Ten challenges for the future of pedometrics

Alexandre M.J.-C. Wadoux, Gerard B.M. Heuvelink, R. Murray Lark, Philippe Lagacherie, Johan Bouma, Vera Mulder, Zamir Libohova, Lin Yang, Alex Mcbratney

Ten challenges for the future of Pedometrics

Article *in* Geoderma · November 2021

DOI: 10.1016/j.geoderma.2021.115155

9 authors, including:


Alexandre Wadoux
French National Institute for Agriculture, Food, and Environment (INRAE)


Gerard Heuvelink
Wageningen University & Research


P. Lagacherie
French National Institute for Agriculture, Food, and Environment (INRAE)


Johan Bouma
Wageningen University & Research

Ten challenges for the future of Pedometrics

Alexandre M.J-C. Wadoux[a,*], Gerard B.M. Heuvelink[b,c], R. Murray Lark[d], Philippe Lagacherie[e], Johan Bouma[f], Vera L. Mulder[b], Zamir Libohova[g], Lin Yang[h], Alex B. McBratney[a]

[a]*Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia*
[b]*Soil Geography and Landscape group, Wageningen University & Research, the Netherlands*
[c]*ISRIC – World Soil Information, the Netherlands*
[d]*School of Biosciences, University of Nottingham, United Kingdom*
[e]*LISAH, University of Montpellier, INRAE, IRD, Montpellier SupAgro, France*
[f]*formerly Department of Soil Science and Geology, Wageningen University & Research, the Netherlands*
[g]*US Department of Agriculture, Natural Resources Conservation Services, USA*
[h]*School of Geography and Ocean Science, Nanjing University, China*

Abstract

Pedometrics, the application of mathematical and statistical methods to the study of the distribution and genesis of soils, has broadened its scope over the past two decades. The primary focus of pedometricians has traditionally been on spatial and spatio-temporal soil inventories with numerical soil classification, geostatistical modelling of spatial variation and mapping. The rapid development of remote and proximal soil sensing as well as data-driven statistical modelling techniques have had a major impact on pedometrics over the past decades. During this time, a general demand for quantitative digital soil information for environmental modelling and management has compelled pedometricians to address other soil-related questions from a quantitative point of view: soil genesis and utility and quality of soil. While scientific progress is largely an autonomous process that is difficult to steer, research efforts could benefit from an agenda with pressing pedometric research topics. This paper defines and discusses ten recent or longstanding pedometrics challenges, with the attempt to identify knowledge gaps and suggest new concepts and methods to overcome them. The ten challenges were selected through a collaborative effort and may serve as a guidance for future pedometrics research and to foster collaboration among soil scientists. The challenges discussed in this paper are also indicators of the current understanding and state of knowledge from which progress can be measured in the future.

Keywords: Soil science, Pedology, Knowledge gaps, Research questions, Scientific agenda

Highlights

- Community-based research effort.
- Define ten challenging and unsolved pedometrics questions.
- Review and proposed agenda for each challenge.
- Guide for future pedometrics research.

*Corresponding author: Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, New South Wales, Australia

Email address: alexandre.wadoux@sydney.edu.au (Alexandre M.J-C. Wadoux)

- Foster collaboration among soil scientists to solve the challenges.

1. A research agenda for pedometrics

Pedometrics currently finds itself in the same situation as many other scientific disciplines which are experiencing a fast growth, stimulated by a rapid development in digital data availability and new tools and techniques in statistics, machine learning and information systems. Although geostatistical modelling of soil variation, and later digital soil mapping (DSM) have been the primary focus of pedometricians for nearly three decades (McBratney et al., 2019b), DSM is one facet of pedometrics only. Pedometrics also encompasses, among other endeavours the study of sampling designs and their optimization, the integration of process models with observations, the quantification of uncertainty in model output, the use of mathematical methods to elucidate variation of the soil at multiple scales and new approaches to experimental design, particularly at the landscape scale. Pedometrics as a knowledge and data provider has a key role to play in participating in a broader soil science agenda and in solving global environmental issues (Finke, 2012), such as food security, land degradation and climate change. The rapid growth of pedometrics among pedological subdisciplines observed since 20 years (Hartemink et al., 2001) and the diversity of activities with a pedometrical dimension, both inside and outside the scope of spatio-temporal soil mapping, calls for a better coordination of the pedometrics research agenda.

In a number of disciplines, researchers have identified a research agenda and work towards its realization. The twenty-three problems in mathematics formulated by David Hilbert in the early 19^{th} century (Hilbert, 1902) invigorated mathematical research for a century and stimulated the development of mathematical sub-disciplines. Following this example, researchers have formulated challenges and important questions facing their own discipline, such as in ecology (e.g. Sutherland et al., 2013), plant sciences (e.g. Grierson et al., 2011) and geology (e.g. Gudmundsson, 2013). In each of these cases, identifying important research questions has focused attention, pressing scientists to find solutions which eventually led to a leap forward in their respective discipline.

In soil science, Adewopo et al. (2014) listed 25 top-ranked priorities related to ecosystem functions, health and well-being, soil nutrients/microbial interactions, soil formation and degradation, and soil information systems. A similar exercise was made in soil physics by Jury et al. (2011). In this paper, eight research areas containing key challenges are identified and linked to potential applications in agriculture and environment. Pedometrics has also been the subject of self-reflection studies (e.g. Webster, 1994; Zhu et al., 2012; Rossiter et al., 2018) and identifications of important research topics for the foreseeable future (e.g. Burrough et al., 1994; Heuvelink & Webster, 2001; Finke, 2012; McBratney et al., 2019b).

Building on these previous efforts, Murray Lark's invited talk at the "Pedometrics 2017" conference presented three important pedometrics questions for the next 25 years. This was followed by Gerard Heuvelink's proposal in the Pedometron newsletter (Heuvelink, 2019), of having the pedometrics community jointly define a list of ten key pedometrics problems. Responses by pedometricians were published in the next Pedometron issue (Bouma et al., 2019a). A community initiative was then initiated to identify ten key challenges for pedometrics research, with partial solutions to solve them. Contributing scientists were invited to define the relevance of each challenge followed by a short review and a proposed agenda. The result of these activities provides an overview of pedometrics research on a much broader scope than just DSM, and may serve as a list to be regularly reassessed and from which progress can be measured.

In the following, we address the ten challenges individually, each time explaining why the challenges matters, providing a review of current status, and defining a research agenda by listing key research questions. The challenges are grouped into three main sections: 1. How to better understand soil formation? 2. How to improve methods to obtain relevant soil data? 3. How to improve our ability to address demands by soil users? Finally, a conclusion is provided in Section 3.

2. The challenges

2.1. How to better understand soil formation?

Challenge 1: Can we produce quantitative models of the complex short and long-term processes of soil formation in the landscape which are predictive of the spatio-temporal variation of soil properties?

Why the challenge matters. Pedologists and field practitioners usually explain the soil spatial variation with hypotheses on the processes of soil formation. Digital soil mapping, conversely, make use of quantitative (numerical) techniques to express the soil pattern over an area. The development of DSM was in part driven by the need for quantitative and spatially exhaustive information on soil properties and classes. However, soil maps only provide information on the current state of the soil. How the soil pattern has evolved over time, and how it will change in the future, cannot be readily understood from the soil map, and interpretation is speculative at best. To model the soil as the product of multiple and interacting processes, we need quantitative models of soil formation that can predict the current state of the soil system from time series information about soil forming factors, and serve as input to model changes in soil under different scenarios of human intervention, land degradation and climate change.

Review. Many modellers and soil scientists consider the spatial variation of soil properties in the landscape using empirical, statistical and machine learning models (Scull et al., 2003). These models are based on the correlation among soil observations and soil forming factors (e.g. climate or organisms/vegetation). Pedometricians have since long recognized the limitation of this approach: i) it is of limited use in complex terrain with few observations, and ii) it assumes a steady state equilibrium of the soil landscape. Ideally, we should use dynamic mechanistic models of the process of soil formation instead of empirical, statistical approaches to express the soil spatial variation (Ma et al., 2019b).

Soil scientists have relied on the quantification of the most important pedogenic processes occurring in a landscape for soil mapping. These processes include physical and chemical weathering; fine clay formation; clay translocation and bioturbation, among others (see also Challenge 6). Various soil-landscape evolution models that are capable to model long-term soil change based on soil forming factors (e.g. landform) have already been developed (e.g. Veldkamp et al., 2017). The LAPSUS model (Schoorl et al., 2000) and LandLab (Hobley et al., 2017; Barnhart et al., 2020) proved suitable to model rock weathering and the relocation of soil (i.e. erosion and sedimentation processes). The LORICA model includes empirical equations describing the effect of tillage, landslides and creep while also modelling soil profile development over time and thus quantifying soil genesis (Temme & Vanwalleghem, 2016). Other examples of models are DNDC (Giltrap et al., 2010) for nutrient cycling, DEMENT (Tompkins et al., 1968) for biological activities, and WaSim-ETH (Schulla & Jasper, 2007) for water cycling, among others (Vereecken et al., 2016).

The soil spatial variation arising from short-term variation in soil processes (e.g. soil carbon change, aggregation, compaction) due to human intervention, such as leaching, acidification, erosion and compaction have also been modelled. The use of these models for mapping has been illustrated by Défossez et al. (2003) and Cerri et al. (2007).

Proposal and agenda. Few attempts have been made to explain the spatial variation of the soil with models that account for the co-evolution of a large number of soil processes, interacting at both short- and long-term time scales. There is a need for quantitative models which are robust but which encompass sufficient processes to explain the variation of soil in the landscape. There is a need to collaborate with geomorphologists, soil biologists hydrologists and physicists to develop quantitative pedological models not only for mapping but also for predicting trends of the spatial and temporal variation of soil processes and soil properties. We need to answer the following questions:

- Will we ever be able to obtain accurate information about the past state of all soil forming factors, and by extension to the future state of these factors?
- Can we make evolutionary models that account for the uncertain information of past soil forming factors?
- Can we find effective ways in which process understanding can be combined with statistical modelling so that we can use mechanistic knowledge but also handle the inevitable uncertainty in obtaining paste and future soil conditions?
- Can we make models that account for both short- and long-term soil variation in the landscape?
- Should we account for the evolutionary pathways in soil classifications, for example, to classify soils based on either origin (genetic) or only on overall similarity regardless of their evolutionary relation (phenetic), as is often done in practice?

Challenge 2: Can we develop a quantitative and numerical global soil classification that unifies the existing systems and enables transfer between them?

