



GlobalSoilMap: Toward a Fine-Resolution Global Grid of Soil Properties

Dominique Arrouays^{*}, Michael G. Grundy[†], Alfred E. Hartemink[‡],
Jonathan W. Hempel[§], Gerard B.M. Heuvelink[¶], S. Young Hong^{||},
Philippe Lagacherie[#], Glenn Lelyk^{**}, Alexander B. McBratney^{††},
Neil J. McKenzie^{‡‡,1}, Maria d.L. Mendonca-Santos^{§§},
Budiman Minasny^{††}, Luca Montanarella^{¶¶}, Inakwu O.A. Odeh^{††},
Pedro A. Sanchez^{||||}, James A. Thompson^{##}, Gan-Lin Zhang^{***}

^{*}INRA, InfoSol Unit, Orleans, France

[†]CSIRO, EcoSciences Precinct, Dutton Park, Queensland, Australia

[‡]University of Wisconsin-Madison, Department of Soil Science, Madison, USA

[§]United States Department of Agriculture, Natural Resources Conservation Service, Lincoln, Nebraska, USA

[¶]ISRIC—World Soil Information, Wageningen, Netherlands

^{||}National Academy of Agricultural Science, Rural Development Administration, Suwon, South Korea

[#]INRA, IRD, Lab Etud Interact Sols Agrosyst Hydrosyst, Montpellier, France

^{**}Agriculture and Agri-Food Canada, University of Manitoba, Winnipeg, Manitoba, Canada

^{††}Faculty of Agriculture and Environment, The University of Sydney, Sydney, New South Wales, Australia

^{‡‡}CSIRO Australia, Campus International de Baillarguet, Montpellier, Cedex, France

^{§§}EMBRAPA—Brazilian Agricultural Research Corporation/The National Centre of Soil Research (Embrapa Solos), Rio de Janeiro, Brazil

^{¶¶}European Commission—DG JRC, Ispra, Varese, Italy

^{||||}The Earth Institute at Columbia University, Palisades, New York, USA

^{##}West Virginia University, Morgantown, West Virginia, USA

^{***}State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, PR China

¹Corresponding author: e-mail address: neil.mckenzie@csiro.au

Contents

1. Introduction	94
2. Information for Soil Assessment	95
2.1 Mapping, modeling, and monitoring	96
2.2 The new global imperative for soil science	96
2.3 Current map coverage	98
2.4 Conventional concepts of soil survey and classification	99
2.5 Grids or polygons or both?	100
3. Digital Soil Mapping and Origins of <i>GlobalSoilMap</i>	101
4. Technical Specifications of <i>GlobalSoilMap</i>	102
4.1 Defining the soil individual	103
4.2 Definition of the grid	107
5. Minimum Data Set	109
5.1 Concept	109
5.2 Time	113

6. Methods for Estimating Soil Properties Using Legacy Data	114
6.1 Legacy data only	114
6.2 Acquiring new data	116
7. Uncertainty	117
8. Information Architecture, Web Services, and Applications	120
9. Examples	121
10. Toward an Enduring Global Soil Information System	127
10.1 Governance	127
10.2 Mutual benefit	128
10.3 Moving from research to operational implementation	128
11. Conclusions	129
Acknowledgments	129
References	129

Abstract

Soil scientists are being challenged to provide assessments of soil condition from local through to global scales. A particular issue is the need for estimates of the stores and fluxes in soils of water, carbon, nutrients, and solutes. This review outlines progress in the development and testing of *GlobalSoilMap*—a digital soil map that aims to provide a fine-resolution global grid of soil functional properties with estimates of their associated uncertainties. A range of methods can be used to generate the fine-resolution spatial estimates depending on the availability of existing soil surveys, environmental data, and point observations. The system has an explicit geometry for estimating point and block estimates of soil properties continuously down the soil profile. This geometry is necessary to ensure mass balance when stores and fluxes are computed. It also overcomes some limitations with existing systems for characterizing soil variation with depth. *GlobalSoilMap* has been designed to enable delivery of soil data via Web services. This review provides an overview of the system's technical specifications including the minimum data set. Examples from contrasting countries and environments are then presented to demonstrate the robustness of the technical specifications. *GlobalSoilMap* provides the means for supplying soil information in a format and resolution compatible with other fundamental data sets from remote sensing, terrain analysis, and other systems for mapping, monitoring, and forecasting biophysical processes. The initial research phase of the core project is nearing completion and attention is now shifting toward establishing the institutional and governance arrangements necessary to complete a full global coverage and maintaining the operational version of the *GlobalSoilMap*. This will be a grand and rewarding challenge for the soil science profession in the coming years.



1. INTRODUCTION

There is a renewed awareness of the finite nature of the world's soil resources, growing concern about soil security, and significant uncertainties

about the carrying capacity of the planet (FAO, 2012; Hartemink and McBratney, 2008; UNEP, 2012). Soil scientists are being challenged to provide assessments of soil condition from local through to global scales (Grunwald et al., 2011). However, only a few countries have the necessary survey and monitoring programs to meet these new needs and global data sets are out-of-date. A particular issue is the demand for estimates of the stores and fluxes in soils of water, carbon, nutrients, and solutes. Conventional systems for mapping and classifying soils were not designed for this purpose. Instead, a global grid with estimates of soil functional properties is required with the explicit geometry necessary for ensuring mass balance when stores and fluxes are computed. The term “functional properties of soils” is used here to refer to *soil characteristics* that can be used to estimate *soil qualities* (*sensu* Dent and Young, 1981; FAO, 1976) or *soil functions* (*sensu* Blum, 1993) relating to flows, transformations, and stores of materials in soils.

This review summarizes progress toward development and testing of a global grid of soil functional properties that can be accessed via Web services. It describes solutions to several theoretical and practical problems along with an overview of the system’s technical specifications. Examples from contrasting countries and environments are then presented to demonstrate the robustness of the technical specifications. We begin with a consideration of the general demand for soil information and then focus on why consistent global soil information is needed.



2. INFORMATION FOR SOIL ASSESSMENT

The primary reasons for assessing soil and land resources are to know what resources are present, what land uses are suitable, and how to manage the soil to produce food and fiber, to secure water supplies, and to conserve the ecosystem services that it delivers to all of us. This information is of most value when it reduces risks in decision-making whether at the local, regional, national, continental, or global scale (Bouma, 2002; Carré et al., 2007a). Around the world, millions of people profitably use existing soil information to either produce more food, protect environmental assets, or both (e.g., Bouma, 1989; Ji and Peters, 2003; Srinivasan et al., 2010). However, many more would benefit if they could easily obtain up-to-date information on soil function in a wider range of formats along with an estimate of its uncertainty. The latter is essential for understanding and dealing with risk in decision-making.

2.1. Mapping, modeling, and monitoring

Three important components of soil knowledge necessary for sustainable land use and management are:

1. an understanding of spatial variations in soil function (e.g., maps and spatial information)
2. an ability to detect and interpret soil-change with time (e.g., via monitoring sites, long-term experiments, dynamic environmental proxies)
3. a capacity to forecast the likely future state of soils under specified systems of land management and future climates (e.g., through the use of simulation models)

This review deals with the first component. Mapping soil resources provides basic information on landscape attributes and it is essential for sound planning and management at all scales. Mapping also provides a spatial framework for monitoring (i.e., component 2 above) and supplies input data for simulation models (i.e., component 3 above) (e.g., [Lin et al., 2005](#)). Simulation modeling is important because it can be used to explore how soil conditions may change in ways that other methods cannot. For example, variations in climate might mask subtle but important changes in soil health, and detection of a statistically significant change through field measurement might be possible only over an impractically long period (i.e., 50 years or more). At present, many simulation models have rudimentary representations of soil processes partly because of the difficulty of obtaining suitable input data. Model developers are likely to build more realistic modules of soil processes if data on functional properties are more readily available.

Our focus on the functional properties of soils is based on the recognition that most practical questions relating to agricultural production, land degradation, and environmental management require estimates of material flows and transformations in soils. Of particular importance are the balances of water (availability to plants, leaching, water-logging, run-off), nutrients (plant growth), carbon (sequestration of greenhouse gases, soil health), solutes (salinization), and protons (e.g., acidification).

2.2. The new global imperative for soil science

There is a qualitative appreciation of the pressures on global soil resources but limited consistent evidence on their condition and trajectories of change. In short, the world's soils need to support at least a 60% increase in agricultural production between 2006 and 2050 to meet the projected

growth in demand (FAO, 2012) but there appears to be serious constraints that include:

- a finite supply of arable land with suitable soils—the current and potential area of such land is not defined with sufficient certainty at present and strongly contrasting conclusions are being drawn (cf. Ausubel et al., 2012; FAO, 2011)
- yield plateaux for major crops—the rate of yield increase in some staple crops in some regions is slowing (van Wart et al., 2013) and soil constraints are a significant consideration
- increasing costs of energy, nutrients, and emissions—improvements in farming systems and soil management practices are needed to conserve nutrients, reduce emissions, and minimize reliance on fossil fuels
- soil degradation—the rates and distribution of threatening processes (e.g., erosion, nutrient depletion, acidification, salinization, compaction) are poorly characterized in most parts of the world but significant areas are known to be degrading (FAO, 2011; Sonnerfeld and Dent, 2009). The distribution, drivers, and remedies for these problems have to be understood with much greater certainty
- water scarcity in key regions—water-use efficiency in agriculture is determined in part by soil management. Again, the actual and potential water use efficiency of agriculture systems is poorly characterized in many parts of the world but substantial gains may be possible in some regions (Molden, 2007; Wallace, 2000).
- climate change—current systems of land use are being disrupted by a changing climate and new patterns of land use will emerge. The sustainability of these systems will depend heavily on interactions between climate, soil, land management, genetic resources, and the policies that enable or undermine them. An ability to track these changes and anticipate consequences requires reliable information on climate change and the character and distribution of soils among other factors.

The constraints and issues listed above are neither static nor evenly spread around the world. They require careful analyses by technical specialists and appropriate responses from international organizations, governments, industries, and those managing the land. The data and information requirements for the analyses are substantial.

As noted earlier, soil data and information can be used for many purposes from local through to national and global scales. For efficiency, the systems of measurement and analysis need to be integrated across this hierarchy of scales so that data collected at lower levels flow through to analyses at the

higher levels. The dramatic advances in Web-based technology make this integration of local, national, and global systems possible. However, data and information have to be collected and managed according to consistent standards to enable integrated analysis. Taking full advantage of the digital revolution requires a significant shift within the soil science community and a much stronger commitment to developing and adhering to standards such as those outlined below and those that define the broader Global Earth Observing System of Systems (GEOSS).¹ The global grid of soil functional properties described here fits into planning for an enduring international system for observing and forecasting soil condition, which includes integration with GEOSS (McKenzie, 2014).

2.3. Current map coverage

Building a consistent global soil data set has to start with a thorough understanding of the status of thousands of data sets collected during survey programs and other activities (collectively referred to as legacy data). Globally, about two-thirds of countries have been mapped at cartographic scales of 1:1 million scale or more detailed, but over two-thirds of the total land area has yet to be mapped even at a 1:1 million scale (Nachtergaele and Van Ranst, 2003). Most of the available mapping was completed more than two decades ago, well before the widespread use of digital methods. The fact that most spatial data on soils are “out-of-date” is a major issue for soil science that is beyond the scope of this review.

There are great differences in the status of mapping (extent, scale, underlying measurement program) and this has to be accommodated in the design of a global soil information system. Experience demonstrates that this difficult task can be done with perhaps the best example being the Soil Map of the World by FAO–UNESCO, which was completed in the mid-1970s (van Baren et al., 2000). This remarkable product, completed at the height of the Cold War, has been widely used for applications including assessment of desertification, delineation of major agroecological zones, evaluation of global land degradation, calculation of population supporting capacity, and the creation of a World Reference Base for Soil Resources.

Several other global data sets have been built on the foundations of the Soil Map of the World, the Harmonized World Soil Database (HWSD) being the most recent example (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012). The HWSD has been used in a wide range of global modeling studies

¹ www.earthobservations.org/geoss.shtml.

even though its spatial resolution is relatively coarse (1 km), uncertainties are large and not quantified, and the data model is rudimentary (two-layer soil with limited information on depths). The intensive use of the HWSO is a clear signal to the soil science community that a concerted effort to build an improved global grid will be very worthwhile.

