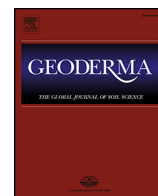




Contents lists available at ScienceDirect

Geoderma

journal homepage: [www.elsevier.com/locate/geoderma](http://www.elsevier.com/locate/geoderma)

## Digital soil mapping: A brief history and some lessons

Budiman Minasny<sup>\*</sup>, Alex.B. McBratney

Faculty of Agriculture & Environment, The University of Sydney, NSW 2006, Australia

### ARTICLE INFO

#### Article history:

Received 6 March 2015

Received in revised form 14 July 2015

Accepted 26 July 2015

Available online xxx

#### Keywords:

Soil survey

Pedology

GlobalSoilMap

Soil history

Case-based reasoning

### ABSTRACT

Digital soil mapping (DSM) is a successful sub discipline of soil science with an active research output. The success of digital soil mapping is a confluence of several factors in the beginning of 2000 including the increased availability of spatial data (digital elevation model, satellite imagery), the availability of computing power for processing data, the development of data-mining tools and GIS, and numerous applications beyond geostatistics. In addition, there was an increased global demand for spatial data including uncertainty assessments, and a rejuvenation of many soil survey and university centres which helped in the spreading of digital soil mapping technologies and knowledge. The theoretical framework for digital soil mapping was formalised in a 2003 paper in *Geoderma*. In this paper, we define what constitutes digital soil mapping, sketch a brief history of it, and discuss some lessons. Digital soil mapping requires three components: the input in the form of field and laboratory observational methods, the process used in terms of spatial and non-spatial soil inference systems, and the output in the form of spatial soil information systems, which includes outputs in the form of rasters of prediction along with the uncertainty of prediction. We also illustrate the history with a number of sleeping beauty papers that seem too precocious and consequently the ideas were not taken up by contemporaries and largely forgotten. It took another 30 to 40 years before the ideas were rediscovered and then flourished. Examples include proximal soil sensing that was developed in the 1920s, soil spectroscopy in 1970s, and soil mapping based on similarity of environmental factors in 1979. In summary, the coming together of emerging topics and timeliness greatly assists in the development of paradigm. We learned that research and ideas that are too precocious are largely ignored — such work warrants (re)discovery.

© 2015 Published by Elsevier B.V.

### 1. Introduction

Digital soil mapping (DSM) has become a successful sub-discipline of soil science. Currently, the number of papers on DSM increases at a rate of 12 papers per year, and the number of citations increases by 384 citations per year (Fig. 1). The use of computer or numerical models to map soil is not new and researches into methods for creating digital soil maps have been produced since the 1990s (e.g. Skidmore et al., 1991; Bell et al., 1992; Odeh et al., 1992a; McKenzie and Austin, 1993; Moore et al., 1993). McBratney et al. (2003) noted their commonalities and proposed a generic framework called the *scorpan*-SSPFe (soil spatial prediction function with spatially autocorrelated errors) as a method to produce digital soil maps. The term digital soil maps has been used since early on, for example Roger Tomlinson, the father of GIS (Tomlinson, 1978) labelled digitised polygon maps as digital soil maps. Similarly, Bliss and Reibold (1989) and Bliss et al. (1995) converted the STATSGO polygon maps into “digital soil maps”. Dobos et al. (2002) used the term “digital soil mapping” as a way of integrating soil maps with DEM and satellite sensing images.

The aim of this paper is to define what constitutes digital soil mapping following the framework of *scorpan*-SSPFe and reviews several research topics that contribute to the development of digital soil mapping. We will sketch a brief history using examples from several pioneering soil mapping studies, highlighting some ‘sleeping beauties’ papers and their rediscovery, and then discuss some lessons for the future.