Why the challenge matters. Soil classifications have been derived for more than a century to create conceptual models of complex soil processes based on observations. Numerous regional soil classification systems have been developed, each fulfilling a specific function and adapted to the regional soil characteristics. With the current growing demand for global soil information, soil scientists are confronted with the difficulty to successfully convey a unified message between the many disjoint soil classification systems. This is largely due to a lack of communication, synchronization, and transfer between regional soil classification systems around the world. Current regional soil classifications are too narrowed to be used at a global level and are not accepted by all countries or organizations. Global soil classification systems exist, but are deficient for many areas of the world and for local use. An effective communication of soil information requires a unified global soil classification that provides an objective (unbiased) basis to regional soil classifications and enables synchronization between soil taxonomic systems around the world.

Review. Taxonomic systems have traditionally been driven by national soil surveys, resulting in a myriad of independent classification schemes. In France, for example, soil classes are defined from morpho-genetic soil characteristics (Baize & Girard, 2009) while in the United States (USDA-SCS, 1986) soil classification focuses on quantifiable and observable soil properties and processes. Examples of country-specific and local soil classifications schemes are found in Australia (Isbell, 2016), Brazil (dos Santos et al., 2018), New-Zealand (Hewitt, 2010), China (Xi, 1998), Zimbabwe (Nyamapfene, 1991), and Germany (Mückenhausen & Vogel, 1962).

The World Reference Base (WRB, IUSS Working Group, 2015) and USDA Soil Taxonomy (ST, Soil Science Division Staff, 2017) are considered international soil classifications. Despite their similarities, a direct linkage between the two systems is complex. The WRB was sought to harmonize ongoing soil classifications, by collective agreement on key grouping. These grouping are quantitatively defined, from properties and morphological characteristics measurable in the field as an expression of pedogenesis. The system contains two hierarchical levels (reference soil groups and uniquely defined qualifiers for specific soil characteristics) and does not consider climate or pedogenesis as differentiating criteria in the classification. The USDA Soil Taxonomy contains six levels (order, sub-order, great group, sub-group and family) and uses soil properties and climate in classifying soils as expressions of pedogenesis.

Research has shown that despite the interest to use both classifications as an international standard, problems remain with various examples in the literature describing their shortcomings in local applications, for example in Australia (Morand, 2013) and Benin (Diekkrüger et al., 2015). Attempts have been made to harmonize soil classification systems (e.g. Michéli et al., 2016), for example through communication between systems (Simonson, 1959; Bockheim & Gennadiyev, 2000) and standardization using taxonomic distance (Verheyen et al., 2001; Carré & Jacobson, 2009). Numerical methods allow for the quantification of the similarities between soil profile descriptions at different levels of taxonomy using membership functions (e.g. McBratney & De Gruijter, 1992), soil profile descriptions (e.g. Hughes et al., 2017a,b, 2018a,b) and between genetic horizons using probabilistic functions (e.g. Beaudette et al., 2016). Soil classifications have been evaluated in the multivariate space of the soil data, by using data dimensionality reduction and by calculating centroids for recognizing classes in any system (e.g. Hughes et al., 2017b, 2018a).

Proposal and agenda. Pedometricians are familiar with many of the numerical methods for soil data analysis, such as dimensionality reduction and computation of taxonomic distances. These numerical methods could serve as a basis for building a system that facilitates translation between the existing taxonomic systems, and eventually lead to an international unifying global soil classification by which all existing systems could subscribe. For achieving a global soil classification, we propose a step-wise approach that need to be tackled with soil classification experts: i) evaluate existing soil classification systems to understand and establish taxonomic relationships of differentiated soil taxa and levels; ii) develop numerical algorithms and allocation procedures for standardizing, linking harmonizing and transferring similar soil pedons from different classification systems; iii) create a global reference taxonomic database populated in taxonomic space; iv) develop taxonomically significant and harmonized soil property and spectra databases. This approach needs to address the following specific questions:

- How can a new observation (a set of data on a particular soil profile) be allocated most efficiently and consistently to a class in the system?
- How to incorporate the time dimension in soil classification?
- How to use modern proximal soil sensing techniques (estimating multiple soil properties) or image analysis (digital soil morphometrics) in conjunction with the soil classification system for near real-time soil classification?
- Can we develop pedogenetic models to classify soil profiles and properties at all points in space-time? This would allow to assess how these classes evolve.

Challenge 3: In what ways can we use data-driven models to learn about pedological processes?

Why the challenge matters. Pedometricians make extensive use of data-driven models, i.e. empirical, statistical and machine learning models, to predict soil classes and properties. These models are particularly useful for their exploratory and predictive capabilities, and have demonstrated to be very successful in situations where the soil data and the underlying soil forming processes are too complex to be modelled mechanistically (Heuvelink & Webster, 2001). In recent years, popularization of complex statistical and algorithmic tools, which increasingly engage machine learning, have steered pedometricians towards use of these models almost exclusively for prediction. But pedometricians should not only be concerned with making good predictions (Wadoux et al., 2020). They should also use models to increase understanding of soil forming processes. We need techniques for interpretation of complex statistical learning models that can reveal the system functioning and expose the role and importance of each driver of soil variation for a region, sub-region, spatial location and depth, as well as in time.

Review. Modelling the soil with the aim to learn about soil processes using empirical, statistical and machine learning models is not new to soil science. Yaalon (1975), for example, recommended statistical and stochastic modelling when little is known about the soil system under study. The functional relationships of a calibrated empirical model can suggest hypotheses or causal mechanisms to the soil scientist. The functional relationships found among variables and their strength is usually characterized by the variable importance of the environmental predictors, such as through the standardized regression coefficients of a multiple linear regression, or other variable importance statistics, such as the mean decrease in impurity (or gini importance) for tree-based models and Garson's algorithm for simple artificial neural network models (Olden & Jackson, 2002). Wadoux & McBratney (2021) contended that the spatial pattern predicted by the data-driven model and the correlation found among the covariates are an opportunity to formulate hypotheses. In digital soil mapping our selection of covariates is driven by soil forming factors and when reporting variable importance we often group them by these factors (e.g. Viscarra Rossel et al., 2015). Pedometricians have since long complemented variable importance statistics with analysis and interpretation of the spatial structure of the soil properties or classes, as is done in geostatistics with analysis of the variogram parameters.

The relative effect of individual or interacting environmental variables can also be quantified by sensitivity analysis. Sensitivity analysis shows how the model output varies under a change of its inputs (e.g. Paul et al., 2003; McNicol et al., 2019). Several frameworks for model interpretation are based on this concept. An alternative is to calibrate a surrogate (simpler) model to approximate the prediction of a more complex model, as is done in LIME (Ribeiro et al., 2016). All these techniques also work for machine learning models (e.g. Pham et al., 2019). However, there has been limited interest to understand how such methods can serve pedological interpretation of complex statistical models of soil variation. Some recent examples from soil microbiology (e.g. DiMucci et al., 2018) soil and ecosystem respiration (e.g. Ilie et al., 2017) and pedology (e.g. Vos et al., 2019) show that this is a challenging but promising line of future research.

Proposal and agenda. Machine learning has so far mainly been used to improve soil predictions, but we should also explore its potential to assist in learning about soil processes. Pedometricians have already implemented some of the available methods for interpretation of complex empirical models, but there is a lack of understanding on how these methods can help us understand and explain why soil varies the way it does. We need to integrate these methods into a framework that formalizes pedological interpretation of complex statistical models. Some important research questions are:

- How can we better interpret the rule sets generated by methods for prediction which do not use explicit statistical models (e.g. machine learning) to gain insight into soil processes?
- Can we use machine learning models to study and interpret causal factors of soil variation in time?
- How can sensitivity analysis of complex machine learning models help us understand soil processes better?

- Can we delineate carefully the study areas so that they match to well-defined pedological systems, whose variations and relationship with landscape features are readily interpretable?
- How does the relation between model inputs and outputs of complex statistical models vary spatially and in time?

2.2. How to improve methods to obtain relevant soil data?

Challenge 4: Can we measure soil properties more efficiently?

Why the challenge matters. Soil scientists have recently witnessed a growing demand for quantitative soil information for feeding agro-environmental models and decision-making tools. Furthermore, the lack of such information is actually the major limiting factor of the production of maps of soil properties. Current routine laboratory techniques and field observation protocols will not be able to fulfill this demand because most soil properties are too costly to measure at the appropriate spatial support (e.g. available water capacity) while others are not easily or readily measurable (e.g. soil structure), or not measured at the correct support for the intended use (Challenge 5). To model the soil and derive accurate products that match user's demand, we need new technologies and methods that allow us to better measure soil properties at the right support, more efficiently and more precisely.

Review. Several ways of extending the acquisition of quantitative soil data have been explored in the past: soil sensing, pedotransfer functions, the translation of qualitative soil information and citizen science.

Numerous successful attempts have been made of using sensing techniques for estimating basic physical and chemical soil properties, both at the local level by laboratory techniques (Viscarra Rossel & McBratney, 2008), at field level by *in situ* technologies (Adamchuk et al., 2004; Kuang et al., 2012) and at larger extent using remote sensing techniques (e.g. Lagacherie & Gomez, 2018). Sensing techniques have also been applied to estimate available water capacity using remote sensing images and inverse modelling (e.g. Sreelash et al., 2017), geophysical data (e.g. Jiang et al., 2007) and vegetation indices (e.g. Coulouma et al., 2020). Other soil properties like soil surface hydraulic conductivity (e.g. Goldshleger et al., 2002) and soil structural stability (e.g. Franz et al., 2018) have also been considered by combination of sensing techniques.

Pedotransfer functions (Bouma, 1989) have been developed to relate simple soil properties (e.g. texture) to more complex soil characteristics such as hydraulic conductivity. Pedotransfer functions were later included in a Soil Inference System (SINFERS, McBratney et al., 2002) to populate digital soil databases. In a SINFERS, the uncertainty of the simple soil properties is propagated in the estimation of the complex soil characteristics. This requires methods for measurement error quantification using expert elicitation or quantitative approach (Brown & Heuvelink, 2006).

Research has also attempted to translate traditional qualitative soil information into a quantitative one (see also Challenge 2). Lagacherie (2005), has translated descriptions of soil profiles and soil classes that were coded in soil databases into a numerical format using possibility theory (fuzzy sets). Furey et al. (2019) have translated description of soil profiles in text form into a numerical format using natural language processing.

Pedometricians have recognized that citizens could assist in producing soil data using a network of non-specialist volunteers. Rossiter et al. (2015) details some examples for digital soil mapping. Examples of such initiatives are mySoil for collecting soil data using smartphones, the OPAL Soil and Earthworm Survey (Bone et al., 2012) and the tea bag index to assess decomposition rate and litter stabilisation (Keuskamp et al., 2013).