2.4. Conventional concepts of soil survey and classification

Many of the concepts underlying conventional surveys reflect the necessity at the time they were devised for manual methods of data analysis (for an overview, see [Dent and Young, 1981](#)). Mapping relied heavily on air photo interpretation, and there were various qualitative strategies for field sampling that were efficient but without the rigor of statistical design (e.g., free survey, land-system survey). Classification was central because it provided the mechanism for summarizing vast quantities of soil morphological and analytical data and it also formed the basis for extrapolation. The cartographic constraints of paper maps resulted in various conventions for depicting soil variation within mapped polygons. For example, multiple soil types and characteristic toposquences (often with estimates of areal extent) were listed for each polygon but they were not geo-referenced with sufficient accuracy to enable co-registration with other environmental data (e.g., terrain, land cover, remote sensing).

A large literature has developed around the validity of assumptions underlying conventional survey including the presumption of high correlation between soil properties, the spatial scales of soil property variation, and the reality or otherwise of sharp soil boundaries. [Heuvelink and Webster \(2001\)](#), [McKenzie and Grundy \(2008\)](#), and [Gallant et al. \(2008\)](#) provide entry points to this literature. The main drawbacks of polygon maps are as follows:

- They are static. The maps do not provide direct information on the dynamics of soil condition (e.g., rates of nutrient depletion) whereas such information is of great interest to farmers and policymakers.
- They are inflexible for quantitative studies. Such studies (e.g., food production, land degradation, carbon balance, greenhouse gas emission) generally require information on the soil's functional properties rather than a soil name.
- They imply that soil variation is abrupt and only occurs at the boundary of the mapping units.
- Some information is lost on polygon maps. As noted above, the traditional map and report presents a highly summarized account of the soils

of a region. The loss occurs because the reporting format requires information to be condensed and data to be classified.

- The information is often presented at a specific scale that is seldom useful for the particular question.
- The data model implicit in polygon maps is difficult to integrate with most other forms of natural resource data that are grid based (e.g., satellite imagery, digital elevation models, climate data).

2.5. Grids or polygons or both?

It is sometimes assumed that the problems of conventional survey can only be resolved by moving to digital systems of spatial estimation that represent individual soil properties on a fine-resolution grid. However, polygons remain useful for conveying soil information for the following reasons.

- Many decisions on land use and management require the delineation of areas with sharp boundaries either for legal or practical reasons (e.g., town planning, zoning decisions, layout of land-management systems).
- Soils are natural bodies of material associated with specific landforms or landform elements, and in some landscapes, they are best delineated using polygons because this accurately depicts physical reality (e.g., distinct sedimentary bodies such as alluvial terraces).
- There is educational value in being able to identify landscape units with distinctive patterns of soil variation that reflect landscape evolution and pedogenesis.
- Stratification of landscapes into zones with a similar evolutionary history is also valuable for digital soil mapping because the relationships between environmental covariates and soil properties are often conditional on this history—the polygons can be used as a nominal environmental covariate.

Spatial information systems for polygons and grids are complementary and both are needed from local through to global scales. There is reasonable agreement on the design of hierarchical polygon systems at the national and global level ([van Engelen and Dijkshoorn, 2012](#)). However, recent attempts at building a consistent polygon-based system have struggled to get the support and resources necessary to complete the global products.

While there have been decades of work devoted to methods for polygon-based information systems, it is only recently that a significant effort has been devoted to building operational grid-based systems. It is relatively easy to produce a grid of soil properties using the vast array of tools available for spatial analysis, but it is much harder to build analytical systems and workflows for the routine production of grids that accurately represent

the three-dimensional patterns of soil variation—[Hengl \(2012\)](#) provides one of the few examples. The technical specifications for *GlobalSoilMap* outlined here are a contribution to this much larger effort.



3. DIGITAL SOIL MAPPING AND ORIGINS OF *GlobalSoilMap*

In the early 2000s, a new impetus was given to soil mapping when statistical and mathematical techniques developed over the previous 30 years started to be used for routine soil mapping ([McBratney et al., 2003](#); [McKenzie et al., 2008](#)). Most significant was the coming together of pedology, with its focus on soil processes and field studies, and pedometrics, with its emphasis on quantitative analysis and statistics. The emerging synthesis is best represented by successive proceedings of the IUSS Working Group on Digital Soil Mapping ([Boettinger et al., 2010](#); [Hartemink et al., 2008](#); [Lagacherie et al., 2007](#); [Minasny et al., 2012](#)). These meetings quickly put debates over qualitative and quantitative approaches into context and started to capitalize on some of the major advances in environmental sensing and analysis (e.g., digital terrain analysis, airborne geophysical survey, time-series remote sensing), proximal sensing of soils (e.g., [Viscarra Rossel et al., 2011](#)), and environmental information systems.

The proposal for a new global grid of the most important soil functional properties (initially, available water capacity) started with discussions at the 2nd Global Workshop on Digital Soil Mapping in Rio de Janeiro in 2006 ([Hartemink et al., 2008](#)). The idea quickly evolved and the *GlobalSoilMap.net* Consortium was formed in December 2006. [Sanchez et al. \(2009\)](#) document this early phase and recent summaries are given by [Fisher \(2012\)](#) and [Hempel et al. \(2014\)](#). The Consortium and broader network associated with the project includes universities, research centers, development organizations, government agencies, and private enterprises around the world. The enthusiasm for the project was driven by the clear demand for a new global coverage with accurate, up-to-date, and spatially referenced soil information as expressed by the modeling community, farmers and land users, and policy and decision makers (e.g., [European Commission, 2006](#); [Hartemink and McBratney, 2008](#); [UNEP, 2007](#)).

Six key ideas have endured throughout the development of *GlobalSoilMap*.

- The final product should provide a full global coverage and be based on the best available data.

- The spatial resolution of soil data has to match and be compatible with other global environmental data sets relating to terrain, hydrology, land cover, and land use.
- Soil functional properties relating to water, carbon and nutrients are the priority.
- Every estimate for a soil property has to have an accompanying estimate of uncertainty.
- An enduring and easy-to-update soil information system with online access has to be built rather than a one-off product.
- The best available soil and environmental data have to be used to generate the estimates and this must be done in a way that respects the sovereignty of nations.

Another key issue in the design of the system was defining what primary or ancillary data had to be included in *GlobalSoilMap*. For example, should (1) soil profile data and environmental covariates used for generating the final grid (e.g., terrain variables) be transferred into a central system or (2) just the estimates of soil properties and their uncertainties provided by contributing organizations? The second option was preferred because it avoided the development of an extremely large and potentially unmanageable central database. More importantly, it affirmed the rights and responsibilities of nations as custodians of information about their sovereign resources. As it turned out, it also took advantage of the rapidly developing technology for distributed spatial information systems. As a result, *GlobalSoilMap* is designed to be a distributed system with a set of standards for Web services and a supporting community of practice rather than a single database operated by a central agency. Aspects of the information architecture of *GlobalSoilMap* are considered in later sections.



4. TECHNICAL SPECIFICATIONS OF *GlobalSoilMap*

The Technical Specifications of *GlobalSoilMap* are at the heart of the venture (Arrouays et al., 2014; *GlobalSoilMap Science Committee*, 2013). These have been developed over 5 years under the direction of the *GlobalSoilMap* Science Committee. Here, we summarize some of the theoretical and technical issues encountered along with solutions. The latest approved version outlines specifications for the estimation of soil properties at points with defined depth intervals. The next version will also include estimates for 100×100 m square blocks.

4.1. Defining the soil individual

Most national soil information systems are hard to manage, in part because of the complexity of the soil entities involved. As a result, database queries for seemingly simple requests (e.g., what is the mass of carbon in the upper 0.30 m of the soil profile across a district?) are complex and often computationally slow because they normally involve calculation of depth-weighted averages for one or more soil horizons and areal-weighted averages of polygons for one or more soil types. Differences in the primary soil entity (sometimes referred to as a soil individual or operational taxonomic unit), both within and between countries, make a significant problem even more difficult. We required an entity that avoided these problems. It also had to

- represent the complete landscape in every environment (i.e., no gaps);
- allow estimation of soil properties and their uncertainties according to a common geometric framework for the globe regardless of the source data (e.g., vastly different geometric supports, different systems of nomenclature for horizons);
- provide a computationally efficient means for data base querying and moving between scales.

We started with [Holmgren's \(1988\)](#) critique of the soil individual in soil science and several investigations into the use of mathematical functions for representing soil variation down the soil profile, particularly the use of spline functions ([Bishop et al., 1999](#); [Ponce-Hernandez et al., 1986](#); [Webster, 1978](#)). The key idea being that every grid cell would have fitted depth functions for each soil property with the system storing the parameters for each of these functions. The alternatives were to use either layers with fixed intervals, or soil horizons with defined upper and lower boundaries, or a combination of both.

In *GlobalSoilMap*, soil is referenced to a geographical point with a subtended linear coordinate for locating layers below. The volume of the soil is defined operationally depending on whether the estimate refers to a point or block (see below). The procedures for sampling, measurement, and estimation do not rely on any notion of a natural soil body (cf. the current definitions of the pedon and polypedon used in *Soil Taxonomy* or the various schemes for defining horizons).

4.1.1 Depth function

[Bishop et al. \(1999\)](#) provide a comprehensive introduction to depth functions and quadratic-smoothing splines, in particular. [Malone et al. \(2009\)](#) applied the latter technique to digital soil mapping and more examples

are presented below. Software and guidance on the fitting of splines to soil profile data are given by [Jacquier and Seaton \(2012\)](#). [Kempen et al. \(2011\)](#) outline some alternative approaches for soils with sharp contrasts between horizons; however, these can also be handled using splines but extra data preparation is necessary (see [Jacquier and Seaton, 2012](#)). [McKenzie et al. \(2004\)](#) illustrate the flexibility of the technique for representing the soil chemical and physical properties of more than 100 diverse soil types. The logic of the process is shown in [Fig. 3.1](#).

Once depth functions have been estimated for all profile points in a region, decisions still have to be made on what depth intervals to use and whether to store the parameters of the function or estimates for finely resolved layers (e.g., 10 mm layers). The latter option was initially thought to be impractical because of the data storage requirements. However, cloud computing and decreasing costs of computer memory are changing this. The Technical Specifications currently require the fitting of functions to standard soil profile data followed by estimation of soil properties for six standard depths ([Fig. 3.1](#)). These estimates can be used to generate estimates with a finer vertical resolution if required, again using the spline or some other form of depth function.

The current and anticipated advantages of using depth functions are as follows:

- they avoid the need to develop a single global system for describing soil layers and profiles
- continuous values for a soil variable are produced from source data that are often discontinuous down the profile
- national soil information systems are not disrupted because they do not need to maintain national and international standards (as is currently the case for soil classification)
- querying is computationally more efficient, particularly for calculations of material quantities integrated over different depth intervals (e.g., carbon, available water capacity)
- quadratic-smoothing splines can be constrained to be mass-preserving (see [Bishop et al., 1999](#)) and this eliminates a potential source of bias in studies involving the computation of stores and fluxes where mass balance must be preserved
- the representation of soil variation using a three-dimensional grid is readily understood by those not familiar with soil science and is compatible with data models used in allied disciplines (e.g., hydrology, atmospheric sciences, terrain analysis, remote sensing).