Since the digital soil mapping *scorpan* concept was introduced, and following a series of global workshops, there has been huge interest in this topic of research. The first global digital soil mapping workshop was held in Montpellier in September 2004. The IUSS working group on Digital Soil Mapping was formed following the first workshop. Successive global workshops were held in Rio De Janeiro, Brazil in 2006, Logan, USA in 2008, Rome, Italy in 2010, Sydney, Australia in 2012 and in Nanjing, China in 2014. It has resulted in a series of books (Lagacherie et al., 2006; Hartemink et al., 2008; Boettinger, 2010; Minasny et al., 2012). Following the second global workshop in 2006, the GlobalSoilMap project was initiated (Arrouays et al., 2014; Hempel et al., 2014).

This paper does not attempt to give a history of soil mapping, which has been covered in many reviews (Yaalon, 1989; Brown, 2006; Legros, 2006; Hartemink et al., 2013; Miller and Schaetzl, 2014). It also will not discuss the history of pedometrics research (Webster, 1994). Bui (2006) provided a review of digital soil mapping development in Australia until

<sup>\*</sup> Corresponding author.

E-mail address: [Budiman.minasny@sydney.edu.au](mailto:Budiman.minasny@sydney.edu.au) (B. Minasny).

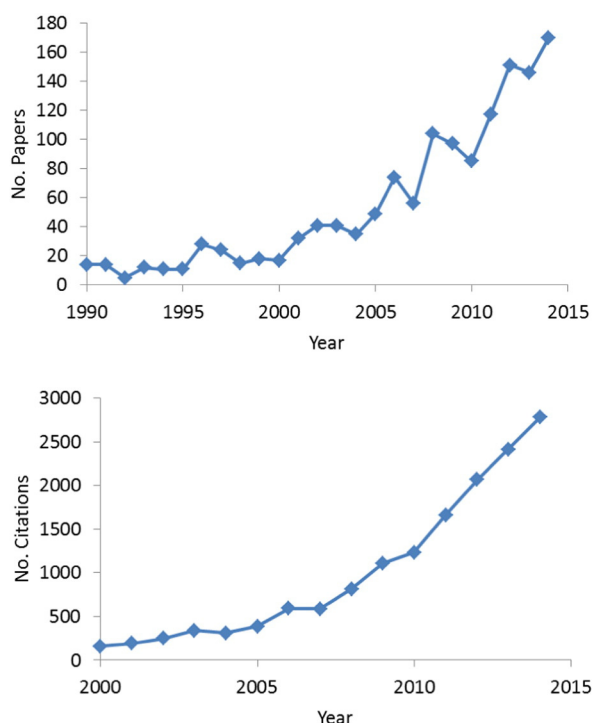


Fig. 1. Number of papers and number of citations from keyword search “digital” “soil” “mapping” from the Scopus database (data extracted in February 2015).

2004. Grunwald (2009) characterised some recent works on digital soil mapping and modelling, and Grunwald et al. (2011) reviewed some works on digital soil mapping at the continental scale.

The structure of this paper is as follows: we will first define what constitute digital soil mapping (Section 2). We will then review several key concepts that lead to the development of digital soil mapping (Fig. 2). We will illustrate these with some sleeping beauty papers. Sleeping beauty in scientific publication refers to a publication whose importance goes unnoticed for a long time and then suddenly attracts a lot of attention or citations (Van Raan, 2004; Ke et al., 2015). Section 3 discusses some sleeping beauties in soil proximal and remote

sensing. Section 4 examines key concepts on factorial models, environmental similarity, digital cartography, pedometrics and environmental correlation. Finally we discuss the confluence of various factors that led to the development of digital soil mapping and discuss where it is going.

## 2. What is digital soil mapping?

A soil map is a graphic representation for transmitting information about the spatial distribution of soil attributes (Yaalon, 1989). The earliest soil maps were produced in the mid 18th century for agricultural purposes. These soil maps were made simply to delineate homogeneous areas with intrinsic soil attributes useful in determining suitable land use, and not for soil classification. A more formalised soil mapping approach was later derived from agrogeologists, where maps of geology and soil types were derived from topographic maps (Miller and Schaetzl, 2014). In the 19th century the Russian school stressed the importance of genetic soil type, while in the USA the stress is on the soil's intrinsic properties.