The development of efficient sampling strategies that may reduce measurement costs independent of the measurement technique has received much attention from pedometricians, such as for digital soil mapping using geostatistical (e.g. Brus & Heuvelink, 2007) and machine learning approaches (e.g. Wadoux et al., 2019; Lagacherie et al., 2020). Sampling for modelling several soil properties has also been considered (e.g. by Vašát et al., 2010).

Proposal and agenda. While Pedometricians have developed and implemented some techniques for efficient soil data collection, we need to develop new technologies and develop standardized methods to collect more soil data, more precisely and more efficiently than with present-day methods and techniques. Important research questions are:

- Can we design a protocol to make traditional description of soil profile more quantitative involving faster and cheaper instruments and techniques to measure *in-situ* soil properties?
- Can we estimate measurement error of various measurement methods, and propagate the uncertainty to applications and functions.
- Could we consider sensing of classical soil attributes and sensing of soil functions to be conducted together within a coherent soil sensing inference system that would involve the synergistic and multi-scale coupling of several sensing techniques?
- To which degree can information from citizen science or farmers be used for applications and soil monitoring?
- Can we develop new sampling strategies, mapping models and decision making tools that can synergistically use and consider different sources of soil data (qualitative soil profiles descriptions, soil maps, soil sensing data, citizen data) that have different precision, geographical supports, resolutions and extents?

Challenge 5: Can we develop workable techniques to derive predictions of soil characteristics at scales appropriate for modelling and decision making, by up- and downscaling observations in 3D space and time? These techniques should account for complex non-linearities and interactions, and quantify the associated prediction uncertainty.

Why the challenge matters. Soil properties and processes vary at spatial scales from the atomic to the global, and in time. The factors causing variation at one scale may be different from those at other scales, which has implications for modelling. We often measure soil properties in the field at much finer support scales than required by users and modellers, while remote sensing may yield observations of soil properties and environmental soil forming factors at too coarse resolution and support. To synthesize different sources of information and deliver appropriate soil information to users, we need up- and downscaling techniques for soil properties and processes that can accomplish a change of support in all dimensions (2D, vertical, and time).

Review. The concept of scale has multiple meanings in soil science, but here we consider it to be the triplet of extent, resolution (i.e. grain) and support (Western & Blöschl, 1999; Bierkens et al., 2000). Of these three the support, i.e. the area, volume or time period over which an observation or prediction is made, is the most challenging from a pedometrics point of view. The support scale is important because it determines the character and degree of soil variation (Burrough, 1983) and consequently influences the assumptions made in mechanistic and statistical modelling of the soil (Heuvelink, 1998). For instance, variation in soil carbon at the farm-scale is largely driven by texture and micro-organisms while these factors are less important at continental scale, where climate and vegetation are dominant (Wiesmeier et al., 2019).

Geostatistics has since long developed block kriging to aggregate from fine to coarse support (e.g. Burgess & Webster, 1980b), while spatial disaggregation using area-to-point kriging developed much later (e.g. Kerry et al., 2012). Techniques to spatially disaggregate polygon soil type maps were also developed (e.g. Odgers et al., 2014). Non-linear aggregation can be achieved using spatial stochastic simulation (Goovaerts, 1999). Handling vertical change of support may be solved using equal-area splines (Bishop et al., 1999), although this is not the ultimate solution because it ignores the uncertainty induced by this process (Orton et al., 2016).

Research has also shown that the optimal support of soil forming factors used in digital soil mapping (e.g. elevation data) is often greater than the support of the soil observations and predictions (e.g. Behrens et al., 2010).

Many modellers and users consider fine-scale soil variation a nuisance and eliminate it using block kriging or composite sampling. But some modelling studies need information at scales even finer than a soil sample to do better justice to the underlying mechanistic processes. Examples are water flow (Šimůnek et al., 2003), nitrification (Boyer et al., 2006) and soil organic matter dynamics (Six et al., 2004). Such processes are often characterized by partial differential equations that can only be realistically represented using very fine model scales. As a result, scaling techniques (Pachepsky & Hill, 2017) are required to apply these models at landscape, regional and continental scales.

Proposal and agenda. Pedometricians have tackled some of the scaling issues in soil science, but there are many more that await us, some of which are best solved in collaboration with soil physicists, chemists and biologists:

- Can we improve our understanding on what defines the optimal scale of explanatory environmental variables used in soil prediction models?
- How to account for the vertical support in soil modelling and digital soil mapping? We need solutions that account for the depth and thickness of soil samples and that recognize that soil vertical variation is often greatest at the boundaries of soil horizons.
- How to deal with temporal scale issues in modelling soil change and space-time soil variation?
- Can we contribute to the development of scaling techniques to support the use of fine-scale mechanistic models at landscape, regional and continental scales? And can we do so such that associated uncertainties are quantified?
- Can we develop statistical validation techniques for models that make predictions at a much larger support than that of the validation measurements?

Challenge 6: Can we incorporate mechanistic pedological knowledge in digital soil mapping?

Why the challenge matters. Most digital soil mapping models are to a high degree empirical, and this is only increasing, now that we rely heavily on statistical models and machine learning algorithms for mapping. The models generated in this way depend almost exclusively on observations, and make no or little use of existing pedological knowledge on soil processes. In many situations, however, mechanistic knowledge on soil processes and the pedologist tacit knowledge of an area learned through experience is available and should be used. Making use of this knowledge in DSM has two important advantages: i) interpretation of the DSM model and its outputs increases insight into dominant soil processes in the area of interest (see also Challenge 3), and ii) it supports extrapolation of the model to areas and times with similar soil characteristics and conditions. Existing techniques for including pedological knowledge into DSM models are very limited. Pedometricians need to develop techniques and models that can include existing knowledge from different sources, including subjective and qualitative soil information.

Review. We usually distinguish between mechanistic and functional (empirical) models to map the soil (Hoosbeek & Bryant, 1992). Pedometricians have extensively used static, functional models that make use of empirical relationships between soil characteristics and environmental covariates (Heuvelink & Webster, 2001; Ma et al., 2019a). These approaches, despite being widely applied and often providing accurate predictions, do not represent the underlying dynamic processes because their model structures at best are a poor representation of the true physical, chemical and biological mechanisms. Mechanistic models for soil mapping, conversely, have a structure based on mechanisms derived from our knowledge on physical processes and soil chemical and biological reactions. Examples of mechanistic models for mapping soil variation in space and time are soil erosion/deposition models (e.g. Milne, 1936; Minasny & McBratney, 2006; Cohen et al., 2010; Stumpf et al., 2017), soil-landscape evolution (e.g. Temme & Vanwalleghem, 2016), soil thickness (e.g. Follain et al., 2006) and soil carbon and nitrogen models (e.g. Wang et al., 2020; Hendriks et al., in Press), see also Challenge 1. The main problems with these models are that they require adequate mechanistic understanding of major soil processes, need a lot of input data that are often not available, have many parameters that are difficult to infer, are computationally challenging, and typically do not quantify prediction uncertainty. This last problem may be tackled using state-space modelling and the space-time Kalman filter (e.g. Heuvelink et al., 2006; Lin et al., 2017).

Walter et al. (2006) detailed five ways of incorporating mechanistic knowledge into empirical DSM models, the most basic of which is using expert knowledge to choose covariates that are proxies of the soil forming factors in regression, as done for example in Nussbaum et al. (2018). Hybrid solutions that take incorporation of mechanistic knowledge a step further are structural equation modelling (SEM, Angelini et al., 2016) and Bayesian belief networks (BNN, Mayr et al., 2010; Taalab et al., 2015). SEM converts a conceptual soil-landscape model into a statistically explicit model that does more justice to underlying mechanistic processes than conventional regression models. Bayesian belief networks formulate expert knowledge into rules that define relations between soil properties and between the soil and the environment. Other solutions are to include constraints and correlations in the empirical model structure, for example by constraining the joint prediction of soil properties to realistic co-occurrences (pH and carbonate content) or to clay, silt and sand contents summing to 100 percent (e.g. Wadoux, 2019).

Proposal and agenda. While pedometricians have developed some solutions to include mechanistic knowledge in DSM, this is only the beginning of a new subdomain within pedometrics. We need a tangible framework that formalizes the integration of mechanistic knowledge in DSM, for different soil properties and processes, and different existing types of knowledge (e.g. tacit, conceptual). Some important research questions for future research are:

- Can we use expert elicitation and/or Bayesian statistics to extract pedological knowledge from soil surveyors such that they provide useful input and/or boundary conditions to DSM?
- Can we extend SEM from static to dynamic model structures, whilst avoiding some of the five major impediments of mechanistic models?
- Can we help address and solve some of the limitations of mechanistic models by finding a workable compromise between mechanistic and empirical models?
- Can we steer machine learning DSM models into directions that make sense mechanistically by incorporating mechanistic knowledge in the loss function that is minimized?
- How to incorporate uncertainty in mechanistic knowledge when using it to support DSM?

2.3. How to improve our ability to address demands by soil users?

Challenge 7: How to recognize, quantify and map soil functionality?

Why the challenge matters. As soil scientists, pedometricians spend a lot of effort on modelling and mapping soil types and basic soil properties. However, most end-users perceive soils through how they behave and what they do for satisfying human needs, which has been expressed in recent years by various sets of soil functions related to ecosystem services (Dominati et al., 2014). Therefore, we urgently need to move towards modelling and mapping these soil functions. Beyond fulfilling a strong end-users demand, this evolution is necessary for enabling soil science to participate in interdisciplinary studies on the key issues of sustainable development (Bouma, 2014). Furthermore, this represents an opportunity to feed models and validation datasets with new data such as citizen observations of soil functions or soil function sensing outputs.

Review. Converting soil maps into mapping products that better match the end-user's demand (e.g. land evaluation maps, irrigation requirement maps) has long been practiced by many soil survey staffs (e.g. FAO, 1977; McKenzie et al., 2008; Soil Science Division Staff, 2017). This tradition has been prolongated and amplified recently with the need to link soils to ecosystem services (e.g. Adhikari & Hartemink, 2016). Quantitative soil functions assessments (SFA) have been proposed by various authors (see a review by Greiner et al., 2017).

One can distinguish two approaches to quantify soil functions. The first uses empirical rules combining basic soil properties while the second uses biophysical model-based approaches that better account for the dynamics of the soil environment (climate, land use, agricultural practices) but are by far the most data-demanding and time-consuming. Another important distinction is made between SFAs that consider the intrinsic potential of soil to fulfill various functions and SFAs that evaluate its actual state for doing so (Vogel et al., 2019; McBratney et al., 2019a). This clearly corresponds to different scales (in time and space) that should be considered (see also Challenge 7).