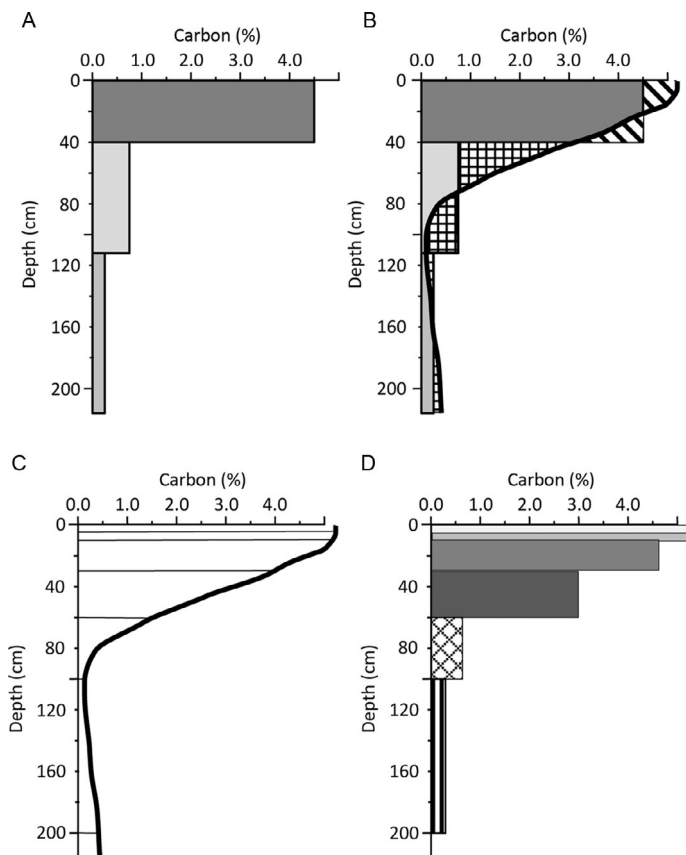


Figure 3.1 The process for estimating depth functions from soil data relating to soil layers or horizons. In this case, (A) soil data are available for three horizons. A quadratic-smoothing spline is fitted (B) and this is constrained so that it is mass-preserving (the hatched area above and below the function are of equal area for each of the original horizons). The continuous function is then used to estimate the soil property at six standard depths (C) and these are the estimates that are stored (D). In the future, estimates for finely resolved layers (e.g., 10 mm intervals) may be stored if computing resources permit.

4.1.2 Pragmatic treatment of soil depth

Estimating the depth of soil is difficult because most soil survey data sets are either censored (typically to 1 or 1.5 m depending on the equipment used for field sampling) or it is difficult to consistently identify barriers to root growth and the lower bounds of biological activity. The Technical Specifications require the estimation of soil functional properties at standard depths down to 2 m, unless the lower limit of the soil occurs at a shallower

depth. Estimates are also provided for total profile depth and plant exploitable depth, both of which may be greater than 2 m. Predictions of soil properties are not provided for grid cells occupied wholly or dominantly (>50%) by nonsoil materials (e.g., permanent water and ice, bare rock, and permanently sealed surfaces (urban areas and pavements)). No attempt is made to specify the types or proportions of nonsoil materials in a grid cell. The Technical Specifications include protocols for defining nonsoil areas.

The selection of 2 m is pragmatic and reflects arguments presented in the USDA-NRCS Soil Survey Manual (Soil Survey Division Staff, 1993, Chapter 2, page 3): “for purposes of most soil surveys, a practical lower limit of a pedon is bedrock or a depth of about 2 m, whichever is shallower. A depth of 2 m provides a good sample of major soil horizons, even in thick soil. It includes much of the volume of soil penetrated by plant roots, and it permits reliable observations of soil properties.” Where the source data are censored at <2 m, it is necessary to infer values for the intervening values. As a result, uncertainty will increase with depth in many areas.

Depth is measured from the soil surface. For mineral soils, the soil surface is the top of the mineral soil. For organic soils (or mineral soils with an O horizon), the top of any surface horizon identified as an O horizon is considered the soil surface. The soil surface is the top of the part of the O horizon that is at least slightly decomposed. Fresh leaf or needle fall that has not undergone observable decomposition is excluded when determining soil depth. For soils with a cover of 80% or more rock fragments on the surface, the depth is measured from the surface of the rock fragments (Soil Survey Division Staff, 1993, Chapter 3, page 4).

Total profile depth is the depth in centimeters to a lithic or paralithic contact. Depth to bedrock refers to the depth to fixed rock. Hard and soft bedrock are distinguished. Hard bedrock is usually indurated but may be strongly cemented and excavation difficulty would be very high or higher. Soft bedrock meets the consistence requirements for paralithic contact (Soil Survey Division Staff, 1993, Chapter 6, page 13).

Plant exploitable depth is defined as “The lower limit of soil is normally the lower limit of biologic activity, which generally coincides with the common rooting depth of native perennial plants” (Soil Survey Staff, 1975; Soil Survey Division Staff, 1993, Chapter 1, page 5).

Root depth is defined by either the evidence of the roots themselves or on the presence of barriers to root extension. The first option requires rules for root abundance to define the lower limit or inferences on the depth of roots from soil morphology. Depths may differ between biomes as given in

Table 3.1 Summary of maximum rooting depth by biome

Biome	<i>N</i>	Mean maximum rooting depth (m)	Highest value for rooting depth (m)
Boreal Forest	6	2.0 ± 0.3	3.3
Cropland	17	2.1 ± 0.2	3.7
Desert	22	9.5 ± 2.4	53
Sclerophyllous shrubland and forest	57	5.2 ± 0.8	40
Temperate coniferous forest	17	3.9 ± 0.4	7.5
Temperate deciduous forest	19	2.9 ± 0.2	4.4
Temperate grassland	82	2.6 ± 0.2	6.3
Tropical deciduous forest	5	3.7 ± 0.5	4.7
Tropical evergreen forest	5	7.3 ± 2.8	18
Tropical savanna	15	15.0 ± 5.4	68
Tundra	8	0.5 ± 0.1	0.9

After [Canadell et al. \(1996\)](#).

Table 3.1. Although logical, the approach is complex. The Technical Specifications have to be applied everywhere and not rely on complex accessory data. The second option defines the depth of a relatively easy-rooting zone from the soil surface to a root boundary. The boundary is defined by one or more morphological barriers. These barriers include massive rock, jointed rock, clean sand, pan, high-density material (bulk density > 1.85), permanent water table, extremely gravelly or densely packed gravel, and chemical toxicity.

The extremely deep rooting ability of some perennial species in dry or seasonally dry environments, where roots penetrate to great depth in jointed rock, is noted and would not be recognized in option two. The choice of option needs to consider the most likely application of the data in the region concerned. If it is to map ecosystem behavior across different biomes then option one is favored. If it is to explore opportunities for regional or global food production then the agronomic depth provided on option 2 is favored.

4.2. Definition of the grid

The geographic reference for *GlobalSoilMap* is the 3 arc-second by 3 arc-second grid exactly matching the global Shuttle Radar Topographic Mission (SRTM) DEM data set but extended to the poles. Block estimates for

100 × 100 m squares (see below) are located at the nodes of this grid. Terrain data derived from NASA's SRTM provide the source for several key covariates used in predicting soil properties via a range of methods described in [Section 5](#). Details on projections, the horizontal and vertical datum, location and tiling, unique identifiers, and reporting conventions are outlined in the Technical Specifications. These facilitate seamless compilation and ensure that there are no gaps, offsets, duplication, or edge-matching issues.

4.2.1 100 m resolution

The resolution of 100 m was one of the most debated aspects of *GlobalSoilMap*. Arguments have been presented for both finer and a coarser resolutions. However, the decision to match *GlobalSoilMap* with the resolution of the SRTM terrain data recognized that in many parts of the world the SRTM products are the only terrain data available to help with spatial estimation.

The case for a finer resolution grid (e.g., 20 or 50 m) recognizes that many landscapes have slopes with lengths of 200–500 m. A large proportion of soil variation occurs along the associated toposequences so this has to be differentiated. Realistic depiction of slopes and curvatures requires at least five cells (e.g., crest, upper slope, mid-slope, lower slope, flat) so this implies a resolution finer than 100 m in many landscapes ([Gallant, 2001](#); [Gallant and Hutchinson, 2008](#)). There is nothing to stop a country preparing its grid of soil properties at a finer resolution (and this is already happening) but the key requirement is to supply the 100 m resolution estimates for *GlobalSoilMap* to ensure consistency.

A second argument for a resolution finer than 100 m is that aggregation (up-scaling) is relatively simple, whereas downscaling is complex (e.g., [Malone et al., 2012](#)) so it is better to produce everything at the finest scale possible at the outset.

The main case for a coarser resolution (e.g., 500–1000 m) is pragmatic and opportunistic. Most environmental covariates are already available at these resolutions, the computing requirements are more tractable, and so product delivery can be achieved quickly. An example of this approach, based partly on the Technical Specifications, is presented for Sub-Saharan Africa by [Hengl et al. \(2013\)](#).

The second argument for a coarser resolution raises a more fundamental issue. Many parts of the world only have reconnaissance-scale surveys (e.g., 1:1 million cartographic scale), so generating a fine-resolution grid for these areas is claimed to be misleading. This highlights a basic difference between

conventional and digital approaches. Conventional surveys have used the cartographic and mapping scale (the latter often being approximately four times more detailed) as an integrating index of accuracy and precision. However, in digital soil mapping, each estimate for a soil property has an explicit estimate of its uncertainty. As a consequence, a fine-resolution grid produced from a 1:1 million source map would have estimates for each cell (possibly based on a spatially weighted mean of the component soils) but each cell would most likely have very large uncertainties for each of the soil properties. In other words, the resolution of the cell (10 m or 100 m) is not the measure of uncertainty—the grid provides the geometric framework for estimation and a fine-grained grid does not imply that we have accurate and precise estimates. While it takes time to change to this new way of thinking, the approach has many scientific and technical advantages, particularly in quantitative studies where the analysis of risk in decision-making is important.

4.2.2 Estimates for points and blocks

As noted earlier, the plan is to provide estimates for both points and 100×100 m blocks. In practice, the estimate of a soil property for both the point and block will be the same (most commonly a simple arithmetic average). However, the estimate of uncertainty will differ significantly. This is because a large proportion of soil variation, especially for soil chemical properties, occurs over short distances (often just a few meters). Some of the methods for estimating the uncertainty for points and blocks are outlined below.

The decision to calculate estimates using a 100 m square regardless of latitude may seem clumsy because a correction factor for the actual area of the cell on the SRTM grid is needed for calculating areal quantities. However, the alternative of calculating estimates of uncertainty for a variable-sized grid was considered to be more complex and harder to interpret. The 100 m cells have a slight overlap at the Equator (3 arc-seconds is approximately $93 \text{ m} \times 93 \text{ m}$), and they have almost complete overlap at the poles. The scheme is shown in [Fig. 3.2](#).



5. MINIMUM DATA SET

5.1. Concept

There have been many proposals for standard data sets to be collected in soil and land resource surveys. Agreement on the ideal or optimum data set is rare, even for districts with a limited range of land uses. However, it is fairly

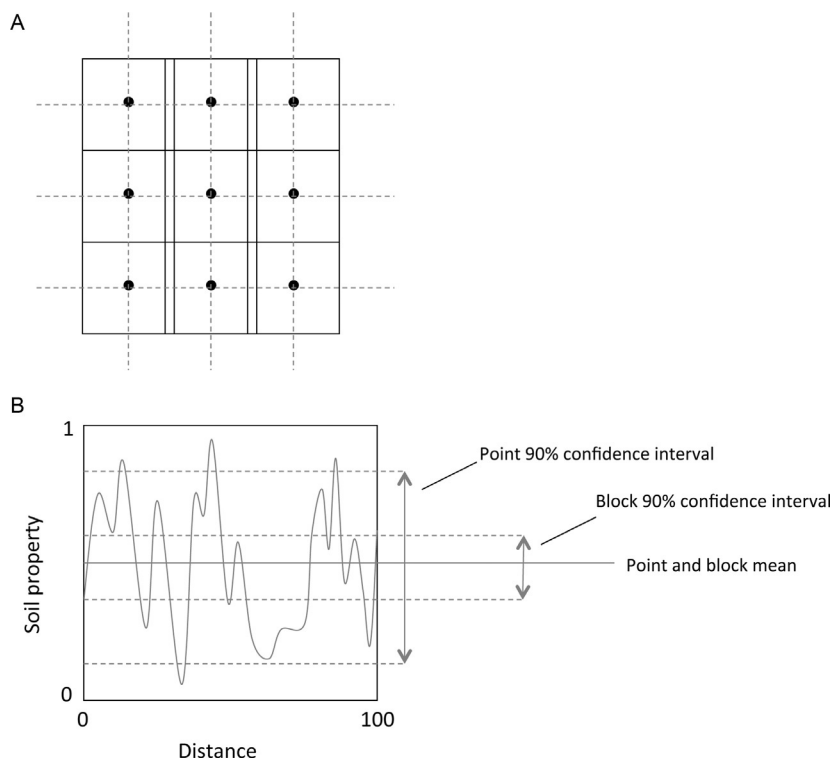


Figure 3.2 The soil property values reported are averages for a 100×100 m square located at the center point of the 3 arc-second grid; the uncertainty reported is the 90% prediction interval of this average. The diagram shows the situation for low latitudes. At high latitudes, the squares overlap to a large degree. A correction factor is used to calculate areal quantities.

straightforward to reach consensus on a minimum data set (Nix, 1984). The minimum data set does not constrain countries to generate estimates of other soil properties. Some countries are already producing estimates for other soil properties that are relevant to their particular environment or systems of land use (e.g., electrical conductivity for dry regions, total phosphorus and trace elements in strongly weathered landscapes). A soil property such as electrical conductivity is important across large areas on several continents; however, including it in the minimum data set would have added an unnecessary cost for the humid regions of the world.

