, 1–38.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111 (9), 3268–3273.

Ruiperez Gonzalez, M., Kempen, B., Bindraban, P.S., Wolters, S., Hicintuka, C., Nibasumba, M., Veerkamp, J., 2015. Digital mapping of soil nutrients for the Republics of Burundi and Rwanda. Soil Sci. Changing World 179.

Samsonova, V.P., Meshalkina, J.L., Blagoveschensky, Y.N., Yaroslavtsev, A.M., Stoorvogel, J.J., 2018. The role of positional errors while interpolating soil organic carbon contents using satellite imagery. Precision Agric. 19, 1085–1099.

Schuler, U., Erbe, P., Zarei, M., Rangubpit, W., Surinkum, A., Stahr, K., Herrmann, L., 2011. A gamma-ray spectrometry approach to field separation of illuviation-type WRB reference soil groups in northern Thailand. J. Plant Nutr. Soil Sci. 174 (4), 536–544.

Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: A review. Progress Phys. Geography 27 (2), 171–197.

Shantz, H.L., Marbut, C.F., Kincer, J.B., 1923. The Vegetation and Soils of Africa. National Research Council and the American Geographical Society.

Silatsa, F.B.T., Tabi, F.O., Yemefack, M., Wilczok, C., Heuvelink, G.B.M., Hengl, T., Leenaars, J.G.B., 2017. Digital soil mapping using SoilGrids and national soil data in Cameroon. In: Arrouays, D., Savin, I., Leenaars, J., McBratney, A.B. (Eds.), GlobalSoilMap - Digital Soil Mapping from Country To Globe. CRC Press, London, pp. 43–48.

Silatsa, F.B.T., Yemefack, M., Tabi, F.O., Heuvelink, G.B.M., Leenaars, J.G.B., 2020. Assessing countrywide soil organic carbon stock using hybrid machine learning modelling and legacy soil data in Cameroon. Geoderma 367, 114260.

Sindayihebura, A., Ottoy, S., Dondeyne, S., Van Meirvenne, M., Van Orshoven, J., 2017. Comparing digital soil mapping techniques for organic carbon and clay content: Case study in Burundi's central plateaus. CATENA 156, 161–175.

Smit, E., van Tol, J., 2022. Impacts of soil information on process-based hydrological modelling in the Upper Goukou Catchment, South Africa. Water 14 (3), 407.

Snapp, S., 2022. Embracing variability in soils on smallholder farms: New tools and better science. Agricult. Sys. 195, 103310.

Sori, G., Iticha, B., Takele, C., 2021. Spatial prediction of soil acidity and nutrients for site-specific soil management in Bedele district, Southwestern Ethiopia. Agric. Food Secur. 10 (1), 1–15.

Stoorvogel, J.J., Kempen, B., Heuvelink, G.B.M., De Bruin, S., 2009. Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. Geoderma 149 (1–2), 161–170.

Swileam, G.S., Shahin, R.R., Nasr, H.M., Essa, K.S., 2019. Assessment of soil variability using electrical resistivity technique for normal alluvial soils, Egypt. Plant Archives 19 (1), 905–912.

Takoutsing, B., Heuvelink, G.B.M., 2022. Comparing the prediction performance, uncertainty quantification and extrapolation potential of regression kriging and random forest while accounting for soil measurement errors. Geoderma 428, 116192.

Takoutsing, B., Heuvelink, G.B.M., Stoorvogel, J.J., Shepherd, K.D., Aynekulu, E., 2022. Accounting for analytical and proximal soil sensing errors in digital soil mapping. Eur. J. Soil Sci. 73 (2), e13226.

Takoutsing, B., Martín, J.A.R., Weber, J.C., Shepherd, K., Sila, A., Tondoh, J., 2017. Landscape approach to assess key soil functional properties in the highlands of Cameroon: Repercussions of spatial relationships for land management interventions. J. Geochem. Explor. 178, 35–44.

Trapnell, C.G., Clothier, J.N., 1937. The Soils, Vegetation and Agricultural Systems of North Western Rhodesia. Report of the Ecological Survey. Government Printer, Lusaka.

Trapnell, C.G., Martin, J.D., Allan, W., 1948. Vegetation-Soil Map of Northern Rhodesia with Accompanying Explanatory Memoir by C.G. Trapnell. Government Printer, Lusaka (2nd Ed. 1950).

Ugbaje, S.U., Reuter, H.I., 2013. Functional digital soil mapping for the prediction of available water capacity in Nigeria using legacy data. Vadose Zone J. 12 (4), 1–13.

Uwiragiye, Y., Ngaba, M.J.Y., Zhao, M., Elrys, A.S., Heuvelink, G.B.M., Zhou, J., 2022. Modelling and mapping soil nutrient depletion in humid highlands of East Africa using ensemble machine learning: A case study from Rwanda. CATENA 217, 106499.

Vågen, T.G., Shepherd, K.D., Walsh, M.G., Winowiecki, L., Desta, L.T., Tondoh, J.E., 2010. AfSIS Technical Specifications: Soil Health Surveillance. World Agroforestry Centre, Nairobi, Kenya.

Vågen, T.-G., Winowiecki, L.A., Tondoh, J.E., Desta, L.T., Gumbricht, T., 2016. Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. Geoderma 263, 216–225.

Vågen, T., Winowiecki, L.A., Walsh, M.G., Tamene, L.D., Tondoh, J.E., 2010. Land degradation surveillance framework (LSDF): Field guide.