SFA has been often coupled with existing soil databases to produce maps of soil functions. However, the lack of fine resolution soil data is acknowledged as a severe limiting factor (Greiner et al., 2017). Although the principle of using digital soil mapping approaches for mapping soil functions has been proposed and evaluated by previous works (e.g. Carré et al., 2007), few applications (e.g. Kidd et al., 2012; Greiner et al., 2017, 2018) of this principle have been undertaken. Recent research (e.g. Styc & Lagacherie, 2019) has investigated the best inference strategy between two choices, i) assessing the soil function at point support and then mapping (i.e. first calculate then interpolate), and ii) mapping of soil properties and then inserting the soil property maps into a soil function (i.e. first interpolate then calculate). The best strategy depends upon the correlation between soil properties, which can be addressed by moving from univariate to multivariate mapping of basic soil properties (e.g. Wadoux, 2019). DSM is already able to provide uncertainty assessment that could be propagated to the soil function maps, which combined with an increase in spatial resolution, will help in decision making (Challenge 9).

Proposal and agenda. To move toward the operational quantification and mapping of soil functions, useful for different users and scales, the following questions need to be answered:

- Can we develop new SFA approaches based on bio-physical models that account for spatial and temporal variation in the soil environment? This requires a critical evaluation of soil models, organizing such models in user-friendly modelling platforms (e.g. Lafolie et al., 2013) and developing meta-modeling approaches to simplify the further mapping.
- Can we develop specifications for the production of standardized soil function maps at field/farm, regional, national and global scales, as was done in the GlobalSoilMap project for

the production of basic soil property maps? Considering the target scales, these functions should deal only with intrinsic potentials rather than with actual states.

- Can we develop local SFA approaches that may be co-built with end-users for matching their own perceptions of soil functions?
- How to build a versatile SFA assessment tool that can be parametrized through an interactive user-friendly interface? This tool should match the users' knowledge of the actual state of the soil functions and of their variation in space in the area under study.

Challenge 8: Can we find ways to connect pedodiversity to soil biodiversity, and translate the connections to relevant soil services and soil management practices?

Why the challenge matters. Soil macro- and microbial diversity is important for sustainable land management and flexible response to environmental change. When the diversity is reduced the soil and the services it provides are less resilient against disturbance and stress. Pedodiversity, the inventory of the variety of pedological entities such as pedotaxa, pedogenetic horizon and soil properties, as well as their spatial and temporal patterns influences soil biodiversity and species richness by providing a heterogeneous habitat and supporting plant, flora and fauna biodiversity. We need techniques to connect pedodiversity to soil biodiversity for varying spatial scales and different habitat types and link these to scenarios of soil change (e.g. loss of soil diversity) and maintenance of soil ecosystem functions.

Review. The concept of diversity contains two components: the richness, i.e. the variety of species of soil types on a sample and the evenness, i.e. the relative number of each individual soil type or specie (Huston, 1994; Ibañez et al., 1995). Most measures of diversity incorporate both components into a single index, or tend to have more emphasis on one of the two. Pedologists have measured and quantified pedodiversity with tools similar to those used by biologists and ecologists to measure biodiversity (Ibanez et al., 2012): species richness indices, indices based on the proportional abundances of species (e.g. the Shannon index) and species abundance models. These indices have been used to quantify pedodiversity for local (e.g. Petersen et al., 2010) or larger areas such as continents (e.g. Ibáñez et al., 1998) or the world (e.g. Ibáñez & Feoli, 2013). McBratney & Minasny (2007) included taxonomic distance in the calculation of diversity indices.

Pedologists have long ago found a positive correlation between pedodiversity and area size (see also Beckett & Bie, 1978; Guo et al., 2003; Phillips & Marion, 2005), and that some areas are less diverse than others. For example, at the global scale and based on the major soil groups, some areas such as South-East Asia, Oceania or Middle-East are less diverse than North America (McBratney et al., 2000).

Several studies have found that the patterns of soil biodiversity and pedodiversity have strong similarities. The distribution pattern of soil micro- and macro-organisms is not random (Ranjard et al., 2013; Karimi et al., 2020), it is a reflection of local and regional environmental heterogeneity (Pino et al., 2019). Pedodiversity greatly influences soil biodiversity (Ibanez et al., 2012), and can be used as a surrogate indicator of soil biodiversity.

Research has attempted to link detected communities in the soil and their associated changes to relevant soil services and soil management practices (e.g. Bardgett, 2005). For example, cover crops and fertilizer application impact the interaction among the species in microbial communities but patterns may vary in different soil habitats (O'Brien et al., 2016; Zheng et al., 2018). Also, clear linkages between soil microbial communities and forest restoration state were found (e.g. by Sun et al., 2017). Moreover, the soil micro- and macrofauna are important for various soil services and are strongly linked to management practices (Hu et al., 1995). For example, earthworm abundance and species have a large effect on the carbon cycle (Angst et al., 2019). Another example is found in Van Den Hoogen et al. (2019), where the regional variations of nematodes abundance at the global scale is shown to provide insights into local patterns of soil fertility and functioning.

Proposal and agenda. We should explore the link between soil biota species richness and edaphic conditions and identify key soil properties that link them. Pedometricians should also take on the task of developing quantitative ways to relate keystone species abundance and distribution to major soil properties and taxa for specific environmental conditions. This can serve as a basis for land management and ecosystem restoration. Pedometricians should become familiar with quantitative methods for hyper-variate data of soil biodiversity analysis, similar to how they became familiar with spectroscopy a decade ago. This approach needs to address the following specific questions:

- Can we develop a common language with ecologists to communicate about diversity?
- How are alpha, beta and gamma soil biodiversity influenced by pedodiversity? Are spatial patterns of soil entities and micro-organisms similar and driven by the same factors? Is there a co-evolution of patterns in time?
- What assemblage of soil properties (pedodiversity) drives the alpha diversity of the biota?
- Can pedometrics support the analysis of microbial properties at the microscale in soil by combining the traditional tools used at the microscale and those used to study the soil spatial variability?
- Can we use proximal soil sensing for estimating microscale biodiversity at a site? That is, can we use spectroscopy to estimate a number of soil properties, which are the basis for estimating soil pedodiversity and biodiversity at a site?
- Can we link pedodiversity and soil biodiversity to soil management?

Challenge 9: Can we find ways to express the uncertainty of predictions of soil properties or class allocations which are meaningful to the users of those predictions? These should allow both the quantitative assessment of risk associated with resulting decisions, and an assessment of the value of further information which might be collected.

Why the challenge matters. Arguably the last measure of uncertainty in soil information developed and applied by soil scientists, and routinely applied by the end-user of that information, was the "purity", of the conventional soil map, the probability that the soil class designated by the map at a random location in the area mapped would be found there in the field. The application of geostatistics in soil science was, in part, motivated by the need for robust measures of uncertainty of quantitative predictions (Webster & Oliver, 2007; Webster, 2015), and the statistical or machine learning methods now used for this purpose all generate uncertainty measures. However, these are rarely linked directly to the decision-making process for which soil information is required. That task is left to the user. This is a failure to communicate the uncertainty effectively. We contend that pedometricians must engage actively with the decision-making processes of end-users if the uncertainty in soil information is to be communicated, and this is essential if the information is to be used.

In general, users of soil information will find it hard to assess the significance of abstract measures of uncertainty (such as variances). Measures of uncertainty which relate directly to the soil behaviour under consideration (e.g. a range of possible outcomes from an intervention), are more likely to be understood. Such approaches are used, for example, in reports from the Intergovernmental Panel on Climate Change (Mastrandrea & Mach, 2011). Additional challenges may be expected when dealing with "wicked" problems that have no single solutions, but, rather, a set of alternative options from which a choice has to be made.

Review. Soil map purity (above) has been used as an uncertainty measure in studies of survey methodology (e.g. Burrough et al., 1971) but has also been used to specify standards in contracts for soil survey and in quality control (Western, 1978). The uncertainty in quantitative prediction of soil properties can be characterized by the mean (expected) square error of the prediction (e.g. kriging variance), or other statistics derived from the prediction distribution in a model-based setting, be this frequentist or Bayesian (Diggle & Ribeiro, 2007). Non-linear geostatistical methods have been applied to map the conditional probability that a soil property falls above or below a management or regulatory threshold (e.g. Webster, 1991). When predictions are used as input to process models the error propagates to model outputs (Heuvelink, 2018), which again may be characterized by a statistical distribution.

Pedometricians have recognized that uncertainty measures can be mapped alongside predictions (Burgess & Webster, 1980a; Vaysse & Lagacherie, 2017), and that the kriging variance can be computed for prospective survey designs to guide decision making (McBratney et al., 1981). However, there has been limited attention to whether such outputs serve the needs of end-users. How should I interpret the uncertainty measure associated with a prediction on which I base a decision? Can I make the decision robust, or should I postpone it until more data are collected, and how much effort is it rational to devote to further sampling? Such questions fall under decision theory.

Decision theory has received limited attention from soil scientists. The consequences of alternative decisions under possible future outcomes can be modelled by a loss function (Ramsey et al., 2002; Lark & Knights, 2015) in combination with a statistical model of the key input variables to characterize their uncertainty. This allows the expected costs of those decisions to be compared (Viscarra Rossel et al., 2001). Decision analysis allows the value of prospective information to be estimated, including imperfect information. This has been illustrated for conventional soil surveys (e.g. Giasson et al., 2000). We propose this is an area with much scope for further development.

Proposal and agenda. We propose a research programme focused on the analysis of decisions made by users of soil information and on how the uncertainties in that information can be assimilated into these processes. Some specific questions to be addressed are:

- Can we analyse the decision processes of key users of soil information (e.g. in precision agriculture or land remediation) in sufficient detail to quantify the impacts of uncertainty in that information?
- How can we analyse decision processes which are more open-ended (e.g. policy decisions about "wicked" problems, Bouma & McBratney, 2013)?
- How does the value of soil information, and the value of imperfect information, differ with the support scale, and the extent of the studied region, and how does it relate to the costs of providing information at these scales and over this extent?
- What are the typical diminishing returns to investment in soil information?

Challenge 10: How to quantify soil contributions to ecosystem services with a framework enabling both local and regional soil management?

Why the challenge matters. For the world's growing population to develop sustainably, we must address six key existential environmental challenges: food, water and energy security, climate change abatement, biodiversity protection and human health, as defined by the UN Sustainable

Development Goals. Such challenges relate to contributions by soils to interdisciplinary ecosystem services (e.g. Veerman et al., 2020). Soils play a pivotal role in each of these but this is not reflected directly in sustainable management on enterprise level and on development policy. Efforts to recognize and include the role of soil in the development of policies is diffuse due to the current lack of targets and indicators that do justice to the essential input that soils can have (Bouma et al., 2019b). The poverty of current data and modelling render significant uncertainty in the assessments of threats to soil functioning across the globe (Bonfante et al., 2020). There is lack of a comprehensive systematic soil monitoring system, and an overall quantitative framework for assessing soil contributions to ecosystem services.