The *GlobalSoilMap* specifications require the estimation of soil property values along with their uncertainty and date of estimation at each of six specified depth increments for the soil properties listed in Table 3.2. These soil

Table 3.2 The minimum data set for *GlobalSoilMap*

No.	Property and units	Precision ^a	Reference	Description of method
<i>Properties related to depth of soil</i>				
1	Depth to rock (cm)	N3.0	Soil Survey Division Staff (1993), Chapter 1 , page 5	Depth in cm to a lithic or paralithic contact as defined in USDA Soil Survey Manual. If depth is <200 cm, record actual depth in cm. If depth is >200 cm, record actual depth if known. If not known exactly, record depth as 999 cm
2	Plant exploitable (effective) depth (cm)	N3.0	Soil Survey Division Staff (1993), Chapter 3 , page 60	Effective depth in cm as defined in the USDA Soil Survey Manual. The lower limit of soil is normally the lower limit of biologic activity, which generally coincides with the common rooting depth of native perennial plants. This depth is where root penetration is strongly inhibited because of physical (including soil moisture or temperature) and/or chemical characteristics
<i>Primary soil properties</i>				
3	Organic carbon (g kg ⁻¹)	N4.0	ISO 10694	Mass fraction of carbon by weight in the <2 mm soil material as determined by dry combustion at 900 °C
4	pH × 10	N3.0	ISO 10390	1:5 soil/water (divide by 10 to get correct pH)
5	Clay (g kg ⁻¹)	N3.0	Burt (2004) , page 347	<2 µm mass fraction of the <2 mm soil material determined using the pipette method
6	Silt (g kg ⁻¹)	N3.0	Burt (2004) , page 347	2–50 µm mass fraction of the <2 mm soil material determined using the pipette method

Continued

Table 3.2 The minimum data set for *GlobalSoilMap*—cont'd

No.	Property and units	Precision	Reference	Description of method
7	Sand (g kg^{-1})	N3.0	Burt (2004) , page 347	50 μm to 2 mm mass fraction of the <2 mm soil material determined using the pipette method
8	Coarse fragments ($\text{m}^3 \text{m}^{-3}$)	N3.0	Burt (2004) , page 36	Volume fraction of the soil material >2 mm
9	ECEC ($\text{mmol}_c \text{kg}^{-1}$)	N4.0	ISO 11260	Cations extracted using barium chloride (BaCl_2) plus exchangeable H + Al
<i>Derived soil properties</i>				
10	Bulk density (Mg m^{-3})	N3.1	ISO 11272	Bulk density of the whole soil (including coarse fragments) in mass per unit volume by a method equivalent to the core method using a pedotransfer function
11	Bulk density (Mg m^{-3})	N3.1	ISO 11272	Bulk density of the fine earth fraction of the soil (<2 mm) in mass per unit volume by a method equivalent to the core method using a pedotransfer function
12	Available water capacity (mm—total over the depth range) ^b	N4.0	GlobalSoilMap Science Committee (2013) , Burt (2004) , page 137	Available water capacity computed for each of the specified depth increments using a specified pedotransfer function that references the values estimated above for organic carbon, sand, silt, clay, and bulk density
<i>Secondary soil property</i>				
13	Electrical conductivity (mS m^{-1})	N4.1		Electrical conductivity in 1:1 saturated paste

^aThe notation for precision (e.g., N3.0) is interpreted as N, number; 3, length of number; 0, number of decimal digits. Wherever possible, values are reported in integer rather than real number formats to minimize data volume.

^bAWC= f (total carbon, sand, silt, clay, % coarse fragments, bulk density) for the six depths. Profile-AWC is AWC summed over the effective depth.

properties are either directly related to soil hydrology, carbon cycling, or nutrient status, or they can be used as inputs to pedotransfer functions that can, with reasonable certainty, generate reliable estimates relevant to the same processes (e.g., available water capacity, bulk density). At present, individual countries have responsibility for the selection and use of pedotransfer functions although a more general global system for soil inference is being developed (Morris et al., 2012).

Definitions and methods of analysis for most of the soil properties are according to ISO standards but particle size distribution is defined according to the USDA Soil Survey Laboratory Methods Manual (Burt, 2004). Many methods are used to measure the soil properties included in the minimum data set (e.g., 80 methods for soil carbon and 18 for pH). All soil properties are translated to a standard method and the Technical Specifications provide guidance on how to do this. Units for properties are reported as integers to reduce the cost of data storage and transmission (i.e., g kg^{-1} instead of % and/or cm instead of m).

5.2. Time

The date of the actual or estimated time of sampling of the source data is attached to each of the estimated soil properties at each grid cell. The date reported is the year of publication for a map or the year of analysis for a sampled soil profile.

The *GlobalSoilMap* project is making maximum use of legacy soils data collected and reported over many decades of field work. Data for any point or any map reflect the state of the soil at the time the point was sampled and analyzed, or the map was produced. A gridded date-map will be made to indicate the date (in years) that the soil property value most closely reflects. It may be possible in future versions (beyond Version 2) to reconcile differences in soil property values reported for different times and under different land uses to one or more standardized reference dates (e.g., harmonized decadal values at 1970, 1980, 1990, 2000, 2010) and under the land-use conditions current at each date. This will first require that regional legacy soil data sets be analyzed to detect and quantify directions and rates of change in soil property values under known land use and land-management regimes. These regional values for rates of change under different land uses could be applied to the original predictions of soil property values, in combination with information on land-use history at each grid cell, to harmonize soil property values to common reference years for each major regional land-use type.



6. METHODS FOR ESTIMATING SOIL PROPERTIES USING LEGACY DATA

Minasny and McBratney (2010) provide an introduction to the range of methods available for global digital soil mapping and this section draws on their account. They distinguish between circumstances where only existing “legacy data” are available and circumstances where new survey and measurement programs are planned.

6.1. Legacy data only

Figure 3.3 is a decision tree for methods in the case when only legacy data are to be used for estimating soil attributes. The options are considered from the most data-rich through to data-poor situations.

6.1.1 Detailed soil maps with legends and soil point data

This situation leads to the best prediction of soil properties. Soil properties can be derived from both detailed soil maps (generally a cartographic scale of

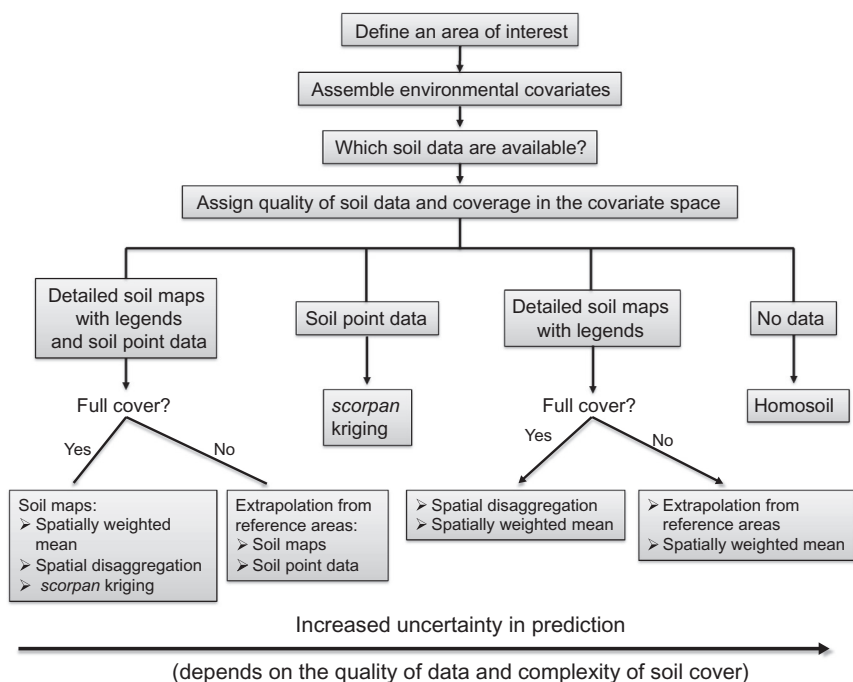


Figure 3.3 A decision tree for digital soil mapping based on legacy data. *Adapted from Minasny and McBratney (2010).*

1:100 000 or more detailed) and soil point data (i.e., measurements down the soil profile at a geo-referenced location). The available methods are extracting soil properties from soil map using a spatially weighted measure of central tendency (e.g., the mean), spatial disaggregation of soil maps, *scorpan* kriging (see below), or a combination of these methods.

6.1.2 Soil point data

When sufficient soil point data are available, soil properties can be interpolated and extrapolated to the whole area by using empirical statistical modeling. This is referred to as the *scorpan* approach and an example for Denmark is provided by [Adhikari et al. \(2013\)](#). The approach can be extended to include an explicit stochastic spatial component where the residuals from the correlation model are modeled by variograms and a form of kriging. This is referred to as either regression kriging, kriging with external drift, or *scorpan* kriging ([McBratney et al., 2003](#)). Examples include [Odeh et al. \(1994, 1995\)](#) and [Hengl et al. \(2004\)](#).

6.1.3 Detailed soil maps with legends

When only soil maps are available, soil properties can be extracted from soil maps according to the distributional concepts underlying the soil-mapping units. In some cases, it will be appropriate to estimate soil properties using an area-weighted mean. An example for the United States is provided by [Ogders et al. \(2012\)](#). However, in most circumstance, the original soil map will have information on the factors controlling soil distribution within an individual map unit. This is most commonly based on terrain (e.g., a catena or characteristic toposequence). The widespread availability of fine-resolution terrain variables now allows the soil properties to be “disaggregated.” Examples of the approach are provided by [Thompson et al. \(2010\)](#), [Wei et al. \(2010\)](#), [Nauman et al. \(2012\)](#), and [Kerry et al. \(2012\)](#).

An extension of this approach is to use *reference areas* where there is a detailed understanding of soil distribution and this model is used as a basis for extrapolation to a broader domain. This is effectively another form of disaggregation and examples include [Lagacherie et al. \(2001\)](#), [Bui and Moran \(2001, 2003\)](#) and [Grinand et al. \(2008\)](#).

6.1.4 No local data—homosoil

This approach is required when no detailed map or soil observations are available in the region of interest ([Mallavan et al., 2010](#)). The method is based on the assumed homology of soil forming factors between a reference

area and the region of interest. In practice, data on climate, terrain, and parent material are most commonly used because of its availability. This approach involves seeking the smallest taxonomic distance of the *scorpan* factors between the region of interest and other reference areas (with soil data) in the world. The rules calibrated in the reference area are applied in the region of interest, realizing its limitations and extrapolation consequences.

6.2. Acquiring new data

The methods for estimating soil properties outlined above have been refined over two decades of research into digital soil mapping. This has involved experimentation with a wide range of geostatistical, data-mining, and expert-based approaches across diverse landscapes. This collective experience has highlighted that the first step in generating digital soil data is to take full advantage of the large quantities of legacy soil data that have been collected during previous decades of soil and land resource survey. The second step in building the global map of soil properties is to focus on a new phase of soil survey using reasoned sampling and new measurement technologies.