Dokuchaev in 1883 produced a map of humus content in Russia based on quantitative point observations, plotted as an isohumus map (Hartemink et al., 2013; Miller and Schaetzl, 2014). There was also a discussion by G.F. Nefedov in 1908 in Russia (as cited by Hartemink et al., 2013) that soil properties have to be mapped first and then soil classes can be delineated based on areas that occupied uniform or similar properties. However this idea seems never realised as there was not enough data at the time, and thus soil classes continued to be mapped.

In conventional soil survey, soil is mapped based on a soil surveyor's conceptual or mental model (Hudson, 1992). Aerial photographs, Landsat images, and digital elevation models (DEMs) are used to identify environmental features relating to geology, landform or vegetation. This process is then verified with field observations (Legros, 2006). The final product is a map with a legend of soil types, which can be difficult to interpret and use. In addition, the subjective nature of the map process yields many drawbacks (Ryan et al., 2000).

The conceptual model of the traditional mapping approach can be harvested and modernised by making use of contemporary quantitative techniques. Digital soil mapping is defined as: *the creation and population of spatial soil information systems by the use of field and laboratory observational methods coupled with spatial and non-spatial soil inference systems* (Lagacherie and McBratney, 2006). Other terminology has also been used or proposed, including: computer-assisted soil cartography, numerical soil cartography, pedometric mapping, environmental correlation, predictive soil mapping, or geographical extrapolation using models (Franklin, 1995; McKenzie and Ryan, 1999; Scull et al., 2003; Legros, 2006; Kempen et al., 2010). However, digital soil mapping is not only about making soil maps using computer or numerical approaches or quantitative relationships between environmental variables and soil properties, according to the definition of Lagacherie and McBratney (2006) it needs to have three main components:

- The *input* in the form of field and laboratory observational methods, this includes the use of legacy soil observations or soil maps, and collecting new samples using statistical sampling techniques.
- The *process* used in terms of spatial and non-spatial soil inference systems, this includes building mathematical or statistical models relating soil observations with their environmental covariates or *scorpan* factors.
- The *output* in the form of spatial soil information systems, which includes outputs in the form of rasters of prediction along with the uncertainty of prediction. This output can be readily updated as new information becomes available.

In the early literature, digitised conventional soil map in the form of polygons is called a digital soil map (Tomlinson, 1978). While the

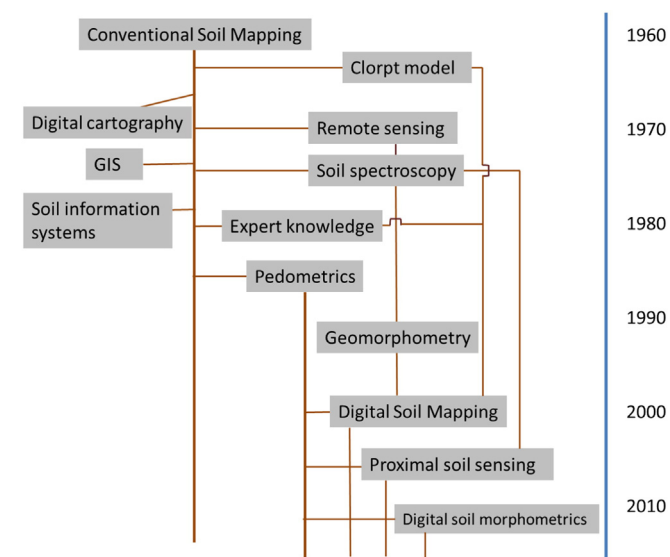


Fig. 2. A diagram illustrating the evolution and confluence of different concepts in digital soil mapping used in this paper.

product is digital, the process of making the map is not based on any statistical inference, and thus it is not digital soil map, but a digitised soil map.