Review. Soil contributions to ecosystem services may be framed withing the UN Sustainable Development Goals (SDG). Soils have been recognized to make important contributions to ecosystem services leading to at least: SDG 2 (food), SDG 6 (water), SDG 13 (climate) and SDG 15 (biodiversity) (Bouma, 2020a). Agronomists, hydrologists, climatologists and ecologists have published widely in the four SDG areas mentioned above, each of which qualifying as focal areas for their disciplines.

Empirical land evaluation schemes (Rossiter, 1996), developed in the 1970's (e.g. the FAO framework in 1976), and dynamic simulation models for the soil-water-atmosphere-plant system (Feddes et al., 1978) are used to quantify soil contributions to ecosystem services (e.g. by Hack-ten Broeke et al., 2019). Bonfante et al. (2020) have shown that application of soil-water-atmosphere-plant simulation models can define soil contributions to ecosystem services for Italian conditions as well as quantitative expressions for soil quality and soil health but more testing is needed in other areas. The abstract notion of cooperation in interdisciplinary projects has been specified in these simulation programs by asking which soil data can be provided to adequately represent dynamic soil behavior in space and time as expressed by modeling, that also requires input by agronomists, hydrologists, climatologists and ecologists (see also Challenge 7).

The soil security concept (McBratney et al., 2014; Field et al., 2016) has been developed as an overall framework to explicitly connect soil with the other global existential challenges. It has five dimensions, i) capability, ii) condition, iii) capital, iv) connectivity and v) codification. When integrated, these dimensions reveal an overall biophysical, economic and sociological understanding of soil for use for on-the-ground management and integration into policy and legal frameworks (Bouma, 2020b).

Proposal and agenda. All the subjects detailed in the previous paragraphs have successfully been studied by the pedometrics community for several years and they need to be, and can be, inserted more effectively in the societal and political discourse. To do so, there is need for quantitative linkages to and between the appropriate SDGs, the five dimensions of soil security, soil functions (Challenge 7), and the global existential challenges (Bouma, 2014). Pedometricians will need to actively cooperate in interdisciplinary projects to achieve this. Some important research questions are:

- Can we build a quantitative framework connecting and assessing soil contributions to realizing the SDGs and soil security which is relevant for both local-scale management decisions and regionally or globally to support policy making?
- How to model and predict the change in soil conditions and optimize preemptive strategies to maintain or improve soil conditions?

To illustrate the potential of pedometrics, case studies should be selected and initiated that clearly demonstrate the impact of introducing quantitative soil measurements. Different approaches should be compared, showing limitations and potentials of each procedure, starting with empirical, qualitative assessments in terms of traditional land evaluation up to innovative and quantitative procedures. It is important to start with applying existing knowledge and expertise and focus new research on gaps that appear when applying existing knowledge (Veerman et al., 2020). In this way the advantages of new research can be clearly demonstrated when using the SDG as an ultimate goal (the "point at the horizon") so that the visibility and relevance of the work can be effectively illustrated, aiming at the policy arena and the public at large.

3. Conclusion and Outlook

We have defined ten challenges reflecting the current state of knowledge but also many areas of soil science where pedometrics can contribute. We hope that this exercise will aid pedometricians to align their research to several new or long-term research questions of immediate importance to soil science, on a much broader scope than just digital soil mapping. Soil science is at the interface of several questions relevant to the sustainable growth of our society for which pedometricians can bring expertise to the quantitative study of the soil. However, many of the challenges described in this paper cannot be faced by pedometricians alone and need to be addressed jointly in inter- and trans-disciplinary research efforts, i.e. also including collaborations outside soil science. Several of the ten challenges are linked and development in one challenge can support development in another.

In the past, defining key research questions and working as a community to solve them has been done with different degrees of success in various fields of science. Scientific progress is largely an autonomous process that is difficult to evaluate and even more so to steer. Nevertheless, the list of challenges presented in this paper can serve as basis from which progress and change can be assessed in the coming years.

Acknowledgements

We thank Nicolas P.A. Saby, INRAE unité Infosol in Orléans (France), for maintaining the website `http://pedometrics.org/` and for building the discussion forum around the challenges. V.L. Mulder and Z. Libohova are members of the research consortium GLADSOILMAP supported by LE STUDIUM Loire Valley Institute for Advanced Studies through its LE STUDIUM Research Consortium Program.

References

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture*, *44*, 71–91.

Adewopo, J. B., VanZomeren, C., Bhomia, R. K., Almaraz, M., Bacon, A. R., Eggleston, E., Judy, J. D., Lewis, R. W., Lusk, M., Miller, B., Moorberg, C., Snyder, E. H., & Tiedeman, M. (2014). Top-ranked priority research questions for soil science in the 21st century. *Soil Science Society of America Journal*, *78*, 337–347.

Adhikari, K., & Hartemink, A. E. (2016). Linking soils to ecosystem services-A global review. *Geoderma*, *262*, 101–111.

Angelini, M. E., Heuvelink, G. B. M., Kempen, B., & Morrás, H. J. M. (2016). Mapping the soils of an Argentine Pampas region using structural equation modelling. *Geoderma*, *281*, 102–118.

Angst, G., Mueller, C. W., Prater, I., Angst, Š., Frouz, J., Jílková, V., Peterse, F., & Nierop, K. G. J. (2019). Earthworms act as biochemical reactors to convert labile plant compounds into stabilized soil microbial necromass. *Communications Biology*, *2*, 1–7.

Baize, D., & Girard, M.-C. (2009). *Référentiel pédologique 2008*. Editions Quae, Versaille, France.

Bardgett, R. (2005). *The biology of Soil: a Community and Ecosystem Approach*. Oxford university press, Oxford, UK.

Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., & Bandaragoda, C. (2020). Landlab v2. 0: a software package for Earth surface dynamics. *Earth Surface Dynamics*, *8*, 379–379.

Beaudette, D. E., Roudier, P., & Skovlin, J. (2016). Probabilistic representation of genetic soil horizons. In A. E. Hartemink, & B. Minasny (Eds.), *Digital Soil Morphometrics* (pp. 281–293). Springer, Berlin, Germany.

Beckett, P. H. T., & Bie, S. W. (1978). *Use of soil and land-system maps to provide soil information in Australia*. Technical Report CSIRO Division of Soils Technical Paper No. 33, Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia.

Behrens, T., Zhu, A.-X., Schmidt, K., & Scholten, T. (2010). Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma*, *155*, 175–185.

Bierkens, M. F. P., Finke, P. A., & de Willigen, P. (2000). *Upscaling and Downscaling Methods for Environmental Research*. Developments in Plant and Soil Sciences. Kluwer Academic Publishing, Dordrecht, NL.

Bishop, T. F. A., McBratney, A. B., & Laslett, G. M. (1999). Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma*, *91*, 27–45.

Bockheim, J. G., & Gennadiyev, A. N. (2000). The role of soil-forming processes in the definition of taxa in Soil Taxonomy and the World Soil Reference Base. *Geoderma*, *95*, 53–72.

Bone, J., Archer, M., Barraclough, D., Eggleton, P., Flight, D., Head, M., Jones, D. T., Scheib, C., & Voulvoulis, N. (2012). Public participation in soil surveys: lessons from a pilot study in England. *Environmental Science & Technology*, *46*, 3687–3696.

Bonfante, A., Basile, A., & Bouma, J. (2020). Targeting the soil quality and soil health concepts when aiming for the United Nations Sustainable Development Goals and the EU Green Deal. *SOIL*, *5*, 1–14.

Bouma, J. (1989). Using soil survey data for quantitative land evaluation. In B. A. Stewart (Ed.), *Advances in Soil Science* (pp. 177–213). Springer, New York, USA Volume 9.

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

Bouma, J. (2020a). Contributing pedological expertise towards achieving the United Nations Sustainable Development Goals. *Geoderma*, *375*, 114508.

Bouma, J. (2020b). Soil security as a roadmap focusing soil contributions on sustainable development agendas. *Soil Security*, *1*, 100001.

Bouma, J., Libohova, Z., Lagacherie, P., & Yang, L. (2019a). The 10PM challenges. *Pedometron*, *44*, 28–36.

Bouma, J., & McBratney, A. B. (2013). Framing soils as an actor when dealing with wicked environmental problems. *Geoderma*, *200*, 130–139.

Bouma, J., Montanarella, L., & Evanylo, G. (2019b). The challenge for the soil science community to contribute to the implementation of the UN Sustainable Development Goals. *Soil Use and Management*, *35*, 538–546.

Boyer, E. W., Alexander, R. B., Parton, W. J., Li, C., Butterbach-Bahl, K., Donner, S. D., Skaggs, R. W., & Grosso, S. J. D. (2006). Modeling denitrification in terrestrial and aquatic ecosystems at regional scales. *Ecological Applications*, *16*, 2123–2142.

Hack-ten Broeke, M. J. D., Mulder, H. M., Bartholomeus, R. P., Van Dam, J. C., Holshof, G., Hoving, I. E., Walvoort, D. J. J., Heinen, M., Kroes, J. G., van Bakel, P. J. T., Supit, I., de Wit, A. J. W., & Ruijtenberg, R. (2019). Quantitative land evaluation implemented in Dutch water management. *Geoderma*, *338*, 536–545.

Brown, J. D., & Heuvelink, G. B. M. (2006). Assessing uncertainty propagation through physically based models of soil water flow and solute transport. In M. G. Anderson, & J. J. McDonnell (Eds.), *Encyclopedia of Hydrological Sciences* chapter 79. Wiley Online Library, Hoboken, USA.

Brus, D. J., & Heuvelink, G. B. M. (2007). Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, *138*, 86–95.

Burgess, T. M., & Webster, R. (1980a). Optimal interpolation and isarithmic mapping of soil properties: I The semi-variogram and punctual kriging. *Journal of Soil Science*, *31*, 315–331.

Burgess, T. M., & Webster, R. (1980b). Optimal interpolation and isarithmic mapping of soil properties: II block kriging. *Journal of Soil Science*, *31*, 333–341.

Burrough, P. A. (1983). Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation. *Journal of Soil Science*, *34*, 577–597.

Burrough, P. A., Beckett, P. H. T., & Jarvis, M. G. (1971). The relation between cost and utility in soil survey (I–III). *Journal of Soil Science*, *22*, 359–394.