6.2.1 Sampling

When resources are available for acquiring new data to improve spatial estimates of soil properties, we need a reliable sampling design to determine where new sampling units should be located. In most cases, it will be profitable to include legacy data into the design process. However, the legacy data available for a given area will have often been sampled at different times and for various purposes. Bias is a significant issue with some areas being relatively over- or under-sampled. The decision tree in [Fig. 3.3](#) provides guidance on how to proceed. The options are listed below starting with the case where only a limited sampling program is required through to the case where a major survey is necessary. More general reviews are provided by [Webster and Oliver \(2007\)](#), [de Gruijter et al. \(2006\)](#), and [Webster and Lark \(2013\)](#).

- *Detailed soil maps with point soil observations:* The hypercube evaluation sampling method ([Carré et al., 2007b](#)) can be applied to inspect the coverage of the *scorpan* variables. New sampling units are first placed in the strata with no sampling units and considering the density of covariates, to ensure that the hypercube is as maximally occupied as possible.
- *Soil point data:* The method of [Brus and Heuvelink \(2007\)](#) selects sites that minimize the variance of universal kriging using simulated annealing. This method assumes that the predictors are linearly related

to the target variable and that the variogram of the residuals is known. This method optimizes sampling in both predictor space and geographic space. However, this method can be limited in practice as it assumes linearity of the prediction function and knowledge of the residual variogram.

- *Detailed soil maps:* Soil samples are needed to cover both representation of the soil mapping units and *scorpan* variables. The conditioned Latin hypercube sampling (cLHS) method (Minasny and McBratney, 2006) can be used in this instance. cLHS attempts to cover the range of values of each of the *scorpan* factors.
- *No prior soil data:* The Latin hypercube sampling method (Minasny and McBratney, 2006) or fuzzy k-means clustering can be applied to cover both the spatial coverage and *scorpan* variables. The fuzzy k-means method (de Gruijter et al., 2008) classifies the *scorpan* variables into k classes, with k equal to the number of sampling units. The pixel with the largest membership for each class is selected as the sampling unit.

6.2.2 New measurement technologies

Until recently, measurement technologies in soil survey were limited to conventional morphological description, invasive techniques for specimen collection, and expensive wet laboratory techniques. This has changed significantly in the past decade due to the emergence of rapid measurement technologies for the laboratory and field. Viscarra Rossel et al. (2011) provide a recent review. These technologies have the potential to revolutionize soil survey, but before this can happen, research methods have to be made operational and general agreement is then needed on standardized systems to ensure the generation of harmonized data sets globally.



7. UNCERTAINTY

Heuvelink (2014) provides an account of the quantification of uncertainty in *GlobalSoilMap* products. In the technical specifications for *GlobalSoilMap*, uncertainty is defined as the 90% Prediction Interval (PI) which reports the range of values within which the true value is expected to occur 9 times out of 10, with a 1 out of 20 probability for each of the two tails. Note that this does not necessarily imply that the PI is symmetric around the predicted value.

The general framework for assessing and representing uncertainties in environmental data provided by Brown (2004) has been followed.

Heuvelink and Brown (2007) observed that “soil data are rarely certain or ‘error free’, and that these errors may be difficult to quantify in practice.” Indeed, the quantification of error (defined here as a “departure from reality”) implies that the “true” state of the environment is known. They reported that “in recent years, a distinct spectrum of methods, not altogether statistical, has emerged for dealing with situations of ‘imperfect knowledge’ in scientific research (see Ayyub, 2002 also).” A spectrum of methods for uncertainty analysis is indeed important with due consideration of the potential sources of uncertainty—namely, from inputs (observed data and covariate information), model parameters, and model structure. Similarly, methods of uncertainty analysis will vary on the basis of whether soil point data or existing soil maps are used for producing outputs (Section 6). Moreover, methods will also vary depending on the density of point data as well.

When there are sufficiently many point observations, there are two general approaches to assess the uncertainty of digital soil maps:

1. Statistical modeling (principally geostatistical models) of the soil properties directly. The uncertainty of predictions is generated from the model as a byproduct. A typical example is kriging, where the kriging predictions are used as predictions of the soil property of interest and where the kriging standard deviation characterizes the prediction uncertainty, from which PIs can be computed.
2. Statistical modeling (principally geostatistical models) of residuals from independent data set or resampling techniques. This approach would typically be used in case of an existing soil property map or when a soil property map is produced using nonstatistical models. In such case, the uncertainty of the map can be assessed by comparison of map predictions with independent observations, either by calculating quality measures such as the Mean Error and Root Mean Squared Error, or by geostatistical modeling of the error (i.e., derive a variogram and kriging the difference between the map predictions and independent observations).

Various methods for validation of soil maps have been reviewed by Brus et al. (2011). Data splitting is often used. The main advantage of this method is that it does not require new field work. A recent example on validation of a soil map by data-splitting is given by Grinand et al. (2008). However, Brus et al. (2011) stress that there is no guarantee of unbiased estimates. Another option is cross-validation. In the leave-one-out cross-validation, for each sampling location, the model is refitted leaving that location out of the

calibration data set. Brus et al. (2011) showed that unbiased and valid estimates of the quality measures can best be obtained by selecting additional test units (measurements at point locations or pixels not used for calibration) by probability sampling. Using this method, validation data are truly independent from the predictions and yield unbiased estimates of validation measures. Furthermore, using probability sampling allows for the calculation of the confidence limits associated with the validation measures and to test whether a more elaborate or novel method produces more accurate results than an existing approach. However, this requires additional field work, soil analyses, and related costs.

An alternative method to estimate PI has been presented and described in Malone et al. (2011). Here, uncertainty is treated as the probability distribution of the output model errors, which comprises all sources of uncertainty (model structure, model parameters, and input data). And since it is estimated through an empirical distribution, it is not necessary to make any assumption about residuals (Solomatine and Shrestha, 2009). This method is particularly useful when dealing with soil spatial prediction functions that include data-mining tools or neural networks (as examples) in combination with the regression kriging approach, where it would be difficult to use other existing methods (of uncertainty analysis) to estimate uncertainty. However, the approach of Malone et al. (2011) requires that there be a sufficient number and density of point observations within any given prediction area (e.g., 30 per class) to support a data-driven assessment of the probability distribution function of a given soil property by class within the geographic extent of the area of interest.

If sufficient information does not exist to support conventional statistical analysis, the PI may be assessed by appropriate local or national experts using formal expert elicitation procedures. Fuzzy logic (Cazemier et al., 2001) and Bayesian beliefs (O'Hagan et al., 2006) have been proposed as suitable frameworks for establishing estimates of uncertainty in the absence of sufficient hard field data. Lilburne et al. (2009) presented a method based on using expert knowledge to estimate the probability distribution function in situations where there is insufficient information to support conventional statistical analysis.

It is a goal of the *GlobalSoilMap* that all output products can be replicated or reproduced, given access to the inputs used to generate them. Achieving this reproducibility requires that each reported soil property has an associated date and estimated uncertainty and full documentation of the methods used to produce these values.



8. INFORMATION ARCHITECTURE, WEB SERVICES, AND APPLICATIONS

Web-based delivery of soil information requires a significant investment into the expertise and facilities for analyzing and managing large and complex information resources. *GlobalSoilMap* is being developed within the principles and practices developed by GEOSS. This provides a logical framework for building and updating *GlobalSoilMap*. It also provides the vehicle for delivering soil information to a wide range of user groups. Before this can occur, the soil science community has to build specific components within the broad architecture provided by GEOSS.

An agreed data sharing format for soil information is a prerequisite. Related disciplines have already developed these standards (e.g., WaterML for hydrology). SoilML is such a standard and it is being developed by the IUSS Working Group on Soil Information Standards. It will eventually allow owners of soil data to publish via the Web in a way that maximizes reuse of the data.

The second major task is to develop Web services for *GlobalSoilMap* and related soil information. Web services deliver data encoded in XML format and they enable the integration of data and modeling systems online. The development of Web services for *GlobalSoilMap* is just starting and there is enormous potential to integrate soil data with other online services (e.g., systems for predicting crop yield; calculators to assess potential for carbon sequestration, input to hydrological modeling). An example of a soil information application that relies on such Web services is SoilMapp.²

A third task for *GlobalSoilMap* is to gain agreement on which countries and institutions host *GlobalSoilMap* and provide the computing infrastructure necessary to maintain the Web services. A single global database without direct links to national or regional databases is not practical because countries are the main providers of data. At the other extreme, a federated international system with more than 190 nation states providing Web services for their jurisdictions is also unlikely to be successful. The main problem being that many countries do not have the technical capacity and infrastructure for providing updated soil information and reliable Web services into the global system. A middle course is therefore necessary. It requires a federated system with a mix of arrangements for delivering

² www.csiro.au/soilmapp.

Web services tailored to match the capabilities of individual countries or groups of countries. In some parts of the world, good collaborative arrangements will allow a leading country or institution to provide the soil information system and services on behalf of several other countries. Such arrangements already exist in some regions (e.g., the European Soil Data Center³) or are being developed. The feasibility of the approach is demonstrated by projects such as OneGeology⁴ where several agencies take responsibility for managing the world portal for geological data.

The aim is to ensure soil data and information is freely available on the Web and in a format that can be readily used for a wide range of purposes. This availability will stimulate the use of soil information and result in many new applications. The Web-based architecture will also create the opportunity for new sources of soil data to be efficiently shared (e.g., from new sensors such as infrared spectroscopy and *in situ* monitoring systems).



9. EXAMPLES

Here, we present several examples of *GlobalSoilMap* products from around the world. Minasny et al. (2012) and Arrouays et al. (2014) provide further examples along with a broader consideration of methodological issues.

Figure 3.4 presents estimates of particle size for Denmark produced by Adhikari et al. (2013). In this case, the estimates have been generated using the *scorpan* approach based on a comprehensive set of point data (1958 soil profiles). The Cubist data mining tool was used and environmental covariates included the national soil map, digital terrain variables, geological data, and land use.

Figure 3.5 presents estimates of percent silt at the 5–15 cm interval for the district of Nabeul in Northern Tunisia. Estimates for the complete minimum data set at the six standard depths have been generated for the 2822 km² study area. Linear regression, ordinary kriging, and regression kriging were used to estimate spatial variation. The study area is sparsely sampled and does not have a large range in silt content and short-range variation is significant. As a consequence, only a small proportion of the variation could be accounted for. The quantitative estimates of uncertainty provide a basis for a targeted sampling plan if more accurate and precise estimates are required.

³ eusoils.jrc.ec.europa.eu.

⁴ www.onegeology.org.

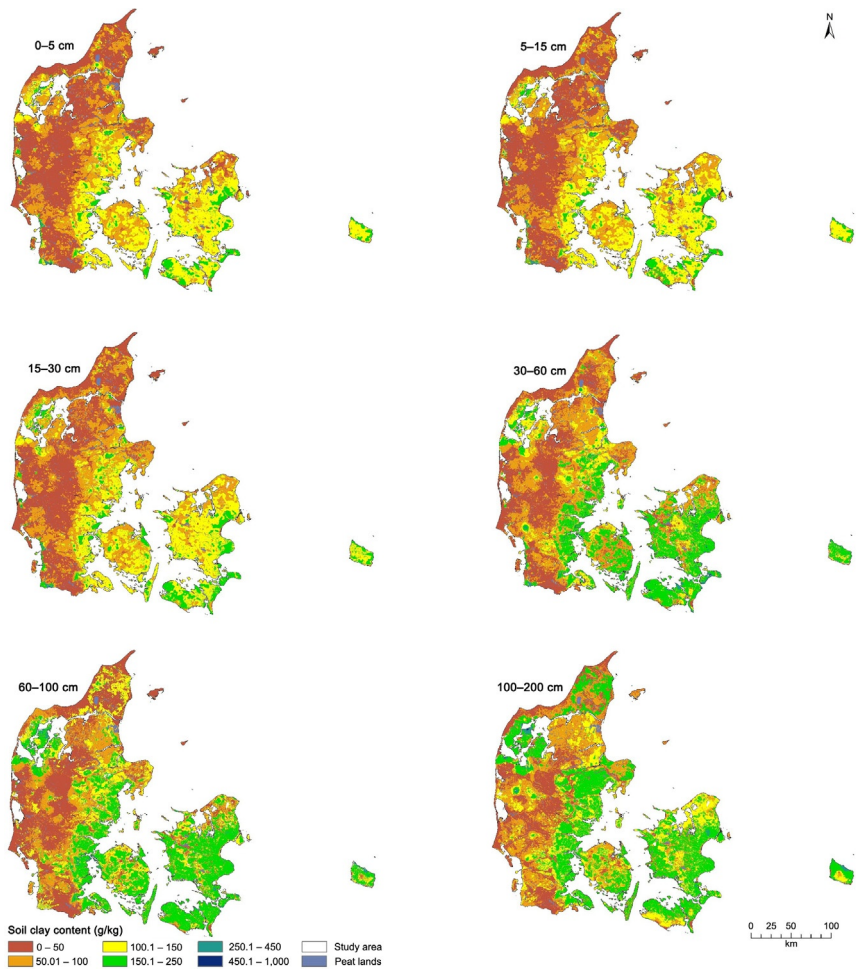


Figure 3.4 Predicted clay content of Danish soils at the six standard depths (Adhikari et al., 2013).