The most common work involves predicting soil class or soil attribute from its so-called *scorpan* factors using a spatial soil prediction function with autocorrelated error (SSPFe) (McBratney et al., 2003):

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

where  $S_c$  and  $S_a$  are soil classes and soil attributes as a function of soil, climate, organisms, relief, parent material, age and spatial position, 'e' is spatially correlated residuals. The *scorpan* model is different from Jenny's factorial model or the Sergey-Zakharaov equation (Florinsky, 2012). The *scorpan* model is used for quantitative prediction of soil classes or continuous soil attributes based on empirical observations and not trying to explain the factors of soil formation. In addition, soil can be used as a factor because soil can be predicted from its properties, or soil properties from its class or other properties.

Most work in digital soil mapping is based on building a numerical model relating field soil observations and *scorpan* factors (e.g. Bou Kheir et al., 2010; Brungard et al., 2015). The model is then applied to spatially exhaustive environmental data. However digital soil mapping can also use conventional soil maps as input for the purpose of extrapolating, updating or disaggregating soil maps (Subburayalu and Slater, 2013; Odgers et al., 2014). The rationale is to discover the knowledge embedded in completed surveys, or reverse engineer the surveyor's soil-landscape mental model through machine learning techniques (Moran and Bui, 2002). Rules can be learned from the soil map, and subsequently these learned rules can be used to spatially extrapolate the mapping areas (Grinand et al., 2008), update or refine existing maps (Collard et al., 2014), or disaggregating the soil mapping units into soil series (Häring et al., 2012; Subburayalu and Slater, 2013; Odgers et al., 2014). In addition the disaggregated maps and *scorpan*-kriging derived maps can be combined to produce better estimates (Malone et al., 2014).

In the following sections, we will discuss several topics that have contributed to the development of digital soil mapping (Fig. 2). We will illustrate this with some sleeping beauties and their rediscovery. We will discuss that although the ideas offered by the sleeping beauties were rediscovered, the papers remain asleep, relatively unknown and uncited.

### 3. Proximal soil sensing, remote sensing, and soil spectroscopy

#### 3.1. Proximal soil sensing

In 1925, long before digital computers and electronic sensors, Bernard Keen and William Haines from Rothamsted Experimental Station in the UK conceived and built the first on-the-go soil strength sensor. They made measurements and a high-resolution soil map, and confronted the reality and importance of soil spatial variation (Keen and Haines, 1925; McBratney and Minasny, 2010). The first on-the-go soil strength sensor measurements were made on Sawyers field (2.4 ha) that had received uniform treatment for many years and was believed to be uniform. (Haines and Keen, 1925a) made a contour map based on average soil mechanical resistance values obtained for each plot. They called these "isodynes" for contour lines of equal force or soil mechanical resistance. (Haines and Keen, 1925b) used their dynamometer to map soil mechanical resistance in the now classical Rothamsted experimental plots. They built a physical 3-D model (pre-GIS and 3D printer) for the soil mechanical resistance and showed that the soil variation is natural and not caused by the manure that was applied from 1843 to 1925 (Fig. 3).

This work which was ahead of its time, not taken up by contemporaries and is now largely forgotten. Nothing was reported in the literature on this kind of work for a long time, and the Haines and Keen