Burrough, P. A., Bouma, J., & Yates, S. R. (1994). The state of the art in pedometrics. *Geoderma*, *62*, 311–326.

Carré, F., & Jacobson, M. (2009). Numerical classification of soil profile data using distance metrics. *Geoderma*, *148*, 336–345.

Carré, F., McBratney, A. B., Mayr, T., & Montanarella, L. (2007). Digital soil assessments: Beyond DSM. *Geoderma*, *142*, 69–79.

Cerri, C. E. P., Easter, M., Paustian, K., Killian, K., Coleman, K., Bernoux, M., Falloon, P., Powlson, D. S., Batjes, N. H., Milne, E., & Cerri, C. C. (2007). Predicted soil organic carbon stocks and changes in the Brazilian Amazon between 2000 and 2030. *Agriculture, Ecosystems & Environment*, *122*, 58–72.

Cohen, S., Willgoose, G., & Hancock, G. (2010). The mARM3D spatially distributed soil evolution model: Three-dimensional model framework and analysis of hillslope and landform responses. *Journal of Geophysical Research: Earth Surface*, *115*.

Coulouma, G., Prevot, L., & Lagacherie, P. (2020). Carbon isotope discrimination as a surrogate for soil available water capacity in rainfed areas: A study in the Languedoc vineyard plain. *Geoderma*, *362*, 114121.

Défossez, P., Richard, G., Boizard, H., & O'Sullivan, M. F. (2003). Modeling change in soil compaction due to agricultural traffic as function of soil water content. *Geoderma*, *116*, 89–105.

Diekkrüger, B., Igwe, C. A. et al. (2015). Soil survey and soil classification of the Koupendri catchment in Benin, West Africa. *African Journal of Agricultural Research*, *10*, 3938–3951.

Diggle, P. J., & Ribeiro, P. J. (2007). *Model-based Geostatistics.* Springer series in statistics, Berlin, Germany.

DiMucci, D., Kon, M., & Segrè, D. (2018). Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem networks. *Msystems*, *3*, e00181–18.

Dominati, E., Mackay, A., Green, S., & Patterson, M. (2014). A soil change-based methodology for the quantification and valuation of ecosystem services from agro-ecosystems: a case study of pastoral agriculture in New Zealand. *Ecological Economics*, *100*, 119–129.

FAO (1977). *A Framework for Land Evaluation.* 32. International Institute for Land Reclamation and Improvement, Agriculture Organization of the United Nations. Soil Resources and Conservation Service.

Feddes, R. A., Kowalik, P. J., & Zaradny, H. (1978). *Simulation of Field Water Use and Crop Yield.* Halsted Press, John Wiley & Sons, New York, USA.

Field, D. J., Morgan, C. L. S., & McBratney, A. B. (2016). *Global Soil Security.* Springer, Berlin, Germany.

Finke, P. A. (2012). On digital soil assessment with models and the Pedometrics agenda. *Geoderma*, *171*, 3–15.

Follain, S., Minasny, B., McBratney, A. B., & Walter, C. (2006). Simulation of soil thickness evolution in a complex agricultural landscape at fine spatial and temporal scales. *Geoderma*, *133*, 71–86.

Franz, D., Acosta, M., Altimir, N., Arriga, N., Arrouays, D., Aubinet, M., Aurela, M., Ayres, E., López-Ballesteros, A., Barbaste, M., Berveiller, D., Biraud, S., Boukir, H., Brown, T., Brümmer, C., Buchmann, N., Burba, G., Carrara, A., Cescatti, A., Ceschia, E., Clement, R., Cremonese, E., Crill, P., Darenova, E., Dengel, S., D'Odorico, P., Filippa, G., Fleck, S., Fratini, G., Fuss, R., Gielen, B., Gogo, S., Grace, J., Graf, A., Grelle, A., Gross, P., Grünwald, T., Haapanala, S., Hehn, M., Heinesch, B., & Ibrom, A. (2018). Towards long-term standardised carbon and greenhouse gas observations for monitoring Europe's terrestrial ecosystems: a review. *International Agrophysics*, *32*, 439–455.

Furey, J., Davis, A., & Seiter-Moser, J. (2019). Natural language indexing for pedoinformatics. *Geoderma*, *334*, 49–54.

Giasson, E., Van Es, C., Van Wambeke, A., & Bryant, R. B. (2000). Assessing the economic value of soil information using decision analysis techniques. *Soil Science*, *165*, 971–978.

Giltrap, D. L., Li, C., & Saggar, S. (2010). DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. *Agriculture, Ecosystems & Environment*, *136*, 292–300.

Goldshleger, N., Ben-Dor, E., Benyamini, Y., Blumberg, D., & Agassi, M. (2002). Spectral properties and hydraulic conductance of soil crusts formed by raindrop impact. *International Journal of Remote Sensing*, *23*, 3909–3920.

Goovaerts, P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*, *89*, 1–45.

Greiner, L., Keller, A., Grêt-Regamey, A., & Papritz, A. (2017). Soil function assessment: review of methods for quantifying the contributions of soils to ecosystem services. *Land Use Policy*, *69*, 224–237.

Greiner, L., Nussbaum, M., Papritz, A. J., Zimmermann, S., Gubler, A., Grêt-Regamey, A., & Keller, A. (2018). Uncertainty indication in soil function maps–transparent and easy-to-use information to support sustainable use of soil resources. *Soil*, *4*, 123–139.

Grierson, C. S., Barnes, S. R., Chase, M. W., Clarke, M., Grierson, D., Edwards, K. J., Jellis, G. J., Jones, J. D., Knapp, S., Oldroyd, G., Poppy, G., Temple, P., Williams, R., & Bastow, R. (2011). One hundred important questions facing plant science research. *New Phytologist*, *192*, 6–12.

Gudmundsson, A. (2013). Great challenges in structural geology and tectonics. *Frontiers in Earth Science*, *1*, 1–10.

Guo, Y., Gong, P., & Amundson, R. (2003). Pedodiversity in the United States of America. *Geoderma*, *117*, 99–115.

Hartemink, A. E., McBratney, A. B., & Cattle, J. A. (2001). Developments and trends in soil science: 100 volumes of Geoderma (1967–2001). *Geoderma*, *100*, 217–268.

Hendriks, C., Stoorvogel, J. J., Álvarez-Martínez, J. M., Claessens, L., Pérez-Silos, I., & Barquín, J. (in Press). Introducing a mechanistic model in digital soil mapping to predict soil organic matter stocks in the Cantabrian region (Spain). *European Journal of Soil Science*, .

Heuvelink, G. B. M. (1998). Uncertainty analysis in environmental modelling under a change of spatial scale. In P. A. Finke, J. Bouma, & M. R. Hoosbeek (Eds.), *Soil and Water Quality at Different Scales* (pp. 255–264). Springer, Berlin, Germany.

Heuvelink, G. B. M. (2018). Uncertainty and uncertainty propagation in soil mapping and modelling. In B. Minasny, A. B. McBratney, & U. Stockmann (Eds.), *Pedometrics* (pp. 439–461). Springer, Berlin, Germany.

Heuvelink, G. B. M. (2019). The "10PM Challenges". *Pedometron*, *43*, 9–13.

Heuvelink, G. B. M., Schoorl, J. M., Veldkamp, A., & Pennock, D. J. (2006). Space–time Kalman filtering of soil redistribution. *Geoderma*, *133*, 124–137.

Heuvelink, G. B. M., & Webster, R. (2001). Modelling soil variation: past, present, and future. *Geoderma*, *100*, 269–301.

Hewitt, A. E. (2010). *New Zealand Soil Classification.*. Manaaki Whenua Press, Lincoln, New Zealand.

Hilbert, D. (1902). Mathematical problems. *Bulletin of the American Mathematical Society*, *8*, 437–479.

Hobley, D., Adams, J. M., Siddhartha Nudurupati, S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker, G. E. (2017). Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. *Earth Surface Dynamics*, *5*, 21–46.

Hoosbeek, M. R., & Bryant, R. B. (1992). Towards the quantitative modeling of pedogenesis-a review. *Geoderma*, *55*, 183–210.

Hu, S., Coleman, D. C., Hendrix, P. F., & Beare, M. H. (1995). Biotic manipulation effects on soil carbohydrates and microbial biomass in a cultivated soil. *Soil Biology and Biochemistry*, *27*, 1127–1135.

Hughes, P., McBratney, A., Huang, J., Minasny, B., Hempel, J., Palmer, D. J., & Micheli, E. (2017a). Creating a novel comprehensive soil classification system by sequentially adding taxa from existing systems. *Geoderma Regional*, *11*, 123–140.

Hughes, P., McBratney, A. B., Huang, J., Minasny, B., Micheli, E., & Hempel, J. (2017b). Comparisons between USDA Soil Taxonomy and the Australian Soil Classification System I: Data harmonization, calculation of taxonomic distance and inter-taxa variation. *Geoderma*, *307*, 198–209.

Hughes, P., McBratney, A. B., Huang, J., Minasny, B., Micheli, E., & Hempel, J. (2018a). A nomenclature algorithm for a potentially global soil taxonomy. *Geoderma*, *322*, 56–70.

Hughes, P., McBratney, A. B., Minasny, B., Huang, J., Micheli, E., Hempel, J., & Jones, E. (2018b). Comparisons between USDA soil taxonomy and the Australian Soil Classification system II: Comparison of order, suborder and great group taxa. *Geoderma*, *322*, 48–55.

Huston, M. A. (1994). *Biological Diversity: the Coexistence of Species*. Cambridge University Press, Cambridge, UK.

Ibañez, J. J., De-Albs, S., Bermúdez, F. F., & García-Álvarez, A. (1995). Pedodiversity: concepts and measures. *CATENA*, *24*, 215–232.

Ibáñez, J. J., De-Alba, S., Lobo, A., & Zucarello, V. (1998). Pedodiversity and global soil patterns at coarse scales (with Discussion). *Geoderma*, *83*, 171–192.

Ibáñez, J. J., & Feoli, E. (2013). Global relationships of pedodiversity and biodiversity. *Vadose Zone Journal*, *12*.

Ibanez, J. J., Krasilnikov, P. V., & Saldana, A. (2012). Archive and refugia of soil organisms: applying a pedodiversity framework for the conservation of biological and non-biological heritages. *Journal of Applied Ecology*, *49*, 1267–1277.

Ilie, I., Dittrich, P., Carvalhais, N., Jung, M., Heinemeyer, A., Migliavacca, M., Morison, J., Sippel, S., Subke, J.-A., Wilkinson, M. et al. (2017). Reverse engineering model structures for soil and ecosystem respiration: the potential of gene expression programming. *Geoscientific Model Development*, *10*, 3519–3545.