Figure 3.6 presents estimates of available water capacity for Korea. In this case, the estimates have been derived from detailed soil series mapping available for the complete country at a cartographic scale of 1:25,000. Each mapping unit is assumed to be represented by a soil series and estimates of soil properties were generated from the modal profiles for each soil class. Hong et al. (2012) provide related examples for organic carbon and clay content based on the

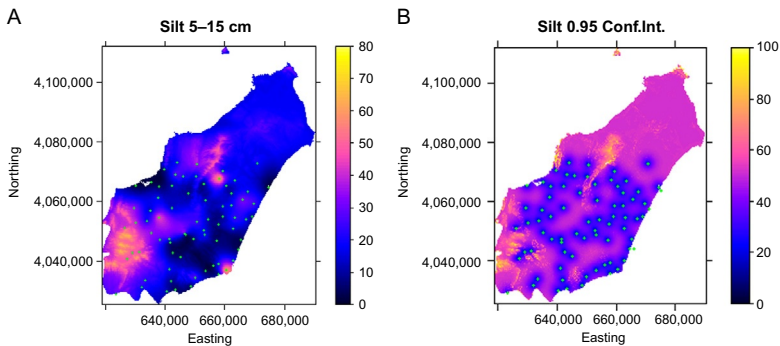


Figure 3.5 Estimates of (A) silt content (%) and (B) the 95% confidence interval, for the 5–15 cm layer for 2822 km² in the district of Nabeul in Northern Tunisia.

Technical Specifications. [Hong et al. \(2013\)](#) outline the procedures employed to estimate available water capacity using pedotransfer functions.

[Figure 3.7](#) presents estimates of soil organic carbon for the United States at the standard depth intervals and 100 m grid resolution (after [Odgers et al., 2012](#)). These estimates have been derived from the State Soil Geographic (STATSGO2) database based on the comprehensive and detailed county soil surveys published by the United States Department of Agriculture Natural Resource Conservation Service. The estimates have been produced using area-weighted means and work is underway to enhance these estimates using spatial disaggregation within the individual map polygons (e.g., [Nauman and Thompson, 2014](#)). [Figure 3.7](#) includes estimates of the standard deviation of organic carbon content. [Figure 3.8](#) shows analogous estimates for test areas in Canada. In this case, the source data come from less-detailed land resource surveys.

[Figure 3.9](#) presents initial estimates of pH for Nigeria produced by [Odeh et al. \(2012\)](#). This example is based on sparse data and *scorpan* kriging. Estimates of pH for upper (0–5 cm) and lower (100–200 cm) layers provide a clear visualization of general trends in soil pH down the profile for the country. [Odeh et al. \(2012\)](#) provide estimates at each standard depth.

The examples provide static snapshots of *GlobalSoilMap* products. The real power of the system will become evident when software interfaces to the Web services have been completed that enable:

- flexible querying online (e.g., how much carbon is in the upper 0.5 or 1.0 m of the soil across a district?);

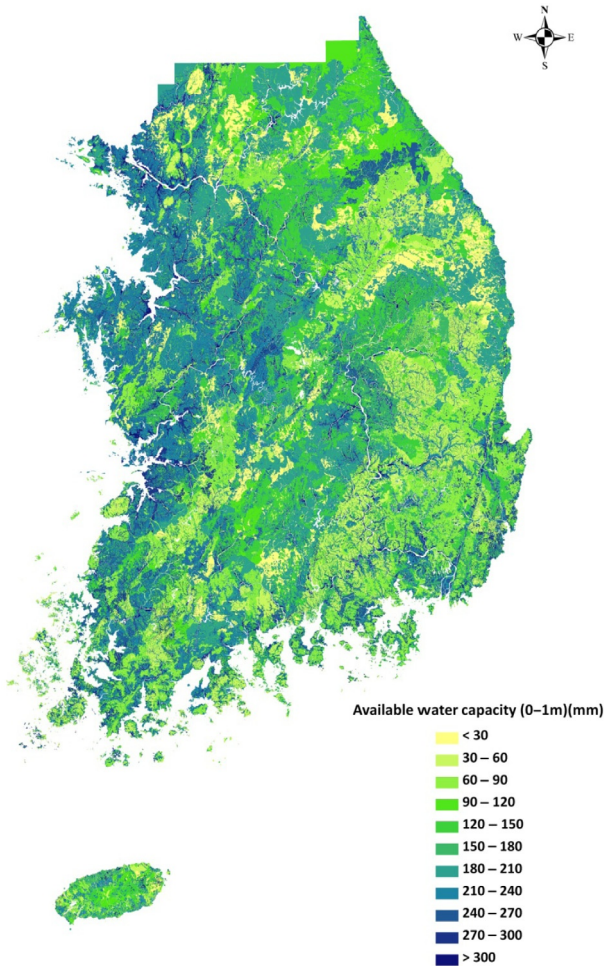


Figure 3.6 Estimated available water capacity for Korea. After [Hong et al. \(2013\)](#).

- inference of soil functional properties (i.e., the soil properties are used to estimate permeability or plant-available water capacity) via systems such as SINFERS ([Morris et al., 2012](#));
- Web services from *GlobalSoilMap* and other sources (e.g., from sensor networks, monitoring systems, remote sensing) to be automatically available to Web-based simulation models that provide forecasts and analyses in near real time. The recent advances in water services ([Salas et al., 2012](#)) provide an indication of the possibilities for applications involving soil data.

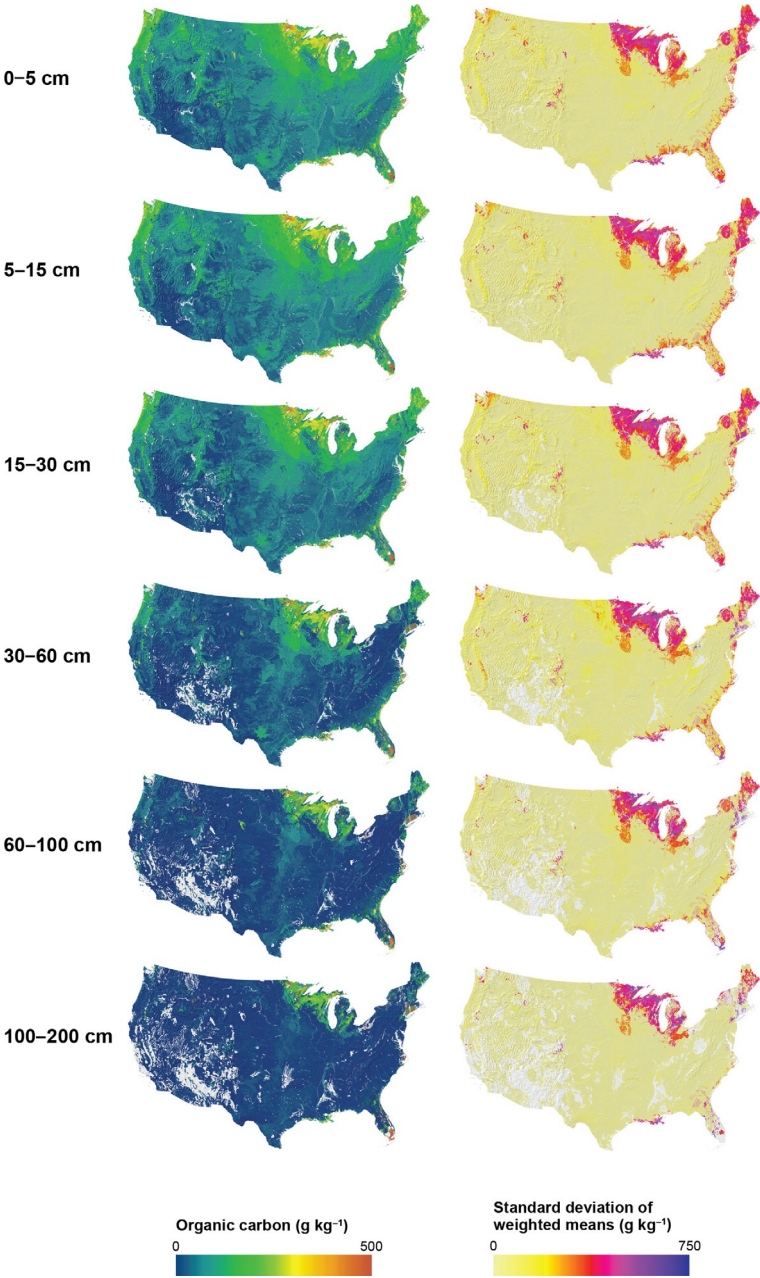
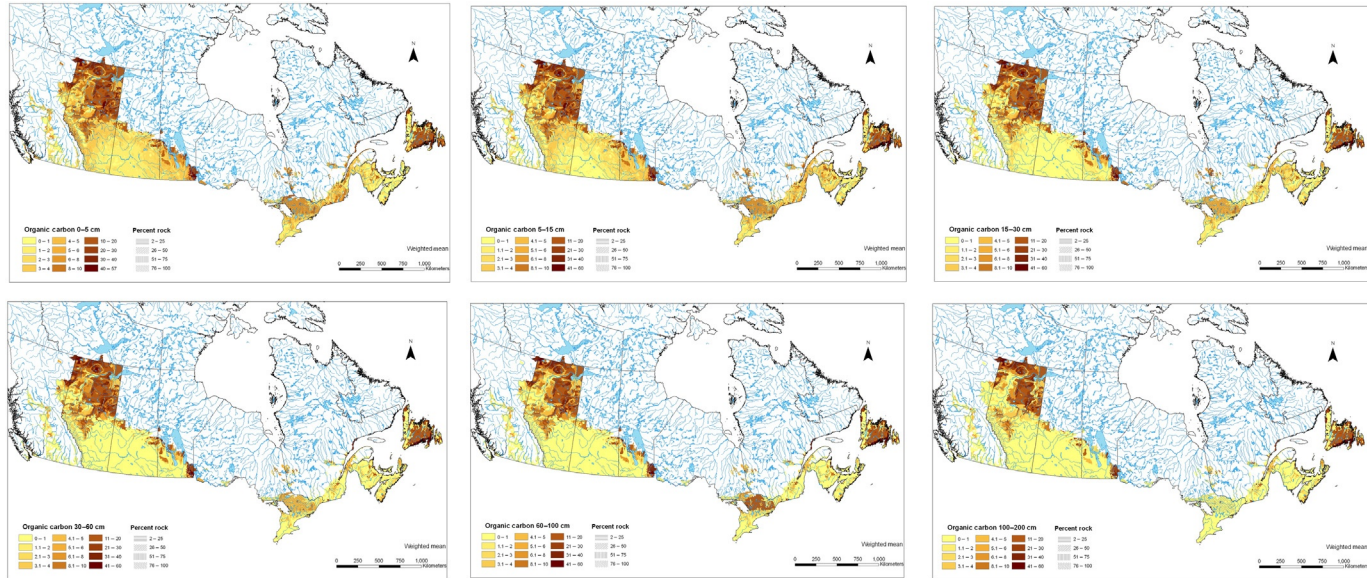


Figure 3.7 Maps of organic carbon content (left) and standard deviation of organic carbon (right) generated by [Odgers et al. \(2012\)](#) in accord with the *GlobalSoilMap* Technical Specifications.