Van Apeldoorn, D.F., Kempen, B., Bartholomeus, H.M., Rusinamhodzi, L., Zingore, S., Sonneveld, M.P.W., Kok, K., Giller, K.E., 2014. Analysing soil organic C gradients in a smallholder farming village of East Zimbabwe. Geod. Reg. 2, 32–40.

van Leeuwen, C.C., Mulder, V.L., Batjes, N.H., Heuvelink, G.B.M., 2022. Statistical modelling of measurement error in wet chemistry soil data. Eur. J. Soil Sci. 73 (1), e13137.

Van Niekerk, A., 2014. Stellenbosch University Digital Elevation Model (SUDEM). Technical Report, Centre for Geographical Analysis SU (ed.).

Van Ranst, E., Verdoodt, A., Baert, G., 2010. Soil mapping in Africa at the crossroads: Work to make up for lost ground. Bull. des Séances d'Acad. R. Sci. d'Outre-Mer 56, 147–163.

van Tol, J., 2020. Hydropedology in South Africa: Advances, applications and research opportunities. South Afr. J. Plant Soil 37 (1), 23–33.

Van Tol, J., Van Zijl, G., 2022. South Africa needs a hydrological soil map: a case study from the upper uMngeni catchment. Water SA 48 (4), 335–347.

Van Tol, J.J., Van Zijl, G.M., Riddell, E.S., Fundisi, D., 2015. Application of hydropedological insights in hydrological modelling of the Stevenson-Hamilton research supersite, Kruger National Park, South Africa. Water SA 41 (4), 525–533.

Van Zijl, G.M., Botha, J.O., 2016. In pursuit of a South African national soil database: Potential and pitfalls of combining different soil data sets. South Afr. J. Plant Soil 33 (4), 257–264.

Van Zijl, G.M., Bouwer, D., Van Tol, J.J., Le Roux, P.A.L., 2014. Functional digital soil mapping: A case study from Namarroi, Mozambique. Geoderma 219, 155–161.

Van Zijl, G., Le Roux, P., 2014. Creating a conceptual hydrological soil response map for the Stevenson Hamilton research supersite, Kruger National Park, South Africa. Water SA 40 (2), 331–336.

Van Zijl, G.M., Le Roux, P.A.L., Smith, H.J.C., 2012. Rapid soil mapping under restrictive conditions in Tete, Mozambique. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and beyond. CRC Press, London, pp. 335–339.

Van Zijl, G.M., Le Roux, P.A.L., Turner, D.P., 2013. Disaggregation of land types using terrain analysis, expert knowledge and GIS methods. South Afr. J. Plant Soil 30 (3), 123–129.

Van Zijl, G., Van Tol, J., Bouwer, D., Lorentz, S., Le Roux, P., 2020. Combining historical remote sensing, digital soil mapping and hydrological modelling to produce solutions for infrastructure damage in Cosmo City, South Africa. Remote Sens. 12 (3), 433.

Venter, Z.S., Hawkins, H.-J., Cramer, M.D., Mills, A.J., 2021. Mapping soil organic carbon stocks and trends with satellite-driven high resolution maps over South Africa. Sci. Total Environ. 771, 145384.

Wadoux, A.M.J.C., Minasny, B., McBratney, A.B., 2020. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Sci. Rev. 210, 103359.

Wedajo Abdi, M., 2019. Localizing ethiosis fertility map based fertilizer type recommendation for maize (Zea mays L.) in coffee and spice production belt in Yeki district, Southwest of Ethiopia (Ph.D. thesis). Jimma University.

Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R., 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol. Indic. 52, 394–403.

Were, K., Singh, B.R., Dick, Ø.B., 2016. Spatially distributed modelling and mapping of soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya. J. Geogr. Sci. 26 (1), 102–124.

Were, K.O., Tien Bui, D., Dick, Ø.B., Singh, B.R., 2017. A novel evolutionary genetic optimization-based adaptive neuro-fuzzy inference system and geographical information systems predict and map soil organic carbon stocks across an Afromontane landscape. Pedosphere 27 (5), 877–889.

van der Westhuizen, S., Heuvelink, G.B.M., Hofmeyr, D.P., Poggio, L., 2022. Measurement error-filtered machine learning in digital soil mapping. Spatial Stat. 47, 100572.

Wiese, L., 2019. Mapping Soil Organic Carbon Stocks by Combining NIR sPectroscopy and Stochastic Vertical Distribution Models: a Case Study in the Mvoti River Catchment, K Z N, South Africa (Ph.D. thesis). Stellenbosch: Stellenbosch University, Chapter 6: Improving input parameters for soil organic carbon assessment – effect on errors from point measurements to final map.

Wiese, L., Ros, I., Rozanov, A., Boshoff, A., de Clercq, W., Seifert, T., 2016. An approach to soil carbon accounting and mapping using vertical distribution functions for known soil types. Geoderma 263, 264–273.

Zhang, S., 2013. Comparison of Statistical Methods for Digital Soil Mapping of Sub-Saharan Africa. Centre for Geo-Information.

Zhu, A.-X., 1997. A similarity model for representing soil spatial information. Geoderma 77 (2–4), 217–242.

van Zijl, G.M., van Tol, J.J., Riddell, E.S., 2016. Digital soil mapping for hydrological modelling. pp. 115–129, Digital soil mapping across paradigms, scales and boundaries.

van Zijl, G., van Tol, J., Tinnefeld, M., Le Roux, P., 2019. A hillslope based digital soil mapping approach, for hydropedological assessments. Geoderma 354, 113888.