Isbell, R. (2016). *The Australian Soil Classification*. CSIRO publishing, Canberra, Australia.

IUSS Working Group (2015). *World Reference Base for Soil Resources 2014: International Soil Classification System for Naming Soils and Creating Legends for Soil Maps* Volume 106. FAO Rome, Italy. World Soil Resources Reports.

Jiang, P., Anderson, S. H., Kitchen, N. R., Sudduth, K. A., & Sadler, E. J. (2007). Estimating plant-available water capacity for claypan landscapes using apparent electrical conductivity. *Soil Science Society of America Journal*, *71*, 1902–1908.

Jury, W. A., Or, D., Pachepsky, Y., Vereecken, H., Hopmans, J. W., Ahuja, L. R., Clothier, B. E., Bristow, K. L., Kluitenberg, G. J., Moldrup, P., Šimůnek, J., Th. van Genuchten, M., & Horton, R. (2011). Kirkham's legacy and contemporary challenges in soil physics research. *Soil Science Society of America Journal*, *75*, 1589–1601.

Karimi, B., Villerd, J., Dequiedt, S., Terrat, S., Chemidlin-Prévost Bouré, N., Djemiel, C., Lelièvre, M., Tripied, J., Nowak, V., Saby, N. P. A., Bispo, A., Jolivet, C., Arrouays, D., Wincker, P., Cruaud, C., & Ranjard, L. (2020). Biogeography of soil microbial habitats across France. *Global Ecology and Biogeography*, *29*, 1399–1411.

Kerry, R., Goovaerts, P., Rawlins, B. G., & Marchant, B. P. (2012). Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma*, *170*, 347–358.

Keuskamp, J. A., Dingemans, B. J. J., Lehtinen, T., Sarneel, J. M., & Hefting, M. M. (2013). Tea bag index: a novel approach to collect uniform decomposition data across ecosystems. *Methods in Ecology and Evolution*, *4*, 1070–1075.

Kidd, D. B., Webb, M. A., Grose, C. J., Moreton, R. M., Malone, B. P., McBratney, A. B., Minasny, B., Viscarra Rossel, R. A., Sparrow, L. A., & Smith, R. (2012). Digital soil assessment: Guiding irrigation expansion in Tasmania, Australia. In B. Minasny, B. P. Malone, & A. B. McBratney (Eds.), *Digital Soil Assessments and Beyond* (pp. 3–8). CRC Press, London, UK.

Kuang, B., Mahmood, H. S., Quraishi, M. Z., Hoogmoed, W. B., Mouazen, A. M., & van Henten, E. J. (2012). Sensing soil properties in the laboratory, in situ, and on-line: a review. In *Advances in Agronomy* (pp. 155–223). Elsevier Volume 114.

Lafolie, F., Cousin, I., Maron, P.-A., Mollier, A., Pot-Genty, V., Moitrier, N., Moitrier, N., & Nouguier, C. (2013). The Soil Virtuel modeling platform. In *Sino-French Training Course AIR ECO SOLS. Interdisciplinary Approach of Soil Ecological Restoration.* Nancy, France. URL: `https://hal.inrae.fr/hal-02805313` [Accessed 15 January 2021].

Lagacherie, P. (2005). An algorithm for fuzzy pattern matching to allocate soil individuals to pre-existing soil classes. *Geoderma*, *128*, 274–288.

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., & Gomez, C. (2018). Vis-NIR-SWIR Remote Sensing Products as New Soil Data for Digital Soil Mapping. In B. Minasny, A. B. McBratney, & U. Stockmann (Eds.), *Pedometrics* (pp. 415–437). Springer, Berlin, Germany.

Lark, R. M., & Knights, K. V. (2015). The implicit loss function for errors in soil information. *Geoderma*, *251*, 24–32.

Lin, L.-F., Ebtehaj, A. M., Wang, J., & Bras, R. L. (2017). Soil moisture background error covariance and data assimilation in a coupled land-atmosphere model. *Water Resources Research*, *53*, 1309–1335.

Ma, Y., Minasny, B., Malone, B. P., & Mcbratney, A. B. (2019a). Pedology and digital soil mapping (DSM). *European Journal of Soil Science*, *70*, 216–235.

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

Mastrandrea, M. D., & Mach, K. J. (2011). Treatment of uncertainties in IPCC Assessment Reports: past approaches and considerations for the Fifth Assessment Report. *Climatic Change*, *108*, 659–673.

Mayr, T., Rivas-Casado, M., Bellamy, P., Palmer, R., Zawadzka, J., & Corstanje, R. (2010). Two methods for using legacy data in digital soil mapping. In J. L. Boettinger, D. W. Howell, A. C. Moore, & A. E. Hartemink (Eds.), *Digital Soil Mapping, Bridging Research, Environmental Application, and Operation* (pp. 191–202). Springer.

McBratney, A. B., & De Gruijter, J. J. (1992). A continuum approach to soil classification by modified fuzzy k-means with extragrades. *Journal of Soil Science*, *43*, 159–175.

McBratney, A. B., Field, D., Morgan, C. L. S., & Huang, J. (2019a). On soil capability, capacity, and condition. *Sustainability*, *11*, 3350.

McBratney, A. B., Field, D. J., & Koch, A. (2014). The dimensions of soil security. *Geoderma*, *213*, 203–213.

McBratney, A. B., de Gruijter, J., & Bryce, A. (2019b). Pedometrics timeline. *Geoderma*, *338*, 568–575.

McBratney, A. B., & Minasny, B. (2007). On measuring pedodiversity. *Geoderma*, *141*, 149–154.

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

McBratney, A. B., Odeh, I. O. A., Bishop, T. F. A., Dunbar, M. S., & Shatar, T. M. (2000). An overview of pedometric techniques for use in soil survey. *Geoderma*, *97*, 293–327.

McBratney, A. B., Webster, R., & Burgess, T. M. (1981). The design of optimal sampling schemes for local estimation and mapping of of regionalized variables- I: Theory and method. *Computers & Geosciences*, *7*, 331–334.

McKenzie, N. J., Grundy, M. J., Webster, R., & Ringrose-Voase, A. J. (2008). *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbooks Series. CSIRO publishing, Australia.

McNicol, G., Bulmer, C., D'Amore, D., Sanborn, P., Saunders, S., Giesbrecht, I., Arriola, S. G., Bidlack, A., Butman, D., & Buma, B. (2019). Large, climate-sensitive soil carbon stocks mapped with pedology-informed machine learning in the North Pacific coastal temperate rainforest. *Environmental Research Letters*, *14*, 014004.

Michéli, E., Láng, V., Owens, P. R., McBratney, A. B., & Hempel, J. (2016). Testing the pedometric evaluation of taxonomic units on soil taxonomy-a step in advancing towards a universal soil classification system. *Geoderma*, *264*, 340–349.

Milne, G. (1936). Normal erosion as a factor in soil profile development. *Nature*, *138*, 548–549.

Minasny, B., & McBratney, A. B. (2006). Mechanistic soil–landscape modelling as an approach to developing pedogenetic classifications. *Geoderma*, *133*, 138–149.

Morand, D. T. (2013). The World Reference Base for Soils (WRB) and Soil Taxonomy: an appraisal of their application to the soils of the Northern Rivers of New South Wales. *Soil Research*, *51*, 167–181.

Mückenhausen, E., & Vogel, F. (1962). *Entstehung, Eigenschaften und Systematik der Böden der Bundesrepublik Deutschland.* DLG Verlags, Frankfurt Am Main, Germany.

Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M. E., & Papritz, A. J. (2018). Evaluation of digital soil mapping approaches with large sets of environmental covariates. *SOIL*, *4*, 1–22.

Nyamapfene, K. W. (1991). *The Soils of Zimbabwe* Volume 1. Nehanda Publishers, Harare.

O'Brien, S. L., Gibbons, S. M., Owens, S. M., Hampton-Marcell, J., Johnston, E. R., Jastrow, J. D., Gilbert, J. A., Meyer, F., & Antonopoulos, D. A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environmental Microbiology*, *18*, 2039–2051.

Odgers, N. P., Sun, W., McBratney, A. B., Minasny, B., & Clifford, D. (2014). Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma*, *214*, 91–100.

Olden, J. D., & Jackson, D. A. (2002). Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, *154*, 135–150.

Orton, T. G., Pringle, M. J., & Bishop, T. F. A. (2016). A one-step approach for modelling and mapping soil properties based on profile data sampled over varying depth intervals. *Geoderma*, *262*, 174–186.

Pachepsky, Y., & Hill, R. L. (2017). Scale and scaling in soils. *Geoderma*, *287*, 4–30.

Paul, K. I., Polglase, P. J., & Richards, G. P. (2003). Sensitivity analysis of predicted change in soil carbon following afforestation. *Ecological Modelling*, *164*, 137–152.

Petersen, A., Gröngröft, A., & Miehlich, G. (2010). Methods to quantify the pedodiversity of 1 km^2 areas—results from southern African drylands. *Geoderma*, *155*, 140–146.

Pham, B. T., Nguyen, M. D., Dao, D. V., Prakash, I., Ly, H.-B., Le, T.-T., Ho, L. S., Nguyen, K. T., Ngo, T. Q., Hoang, V., Son, L. H., Ngo, H. T. T., Tran, H. T., Do, N. M., Van Le, H., Ho, H. L., & Tien Bui, D. (2019). Development of artificial intelligence models for the prediction of Compression Coefficient of soil: An application of Monte Carlo sensitivity analysis. *Science of The Total Environment*, *679*, 172–184.

Phillips, J. D., & Marion, D. A. (2005). Biomechanical effects, lithological variations, and local pedodiversity in some forest soils of Arkansas. *Geoderma*, *124*, 73–89.

Pino, V., McBratney, A. B., Fajardo, M., Wilson, N., & Deaker, R. (2019). Understanding soil biodiversity using two orthogonal 1000km transects across New South Wales, Australia. *Geoderma*, *354*, 113860.

Ramsey, M. H., Taylor, P. D., & Lee, J.-C. (2002). Optimized contaminated land investigation at minimum overall cost to achieve fitness-for-purpose. *Journal of Environmental Monitoring*, *4*, 809–814.

Ranjard, L., Dequiedt, S., Prévost-Bouré, N. C., Thioulouse, J., Saby, N. P. A., Lelievre, M., Maron, P. A., Morin, F. E. R., Bispo, A., Jolivet, C., Arrouays, D., & Lemanceau, P. (2013). Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nature Communications*, *4*, 1–10.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In B. Krishnapuram (Ed.), *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144).

Rossiter, D. G. (1996). A theoretical framework for land evaluation. *Geoderma*, *72*, 165–190.