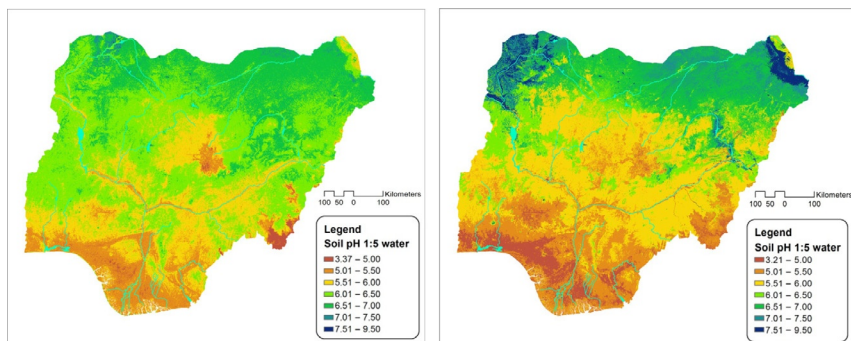


Figure 3.9 Estimates of pH according to the Technical Specifications for the upper (0–5 cm) and lower (100–200 cm) layers across Nigeria.

10. TOWARD AN ENDURING GLOBAL SOIL INFORMATION SYSTEM

The scientific and technical challenges of *GlobalSoilMap* are substantial but so too are the institutional aspects. Some of the institutional issues facing *GlobalSoilMap* are an inevitable consequence of soil information institutions in some parts of the world having limited experience in building collaborative networks (e.g., compared to disciplines dealing with, for example, climate, oceans and the geosciences). Other issues are a consequence of the global disinvestment in agriculture sciences, and soil science, in particular, between 1980 and 2010. This is not the place for a detailed analysis of such institutional matters. However, three matters deserve mention because they will ultimately determine whether *GlobalSoilMap* can achieve its desired comprehensive global coverage within the next decade.

10.1. Governance

Developing a global soil information system, with key components such as *GlobalSoilMap*, requires some form of governance or set of institutional agreements to establish standards, achieve quality control, and ensure completion in a reasonable time frame and ongoing support. To date, *GlobalSoilMap* has relied on voluntary mechanisms and the generous support of a relatively small number of institutions, governments, and donor agencies. A more formal mechanism is needed if the venture is to fulfill its potential. Other disciplines have such mechanisms and they include United Nations

agencies (e.g., the World Meteorological Organization for weather, climate, and water), various forms of multilateral arrangements (e.g., Future Earth), formal committees (e.g., Committee for Earth Observation Systems for remote sensing), and the broader arrangements through the Group on Earth Observations. Governance structures are being developed through the United Nations Food and Agriculture Organization's Global Soil Partnership and its associated Intergovernmental Technical Panel on Soils (McKenzie, 2014). This is a promising development for soil science but it will take several years before an assessment can be made on whether this provides an appropriate and robust governance mechanism for global soil information.

10.2. Mutual benefit

A second institutional challenge facing *GlobalSoilMap* and related ventures is finding ways to ensure all participants obtain benefits. There are significant costs associated with the delivery of data from national systems into the global system. However, if managed well, the global program can generate a range of benefits for participating countries and this provides the essential incentive for engagement. Some of the primary benefits of being a partner (rather than just a contributor) include the following:

- Adoption of standards for system design and Web-based delivery will save costs and avoid duplication for each country.
- Adoption of data formats supported by the global system will provide access to a range of third-party tools such as farming systems models, hydrologic models, and other Web-based services.
- Participation in the broader international scientific and technical community generates benefits for each country through training, mentoring, and capacity development.
- Countries are able to ensure that the international assessments of soil health for their jurisdiction are based on the best available data rather than outdated or incomplete data sets (as is often the case now).
- Better decision-making based on the knowledge of current and forecast soil conditions.

10.3. Moving from research to operational implementation

GlobalSoilMap has provided a fertile research agenda for participating scientists and their organizations (e.g., Arrouays et al., 2014). This research has resolved many issues and identified new questions. However, a natural evolution is occurring and the venture is moving from a primarily research phase

into operational implementation. This brings with it a new set of tasks and the need for some institutions to take responsibility for the long-term management of the operational system. New teams with different skills are required and there has to be a shift in culture from risk taking (an essential characteristic for a productive scientific organization) to risk aversion (an essential characteristic of an operational agency providing highly reliable information services 24 h a day). While there are many outstanding soil research organizations around the world, the number of operational soil information agencies has decreased in recent decades. This is a significant issue that may constrain the implementation of *GlobalSoilMap*.



11. CONCLUSIONS

Unprecedented demands are being placed on the world's soil resources. Responding to these challenges requires reliable information because we manage what we measure. The proof-of-concept studies for *GlobalSoilMap* and the associated development of the technical specifications have demonstrated how to integrate the best available data from local and national sources and provided directions on how to deliver the information online as part of the global Earth observing system. *GlobalSoilMap* provides the means for supplying soil information in a format and resolution compatible with other fundamental data sets from remote sensing, terrain analysis, and other systems for mapping, monitoring, and forecasting biophysical processes. Building the operational version of *GlobalSoilMap* is a grand challenge for soil science and its inauguration cannot come soon enough.

ACKNOWLEDGMENTS

We are very grateful to all our colleagues across the globe who worked daily on this project and have provided data and insights for this chapter. We thank them all.

REFERENCES

- Adhikari, K., Kheir, R.B., Greve, M.B., Böcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B., Greve, M.H., 2013. High-resolution 3-D mapping of soil texture in Denmark. *Soil Sci. Soc. Am. J.* 77, 860–876.
- Arrouays, D., McKenzie, N.J., Hempel, J.W., de Forges, A., McBratney, A.B. (Eds.), 2014. *GlobalSoilMap. Basis of the Global Spatial Soil Information System. Proceedings of the First GlobalSoilMap Conference, Orleans, France, 7–9 October.* Taylor & Francis, Balkema, The Netherlands.
- Arrouays, D., McBratney, A.B., Minasny, B., Hempel, J.W., Heuvelink, G.B.M., MacMillan, R.A., Hartemink, A.E., Lagacherie, P., McKenzie, N.J., 2014. The *GlobalSoilMap* project specifications. In: Arrouays, D., McKenzie, N.J., Hempel, J.W., Richer de Forges, A., McBratney, A.B. (Eds.), *GlobalSoilMap. Basis*

- of the Global Spatial Soil Information System. Proceedings of the First GlobalSoilMap Conference, Orleans, France, 7–9 October 2013. Taylor & Francis, Balkema, The Netherlands.
- Ausubel, J.H., Wernick, I.K., Waggoner, P.E., 2012. Peak farmland and the prospect for land sparing. *Popul. Develop. Rev.* 38 (suppl.), 221–242.
- Ayyub, B.M., 2002. Elicitation of Expert Opinions for Uncertainty and Risks. CRC Press, Boca Raton, Florida.
- Bishop, T.F.A., McBratney, A.B., Laslett, G.M., 1999. Modeling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma* 91, 27–45.
- Blum, W.E.H., 1993. Soil protection concept of the Council of Europe and integrated soil research. In: Eijsackers, H.J.P., Hamers, T. (Eds.), *Soil and Environment*, vol. 1. Kluwer Academic Publisher, Dordrecht, pp. 37–47.
- Boettinger, J.L., Howell, D., Moore, A.C., Hartemink, A.E., Kienast-Brown, S. (Eds.), 2010. *Digital Soil Mapping—Bridging Research, Environmental Application, and Operation*. Springer, Dordrecht, pp. 137–149.
- Bouma, J., 1989. Using soil survey data for quantitative land evaluation. *Adv. Soil Sci.* 9, 177–213.
- Bouma, J., 2002. Land quality indicators of sustainable land management across scales. *Agric. Ecosyst. Environ.* 88, 129–136.
- Brown, J.D., 2004. Knowledge, uncertainty and physical geography: towards the development of methodologies for questioning belief. *Trans. Inst. Br. Geogr.* 29, 367–381.
- Brus, D.J., Heuvelink, G.B.M., 2007. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma* 138, 86–95.
- Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. *Eur. J. Soil Sci.* 62, 394–407.
- Bui, E.N., Moran, C.J., 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. *Geoderma* 103, 79–94.
- Bui, E.N., Moran, C.J., 2003. A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray–Darling basin of Australia. *Geoderma* 111, 21–44.
- Burt, R. (Ed.), 2004. *Soil Survey Laboratory Methods Manual: Soil Survey Investigations Report No. 42, Version 4.0*. United States Department of Agriculture, Natural Resources Conservation Service, November, 2004, 700 pp.
- Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D., 1996. Maximum rooting depth of vegetation types at the global scale. *Oecologia* 108, 583–595.
- Carré, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007a. Digital soil assessments: beyond DSM. *Geoderma* 142, 69–79.
- Carré, F., McBratney, A.B., Minasny, B., 2007b. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. *Geoderma* 141, 1–14.
- Cazemier, D.R., Lagacherie, P., Martin-Clouaire, R., 2001. A possibility theory approach for estimating available water capacity from imprecise information contained in soil databases. *Geoderma* 103, 113–132.
- de Gruijter, J.J., Brus, D., Bierkens, M., Knotters, M., 2006. *Sampling for Natural Resource Monitoring*. Springer, Berlin.
- de Gruijter, J.J., McBratney, A.B., Taylor, J., 2008. Sampling for high resolution soil mapping. In: *Global Workshop on High Resolution Digital Soil Sensing & Mapping*. The University of Sydney, Australia, 5–8 February.
- Dent, D., Young, A., 1981. *Soil Survey and Land Evaluation*. Allen & Unwin, London.
- European Commission, 2006. *Soil Protection—The Long Story Behind the Strategy*. Office for Official Publications of the European Communities, Luxembourg.
- FAO, 1976. *A Framework for Land Evaluation: Soils Bulletin 32*. FAO, Rome.
- FAO, 2011. *The State of the World's Land and Water Resources for Food and Agriculture (SOLAW)—Managing Systems at Risk*. FAO, Earthscan, Rome, London.

- FAO, 2012. World agriculture towards 2030/2050: the 2012 revision. In: Alexandratos, N., Bruinsma, J. (Eds.), *ESA Working Paper No. 12–03*, June. FAO, Rome.
- FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (Version 1.2). FAO, IIASA, Rome, Italy; Laxenburg, Austria.
- Fisher, M., 2012. Creating a global digital soil map. *CSA News*, October, 4–9.
- Gallant, J.C., 2001. Topographic Scaling for the NLWRA Sediment Project: CSIRO Land and Water Technical Report 27/01. CSIRO Land and Water, Canberra.
- Gallant, J.C., Hutchinson, M.F., 2008. Digital terrain analysis. In: McKenzie, N.J., Grundy, M.J., Webster, R., Ringrose-Voase, A.J. (Eds.), *Guidelines for Surveying Soil and Land Resources. Australian Soil and Land Survey Handbook*, vol. 2. CSIRO Publishing, Melbourne.
- Gallant, J.C., McKenzie, N.J., McBratney, A.B., 2008. Scale. In: McKenzie, N.J., Grundy, M.J., Webster, R., Ringrose-Voase, A.J. (Eds.), *Guidelines for Surveying Soil and Land Resources. Australian Soil and Land Survey Handbook*, vol. 2. CSIRO Publishing, Melbourne.
- GlobalSoilMap Science Committee, 2013. GlobalSoilMap specifications—Tiered GlobalSoilMap products. Release 2.3.
- Grinand, C., Arrouays, D., Laroche, B., Martin, M.P., 2008. Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. *Geoderma* 143, 180–190.
- Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. *Soil Sci. Soc. Am. J.* 75, 1201–1213.
- Hartemink, A.E., McBratney, A.B., 2008. A soil science renaissance. *Geoderma* 148, 123–129.
- Hartemink, A.E., McBratney, A.B., Mendonça-Santos, M.L. (Eds.), 2008. *Digital Soil Mapping with Limited Data*. Springer, Dordrecht.
- Hempel, J.W., McBratney, A.B., Arrouays, D., McKenzie, N.J., Hartemink, A.E., 2014. GlobalSoilMap project history. In: Arrouays, D., McKenzie, N.J., Hempel, J.W., Richer de Forges, A., McBratney, A.B. (Eds.), *GlobalSoilMap. Basis of the Global Spatial Soil Information System. Proceedings of the First GlobalSoilMap Conference*, Orleans, France, 7–9th October 2013. Taylor & Francis.
- Hengl, T., 2012. Global Soil Information Facilities (GSIF). ISRIC—World Soil Information, Wageningen, The Netherlands (accessed 14th November 2013). www.isric.org/projects/global-soil-information-facilities-gsif.
- Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 120, 75–93.
- Hengl, T., Heuvelink, G.B.M., Kempen, B., 2013. Mapping soil properties for Africa using regression kriging. In: *Global Soil Information Facilities. A Methodological Framework for Open Soil Information*. ISRIC—World Soil Information, Wageningen, The Netherlands.
- Heuvelink, G.B.M., 2014. Uncertainty quantification of GlobalSoilMap products. In: Arrouays, D., McKenzie, N.J., Hempel, J.W., Richer de Forges, A., McBratney, A.B. (Eds.), *GlobalSoilMap. Basis of the Global Spatial Soil Information System. Proceedings of the First GlobalSoilMap Conference*, Orleans, France, 7–9th October 2013. Taylor & Francis, Balkema, The Netherlands.
- Heuvelink, G.B.M., Brown, J.D., 2007. Towards a soil information system for uncertain soil data. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), *Digital Soil Mapping: An Introductory Perspective. Developments in Soil Science*. Elsevier, Amsterdam, pp. 97–106.
- Heuvelink, G.B.M., Webster, R., 2001. Modelling soil variation: past, present and future. *Geoderma* 100, 269–301.
- Holmgren, G.G., 1988. The point representation of soil. *Soil Sci. Soc. Am. J.* 52, 712–716.