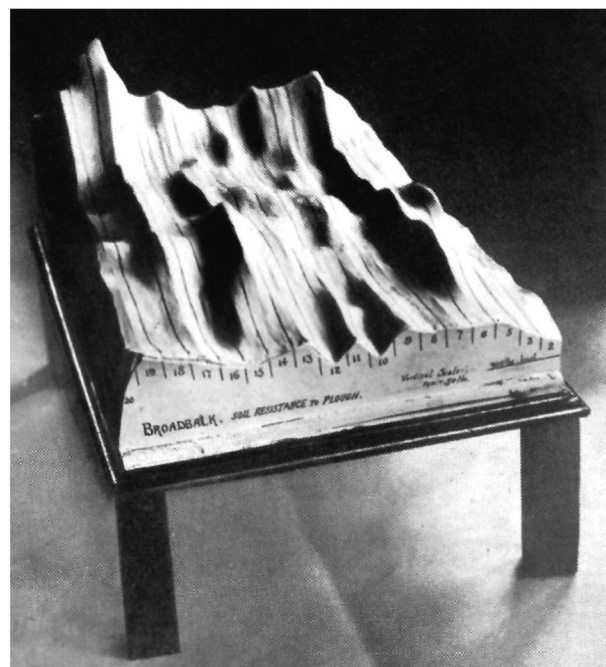


Fig. 3. Physical 3D model of variation of soil mechanical resistance across the Broadbalk continuous wheat experiment (Chromic Luvisols) in 1924. (From Haines and Keen (1925b), used with permission).

(1925a) paper has been cited only 5 times (Google Scholar, February 2015). Published work on the 'on-the-go' soil strength measurement seemed to begin with Stafford and Hendrick (1985) and with more recent developments by Sun et al. (2006) and with the advent of site-specific soil and crop management. Although this study is not digital soil mapping, it is the first proximal soil sensing instrument and study. Proximal soil sensing was formalised in early 2000s and motivated by the need for high-resolution spatial and temporal soil information. Proximal soil sensing, *sensu stricto*, is defined as in-situ, mobile (on-the-go), field methods for sensing soil information. In digital soil mapping the methodology is a function of the resolution sought, and thus proximal soil sensing is the methodology for high-resolution digital soil mapping (McBratney et al., 2011). Related to proximal soil sensing is the emergent study of digital soil morphometrics which use proximal sensors for measuring and mapping soil profile properties, and deriving continuous depth functions (Hartemink and Minasny, 2014). Digital soil morphometrics is distinct from proximal soil sensing and aims to yield new insights in soil horizonation and enhance pedological understanding. An IUSS working group was recently established to proffer these new insights.

#### 3.2. Remote sensing

The use of remotely-sensed aerial photographs for soil mapping started as early as the 1920s, and subsequent use of air-photo interpretation in the USA (Bushnell, 1929) and Australia (Prescott and Taylor, 1930). However soil maps produced from topographic information and aerial photographic interpretation does not constitute digital soil mapping, as the pattern recognition and identification was done manually and subjectively. There is no statistical or mathematical inference behind the process.

Statistical inference for mapping started with the introduction of remotely-sensed infrared spectroscopy. The Laboratory for Applications of Remote Sensing (LARS) at Purdue University (Kristof et al., 1973) proposed 'Spectral Mapping of Soil Organic Matter' using hyperspectral soil reflectance from the visible and near infrared signal of an airborne scanner. A numerical classification procedure was used to group data

into classes that represent the distribution of soil organic matter. This type of work has continued, as exemplified by [Frazier and Cheng \(1989\)](#) who devised a linear relationship between with organic carbon content and Landsat TM multispectral Band ratio 5/4 (SWIR/NIR) for mapping.

LARS at Purdue University produced innovative research in the field of remote sensing, generating the first hyperspectral soil spectral library, predicting soil carbon content from NIR, and mapping soil carbon digitally. It proposed mapping soil type based on unsupervised classification of Landsat images ([Cipra, 1973](#); [Kornblau and Cipra, 1983](#)). At the same time, researchers in India also used Landsat images in conjunction with topographical and geological information to map soils ([Singh, 1980](#); [Karale et al., 1983](#)). A soil spectral database or library was developed in the 1980s at LARS ([Stoner and Biehl, 1980](#)), where the database contained near infrared reflectance spectra along with the physical, chemical, engineering, and site information for 500 soil samples from 39 states in the USA, as well as Brazil, Costa Rica, Sudan, Spain, and Jordan.

Remote sensing has remained an essential data source for soil mapping ([Agbu et al., 1990](#); [Mulder et al., 2011](#)). The spectral range from remote sensing used for soil mapping is not limited to the visible and infrared; gamma ray ([Cook et al., 1996a](#)) and radio frequencies ([Mendonça-Santos and dos Santos, 2006](#)) can and have also be used.