Rossiter, D. G., Hewitt, A. E., & Dominati, E. J. (2018). Pedometric valuation of the soil resource. In B. Minasny, A. B. McBratney, & U. Stockmann (Eds.), *Pedometrics* (pp. 521–546). Springer, Berlin, Germany.

Rossiter, D. G., Liu, J., Carlisle, S., & Zhu, A.-X. (2015). Can citizen science assist digital soil mapping? *Geoderma*, *259*, 71–80.

dos Santos, H. G., Jacomine, P. K. T., Dos Anjos, L. H. C., De Oliveira, V. A., Lumbreras, J. F., Coelho, M. R., De Almeida, J. A., de Araujo Filho, J. C., De Oliveira, J. B., & Cunha, T. J. F. (2018). *Sistema Brasileiro de Classificação de Solos.*. Brasília, DF: Embrapa, Brazil.

Schoorl, J. M., Sonneveld, M. P. W., & Veldkamp, A. (2000). Three-dimensional landscape process modelling: the effect of DEM resolution. *Earth Surface Processes and Landforms*, *25*, 1025–1034.

Schulla, J., & Jasper, K. (2007). *Model description WaSiM-ETH*. Technical Report Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich.

Scull, P., Franklin, J., Chadwick, O. A., & McArthur, D. (2003). Predictive soil mapping: a review. *Progress in Physical Geography*, *27*, 171–197.

Simonson, R. W. (1959). Outline of a generalized theory of soil genesis. *Soil Science Society of America Journal*, *23*, 152–156.

Šimůnek, J., Jarvis, N. J., Van Genuchten, M. T., & Gärdenäs, A. (2003). Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. *Journal of Hydrology*, *272*, 14–35.

Six, J., Bossuyt, H., Degryze, S., & Denef, K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil and Tillage Research*, *79*, 7–31.

Soil Science Division Staff (2017). *Soil Survey Manual.* Government Printing Office, Washington, D.C., USA. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18.

Sreelash, K., Buis, S., Sekhar, M., Ruiz, L., Tomer, S. K., & Guerif, M. (2017). Estimation of available water capacity components of two-layered soils using crop model inversion: Effect of crop type and water regime. *Journal of Hydrology*, *546*, 166–178.

Stumpf, F., Goebes, P., Schmidt, K., Schindewolf, M., Schönbrodt-Stitt, S., Wadoux, A., Xiang, W., & Scholten, T. (2017). Sediment reallocations due to erosive rainfall events in the Three Gorges Reservoir Area, Central China. *Land Degradation & Development*, *28*, 1212–1227.

Styc, Q., & Lagacherie, P. (2019). What is the best inference trajectory for mapping soil functions: An example of mapping soil available water capacity over Languedoc Roussillon (France). *Soil Systems*, *3*, 34.

Sun, S., Li, S., Avera, B. N., Strahm, B. D., & Badgley, B. D. (2017). Soil bacterial and fungal communities show distinct recovery patterns during forest ecosystem restoration. *Applied and Environmental Microbiology*, *83*.

Sutherland, W. J., Freckleton, R. P., Godfray, H. C. J., Beissinger, S. R., Benton, T., Cameron, D. D., Carmel, Y., Coomes, D. A., Coulson, T., Emmerson, M. C., Hails, R. S., Hays, G. C., Hodgson, D. J., Hutchings, M. J., Johnson, D., Jones, J. P. G., Keeling, M. J., Kokko, H., Kunin, W. E., Lambin, X., Lewis, O. T., Malhi, Y., Mieszkowska, N., Milner-Gulland, E. J., Norris, K., Phillimore, A. B., Purves, D. W., Reid, J. M., Reuman, D. C., Thompson, K., Travis, J. M. J., Turnbull, L. A., Wardle, D. A., & Wiegand, T. (2013). Identification of 100 fundamental ecological questions. *Journal of Ecology*, *101*, 58–67.

Taalab, K., Corstanje, R., Mayr, T. M., Whelan, M. J., & Creamer, R. E. (2015). The application of expert knowledge in Bayesian networks to predict soil bulk density at the landscape scale. *European Journal of Soil Science*, *66*, 930–941.

Temme, A. J. A. M., & Vanwalleghem, T. (2016). LORICA–a new model for linking landscape and soil profile evolution: development and sensitivity analysis. *Computers & Geosciences*, *90*, 131–143.

Tompkins, G. A., McIntosh, T. H., & Dunigan, E. P. (1968). Use of the Stanford-DeMent Bioassay to Study Atrazine-Soil Reactions. *Soil Science Society of America Journal*, *32*, 373–377.

USDA-SCS (1986). *The Guy Smith Interviews: Rationale for Concepts in Soil Taxonomy.* 11. Cornell University, Department of Agronomy, Cornell, USA.

Van Den Hoogen, J., Geisen, S., Routh, D., Ferris, H., Traunspurger, W., Wardle, D. A., De Goede, R. G. M., Adams, B. J., Ahmad, W., Andriuzzi, W. S. et al. (2019). Soil nematode abundance and functional group composition at a global scale. *Nature*, *572*, 194–198.

Vašát, R., Heuvelink, G. B. M., & Borůvka, L. (2010). Sampling design optimization for multivariate soil mapping. *Geoderma*, *155*, 147–153.

Vaysse, K., & Lagacherie, P. (2017). Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma*, *291*, 55–64.

Veerman, C., Correia, T. P., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E.-A., Grand, A., & Filchev, L. H. (2020). *Caring for soil is caring for life: ensure 75% of soils are healthy by 2030 for healthy food, people, nature and climate: Report of the Mission Board for Soil health and food.* Technical Report Publications Office of the European Union, Luxembourg.

Veldkamp, A., Baartman, J. E. M., Coulthard, T. J., Maddy, D., Schoorl, J. M., Storms, J. E. A., Temme, A. J. A. M., van Balen, R., van De Wiel, M. J., Van Gorp, W., Viveen, W., Westaway, R., & Whittaker, A. C. (2017). Two decades of numerical modelling to understand long term fluvial archives: Advances and future perspectives. *Quaternary Science Reviews*, *166*, 177–187.

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M. H., Amelung, W., Aitkenhead, M., Allison, S. D., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Hendricks Franssen, H. J., Heppell, J., Horn, R., Huisman, J. A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, K., Leitner, D., McBratney, A. B., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Rothfuss, Y., Rowe, E. C., Schwen, A., Šimůnek, J., Tiktak, A., Van Dam, J., van der Zee, S. E. A. T. M., Vogel, H. J., Vrugt, J. A., Wöhling, T., & Young, I. M. (2016). Modeling soil processes: Review, key challenges, and new perspectives. *Vadose Zone Journal*, *15*.

Verheyen, K., Adriaens, D., Hermy, M., & Deckers, S. (2001). High-resolution continuous soil classification using morphological soil profile descriptions. *Geoderma*, *101*, 31–48.

Viscarra Rossel, R. A., Chen, C., Grundy, M. J., Searle, R., Clifford, D., & Campbell, P. H. (2015). The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. *Soil Research*, *53*, 845–864.

Viscarra Rossel, R. A., Goovaerts, P., & McBratney, A. B. (2001). Assessment of the production and economic risks of site-specific liming using geostatistical uncertainty modelling. *Environmetrics*, *12*, 699–711.

Viscarra Rossel, R. A., & McBratney, A. B. (2008). Diffuse reflectance spectroscopy as a tool for digital soil mapping. In A. E. Hartemink, A. B. McBratney, & M. d. L. Mendonça-Santos (Eds.), *Digital Soil Mapping with Limited Data* (pp. 165–172). Springer, Berlin, Germany.

Vogel, H.-J., Eberhardt, E., Franko, U., Lang, B., Ließ, M., Weller, U., Wiesmeier, M., & Wollschläger, U. (2019). Quantitative evaluation of soil functions: potential and state. *Frontiers in Environmental Science*, *7*, 164.

Vos, C., Don, A., Hobley, E. U., Prietz, R., Heidkamp, A., & Freibauer, A. (2019). Factors controlling the variation in organic carbon stocks in agricultural soils of Germany. *European Journal of Soil Science*, *70*, 550–564.

Wadoux, A. M. J.-C. (2019). Using deep learning for multivariate mapping of soil with quantified uncertainty. *Geoderma*, *351*, 59–70.

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

Wadoux, A. M. J.-C., & McBratney, A. B. (2021). Hypotheses, machine learning and soil mapping. *Geoderma*, *383*, 114725.

Wadoux, A. M. J.-C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth-Science Reviews*, *210*, 103359.

Walter, C., Lagacherie, P., & Follain, S. (2006). Integrating pedological knowledge into digital soil mapping. In P. Lagacherie, A. B. McBratney, & M. Voltz (Eds.), *Digital Soil Mapping, An Introductory Perspective* (pp. 281–615). Elsevier Volume 31.

Wang, J., Li, Y., Bork, E. W., Richter, G. M., Eum, H.-I., Chen, C., Shah, S. H. H., & Mezbahuddin, S. (2020). Modelling spatio-temporal patterns of soil carbon and greenhouse gas emissions in grazing lands: Current status and future prospects. *Science of The Total Environment*, (p. 139092).

Webster, R. (1991). Local disjunctive kriging of soil properties with change of support. *Journal of Soil Science*, *42*, 301–318.

Webster, R. (1994). The development of pedometrics. *Geoderma*, *62*, 1–15.

Webster, R. (2015). Technological developments for spatial prediction of soil properties, and Danie Krige's influence on it. *Journal of the Southern African Institute of Mining and Metallurgy*, *115*, 165–172.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. John Wiley & Sons, Chichester, UK.

Western, A. W., & Blöschl, G. (1999). On the spatial scaling of soil moisture. *Journal of Hydrology*, *217*, 203–224.

Western, S. (1978). *Soil Survey Contracts and Quality Control*. Monographs on Soil Survey. Oxford University Press, Oxford, UK.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N. et al. (2019). Soil organic carbon storage as a key function of soils-A review of drivers and indicators at various scales. *Geoderma*, *333*, 149–162.

Xi, C. F. (1998). *Soils of China*. Chinese Agriculture Press, Beijing, China. 1253 pp. (in Chinese).

Yaalon, D. H. (1975). Conceptual models in pedogenesis: Can soil-forming functions be solved? *Geoderma*, *14*, 189–205.

Zheng, W., Zhao, Z., Gong, Q., Zhai, B., & Li, Z. (2018). Responses of fungal–bacterial community and network to organic inputs vary among different spatial habitats in soil. *Soil Biology and Biochemistry*, *125*, 54–63.

Zhu, A. X., R Lark, M., Minasny, B., & Huang, Y. (2012). Entering the digital world (Pedometrics 2009). *Geoderma*, *171*, 1–2.