- Hong, S.Y., Kim, Y.H., Han, K.H., Hyun, B.K., Zhang, Y.S., Song, K.C., Minasny, B.M., McBratney, A.B., 2012. Digital soil mapping of soil properties for Korean soils. In: Minasny, B., Malone, B., McBratney, A.B. (Eds.), *Digital Soil Assessment and Beyond*. CRC Press.
- Hong, S.Y., Minasny, B., Han, K.H., Kim, Y., Lee, K., 2013. Predicting and mapping soil available water capacity in Korea. *PeerJ* 1, e71. <http://dx.doi.org/10.7717/peerj.71>.
- Jacquier, D.W., Seaton, S., 2012. Spline Tool for Estimating Soil Attributes at Standard Depths. CSIRO Land and Water, Australia. http://www.asris.csiro.au/methods.html#Method_Downloads.
- Ji, L., Peters, A.J., 2003. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sens. Environ.* 87, 85–98.
- Kempen, B., Brus, D.J., Stoorvogel, J.J., 2011. Three-dimensional mapping of soil organic matter content using soil type-specific depth functions. *Geoderma* 162, 107–123.
- 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.
- Lagacherie, P., Robbez-Masson, J.M., Nguyen-The, N., Barthès, J.P., 2001. Mapping of reference area representativity using a mathematical soilscape distance. *Geoderma* 101, 105–118.
- Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), 2007. *Digital Soil Mapping: An Introductory Perspective*. In: *Developments in Soil Science*, Elsevier, Amsterdam, pp. 465–475.
- Lilburne, L., Hewitt, A., Ferriss, S., 2009. Progress with the design of a soil uncertainty database, and associated tools for simulating spatial realisations of soil properties. In: Caetano, M., Painho, M. (Eds.), *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, pp. 5510–5519.
- Lin, H., Bouma, J., Wilding, L.P., Richardson, J.L., Kutilek, M., Nielsen, D.R., 2005. *Advances in hydropedology*. *Adv. Agron.* 85, 1–89.
- Mallavan, B.P., Minasny, B., McBratney, A.B., 2010. Homosoil, a methodology for quantitative extrapolation of soil information across the globe. In: Boettinger, J.L., Howell, D., Moore, A.C., Hartemink, A.E., Kienast-Brown, S. (Eds.), *Digital Soil Mapping—Bridging Research, Environmental Application, and Operation*. Springer, Dordrecht, pp. 137–149.
- Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138–152.
- Malone, B.P., McBratney, A.B., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma* 160, 614–626.
- Malone, B.P., McBratney, A.B., Minasny, B., Wheeler, I., 2012. A general method for downscaling earth resource information. *Comput. Geosci.* 41, 119–125.
- McBratney, A.B., Mendonça-Santos, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52.
- McKenzie, N.J., 2014. A new deal for soils—implications of the Global Soil Partnership for GlobalSoilMap. In: Arrouays, D., McKenzie, N.J., Hempel, J.W., Richer de Forges, A., McBratney, A.B. (Eds.), *GlobalSoilMap. Basis of the Global Spatial Soil Information System*. Proceedings of the First GlobalSoilMap Conference, Orleans, France, 7–9th October 2013. Taylor & Francis, Balkema, The Netherlands.
- McKenzie, N.J., Grundy, M.J., 2008. Approaches to land resource survey. In: McKenzie, N.J., Grundy, M.J., Webster, R., Ringrose-Voase, A.J. (Eds.), *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook, vol. 2. CSIRO Publishing, Melbourne.
- McKenzie, N.J., Isbell, R.F., Jacquier, D.W., Brown, K.L., 2004. *Australian Soils and Landscapes: An Illustrated Compendium*. CSIRO Publishing, Melbourne, p. 432.

- McKenzie, N.J., Grundy, M.J., Webster, R., Ringrose-Voase, A.J. (Eds.), 2008. Guidelines for Surveying Soil and Land Resources. In: Australian Soil and Land Survey Handbook vol. 2, CSIRO Publishing, Melbourne.
- Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci.* 32, 1378–1388.
- Minasny, B., McBratney, A.B., 2010. Methodologies for global soil mapping. In: Boettinger, J.L., Howell, D., Moore, A.C., Hartemink, A.E., Kienast-Brown, S. (Eds.), *Digital Soil Mapping—Bridging Research, Environmental Application, and Operation*. Progress in Soil Science, Springer, Dordrecht.
- Minasny, B., Malone, B.P., McBratney, A.B., 2012. Digital soil assessment and beyond. In: *Proceedings of the Fifth Global Workshop on Digital Soil Mapping*, Sydney, Australia, 10–13 April 2012. CRC Press/Balkema, The Netherlands.
- Molden, D. (Ed.), 2007. *Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture*. Earthscan, International Water Management Institute, London, Colombo.
- Morris, J.C., Minasny, B., McBratney, A.B., 2012. The role of soil inference systems in digital soil assessments. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), *Proceedings of the Fifth Global Workshop on Digital Soil Mapping*, Sydney, Australia, 10–13 April 2012. CRC Press/Balkema, The Netherlands.
- Nachtergaele, F.O., Van Ranst, E., 2003. Qualitative and quantitative aspects of soil databases in tropical countries. In: Stoops, G. (Ed.), *Evolution of Tropical Soil Science: Past and Future*. Koninklijke Academie voor Overzeese Wetenschappen, Brussels, pp. 107–126.
- Nauman, T., Thompson, J.A., Odgers, N., Libohova, Z., 2012. Fuzzy disaggregation of conventional soil maps using database knowledge extraction to produce soil property maps. In: Minasny, B., Malone, B., McBratney, A.B. (Eds.), *Digital Soil Assessments and Beyond: 5th Global Workshop on Digital Soil Mapping*, Sydney, Australia.
- Nauman, T.W., Thompson, J.A., 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. *Geoderma* 213, 385–399. <http://dx.doi.org/10.1016/j.geoderma.2013.08.024>.
- Nix, H.A., 1984. Minimum data sets for Agrotechnology Transfer. In: Kimble, V. (Ed.), *Proceedings of the International Symposium for Agrotechnology Transfer*. ICRISAT, Patancheru, India, pp. 181–188.
- O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., Oakley, J., Rakow, T., 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, Ltd., Chichester.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma* 63, 197–215.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* 67, 215–226.
- Odeh, I.O.A., Leenaars, J., Hartemink, A.E., Amapu, I., 2012. The challenges of collating legacy data for digital mapping of Nigerian soils. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), *Proceedings of the Fifth Global Workshop on Digital Soil Mapping*, Sydney, Australia, 10–13 April 2012. CRC Press/Balkema, The Netherlands, pp. 453–458.
- Odgers, N.P., Libohova, Z., Thompson, J.A., 2012. Equal-area spline functions applied to a legacy soil database to create weighted-means maps of soil organic carbon at a continental scale. *Geoderma* 189, 153–163.
- Ponce-Hernandez, R., Marriott, F.H.C., Beckett, P.H.T., 1986. An improved method for reconstructing a soil-profile from analysis of a small number of samples. *J. Soil Sci.* 37, 455–467.

- Salas, F.R., Boldrini, E., Maidment, D.R., Nativi, S., Domenico, B., 2012. Crossing the digital divide: an interoperable solution for sharing time series and coverages in Earth sciences. *Nat. Hazards Earth Syst. Sci.* 12, 3013–3029.
- Sanchez, P.A., Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.L., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vågen, T.-G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., Zhang, G.-L., 2009. Digital soil map of the world. *Science* 325, 680–681.
- Soil Survey Division Staff, 1993. Soil survey manual. Soil Conservation Service, In: U.S. Department of Agriculture Handbook 18. US Government Printing Office, Washington, DC.
- Soil Survey Staff, 1975. Soil taxonomy: a basic system of soil classification for making and interpreting soil surveys. In: Agriculture Handbook No. 436. United States Government Printing Office, Washington, DC.
- Solomatine, D.P., Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. *Water Resour. Res.* 45, W00B11. <http://dx.doi.org/10.1029/2008WR006839>, Link to external resource: 10.1029/2008WR006839.
- Sonnerfeld, B.G.J.S., Dent, D.L., 2009. How good is GLASOD? *J. Environ. Manage.* 90, 274–283.
- Srinivasan, R., Zhang, X., Arnold, J., 2010. SWAT ungauged: hydrological budget and crop yield predictions in the Upper Mississippi River Basin. *Trans. ASABE* 53, 1533–1546.
- Thompson, J.A., Prescott, T., Moore, A., Bell, J., Kautz, D., Hempel, J., Waltman, S., Perry, C.H., 2010. Regional approach to soil property mapping using legacy data and spatial disaggregation techniques. In: 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia.
- UNEP, 2007. Global Environmental Outlook GEO4—Environment for Development. United Nations Environment Program, Nairobi, Kenya.
- UNEP, 2012. UNEP Year Book 2012. United Nations Environment Program, Nairobi, Kenya.
- van Baren, H., Hartemink, A.E., Tinker, P.B., 2000. 75 years of The International Society of Soil Science. *Geoderma* 96, 1–18.
- van Engelen, V.W.P., Dijkshoorn, J.A., 2012. Global and National Soils and Terrain Databases (SOTER): ISRIC Report 2012/04. ISRIC—World Soil Information, Wageningen, Procedures Manual version 2.0, draft for comments.
- van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. *Field Crop Res* 143, 34–43.
- Viscarra Rossel, R.A.V., Adamchuk, V.I., Sudduth, K.A., McKenzie, N.J., Lobsey, C., Sparks, D., 2011. Proximal soil sensing: an effective approach for soil measurements in space and time. *Adv. Agron.* 113, 237–282.
- Wallace, J.S., 2000. Increasing agricultural water use efficiency to meet future food production. *Agric. Ecosyst. Environ.* 82, 105–119.
- Webster, R., 1978. Mathematical treatment of soil information. In: Transactions of the 11th International Congress of Soil Science, 3, pp. 161–190.
- Webster, R., Lark, M., 2013. Field Sampling for Environmental Science and Management. Routledge Taylor & Francis Group, London.
- Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists, second ed. Wiley, Chichester.
- Wei, S., McBratney, A.B., Hempel, J.W., Minasny, B., Malone, B.P., D’Avello, T., Burras, L., Thompson, J.A., 2010. Digital harmonisation of adjacent soil survey areas—4 Iowa Counties. In: 19th World Congress of Soil Science, Soils Solutions for a Changing World, Brisbane, Australia.