### 3.3. Soil spectroscopy

The spectroscopy work in the 1970–1980s using remotely or proximally-sensed visual to NIR spectrometers generating multispectral and hyperspectral data was ahead of its time ([Al-Abbas et al., 1972](#); [Kristof et al., 1973](#); [Mathews et al., 1973](#); [Stoner and Baumgardner, 1981](#)). It took another 30 years before the soil spectroscopy work became mainstream ([Ben-Dor et al., 1997](#)), and soil spectral libraries were developed again ([Shepherd and Walsh, 2002](#); [Brown et al., 2006](#)). The interest in soil NIR followed the developments of NIR spectroscopy in agriculture in the late 80s, where desktop NIR spectrometers were used to rapidly measure the chemical composition of cereals, fruit and vegetables, meat, etc. ([Bellon-Maurel and McBratney, 2011](#)). The use of desktop or portable NIR instruments in soil science (as opposed to remotely sensed NIR) started in the 1990s with the development of multivariate data analysis of ultraspectral data. Proximally-sensed soil spectroscopy is still an active area of enquiry ([Bellon-Maurel and McBratney, 2011](#); [Sarathjith et al., 2014](#)). A detailed review on NIR spectroscopy in soil science can be found in [Stenberg et al. \(2010\)](#).

The paper by [Stoner and Baumgardner \(1981\)](#) “Characteristic variations in reflectance of surface soils” is an example of the Sleeping Beauties in soil science, in 1981–1990 it only has an average of 1 citation per year, and from 1991 to 2000 the average is 6 citations per year. The paper then started to gain citation from 2001 to 2010 with an average of 11 citations per year, and in 2013, it was cited 18 times (citation data from Web of Knowledge).

## 4. Factorial models, environmental similarity, and environmental correlation

### 4.1. Factorial models of soil formation

Jenny's factorial model of soil formation is used conceptually in conventional mapping. Jenny's clorpt model is also used quantitatively, where “the clorpt formula was taken apart and explicit factor sequences were secured”, i.e. one factor is varied, while the other factors are held constant in the form of chronofunctions, topofunctions as some such examples. The topofunction is most popular, such as the work by ([Bushnell, 1943](#)) on soil distribution based on the catena concept. In fact the American (USA) soil-landscape model is largely:  $s = f(r)$ . [Troeh \(1964\)](#) demonstrated that soil drainage classes could be differentiated by their slope and curvature. [Walker et al. \(1968\)](#) found that soil

morphological properties collected along a transect, such as A horizon depth, depth to mottling and carbonates were related to slope, curvature, aspect and distance from the local summit in a multiple linear regression. They also realised that considerable soil variation remained unaccounted for by the regression. Most of these functions were used for exploratory data analysis and not for mapping. Perhaps there were no resources to predict soil across the landscapes using the empirical equations, or the idea was too radical for the time and not attempted.

A multivariate factors model was proposed by ([Jenny et al., 1968](#)) as an “integrated clorpt”, where all factors were simultaneously modelled in the form of a multivariate linear regression:

$$S = a + k_1 \text{ MAP} + k_2 \text{ MAT} + k_3 \text{ Parent Material} + k_4 \text{ slope} + k_5 \text{ Vegetation} + k_6 \text{ Latitude},$$

where S indicates soil properties (texture, C, N, cations, and 1st principal component of clay minerals), and k indicates empirical constants, and MAP and MAT correspond to mean annual precipitation and temperature respectively. This model was calibrated using 97 topsoil samples (0–20 cm) collected along a moisture transect from deserts at 35°N latitude to a high rainfall area at 41°N. In the process of modelling, Jenny understood there was correlation among the independent factors, which he addressed using principal component analysis (PCA). It is one of the earliest examples using PCA for soil data analysis. It is also one of the first attempts to link soil carbon to some of the factors of soil formation. Jenny found that precipitation was the strongest exploratory variable ( $R^2 = 0.45$ ), followed by temperature ( $R^2 = 0.09$ ), and parent material ( $R^2 = 0.06$ ). In this regional dataset, the influence of slope was minimal ( $R^2 = 0.01$ ). The model was used to explain the factors that control the distribution of soil carbon along with other properties and not as a prediction function or for mapping. Jenny's 1968 paper is not well known and has only 21 citations in Google Scholar (Feb. 2015). This work has been rediscovered and extended to digital soil mapping. For example [Gray et al. \(2015\)](#) express this type of approach as pragmatic soil spatial modelling in their recent work in NSW, Australia. Various forms of the factorial models have been proposed. For example, ([Runge, 1973](#)) considered the energy components of soil formation, but that equation is mostly used as qualitative interpretation. Quantitative energy model for soil formation had been proposed by V.R. Volobuyev and detailed in his book “Ecology of Soils” ([Volobuyev, 1964](#)). [Volobuyev \(1964\)](#) defined the energy balance for soil formation as the sum of energy in weathering of rock or parent materials, soil organic matter accumulation and transformation, evapotranspiration, and leaching. Most of Volobuyev's works are published in Russian with some English translations published in the Soviet Soil Science Journal. His work on the energy of soil formation is generally unknown in the soil science literature and has rarely been cited ([Field and Minasny, 2008](#)). It took 40 years before this energy model was rediscovered by [Rasmussen et al. \(2005\)](#) and [Rasmussen and Tabor \(2007\)](#) as a quantitative pedologic energy model for predicting soil properties across the USA.

### 4.2. Soil mapping based on expert knowledge and environmental similarity

An evolution of Jenny's factorial model is the soil-landscape relationships following the work of [Ruhe \(1956\)](#) and [Butler \(1959\)](#) that recognised the periodic nature of soil-landscape development. This soil-landscape relationship approach influenced many of the conventional soil mapping approaches. The work by [Legros and Bonneric \(1979\)](#) is considered as the first French attempt in computer-assisted soil mapping. This work has most of the elements of digital soil mapping (input data, covariates and statistical inference). The procedure for generating a map attempted to mimic the reasoning of an expert pedologist in a soil-landscape relationship: if a particular soil is found in a particular environment, then the same soil will be found in a similar environment. They used observations of 38 soil profiles from a survey of

the Pilat Regional Natural Park, and deduced rules that enabled geographical extrapolation of the observations to the whole park of 624 km<sup>2</sup>. The area was gridded into 2496 pixels, each with a size of 500 × 500 m. The Gower statistical distance was calculated for the similarity of the environments of each pixel to the environments of each of the 38 profiles. The environments were then grouped by the four soil types found in the area. A similarity measure was calculated between each pixel and the four soil types. The soil type with the highest similarity index was allocated to the pixel (Fig. 4). Lagacherie (1992) and Lagacherie et al. (1995) proposed another method by extracting the knowledge from a reference area where it has been mapped by the conventional technique. The rules extracted from this area were then extrapolated to a larger area within the region.

The digital mapping of environmental similarity by Legros and Bonneric (1979) was largely unknown as it was published in an obscure French journal, but it was rediscovered and called case-based reasoning

by Shi et al. (2004), where they used a fuzzy membership function to calculate the similarity of the environmental covariates of a location and of the modal soil profile. The judgement on the use of covariates and how to weigh each of the covariates for the calculation of the similarity measure can be subjective however.

Another example of computer assisted mapping using expert knowledge is by Skidmore et al. (1991), Cook et al. (1996b), and Skidmore et al. (1996) who tried to quantify the knowledge from an experienced soil scientist by building rules or relationships between the soil-landscape units and their topographical attributes (slope gradient, soil wetness index, and topographical position). A Bayesian inference algorithm was used to calculate the likelihood of a hypothesis (soil-type) occurring, given a piece of evidence (soil-landscape condition). A forward chaining process is then applied, in other words all the rules were evaluated at each pixel to give the probability of occurrence of a particular soil class (Skidmore et al., 1996). This type of

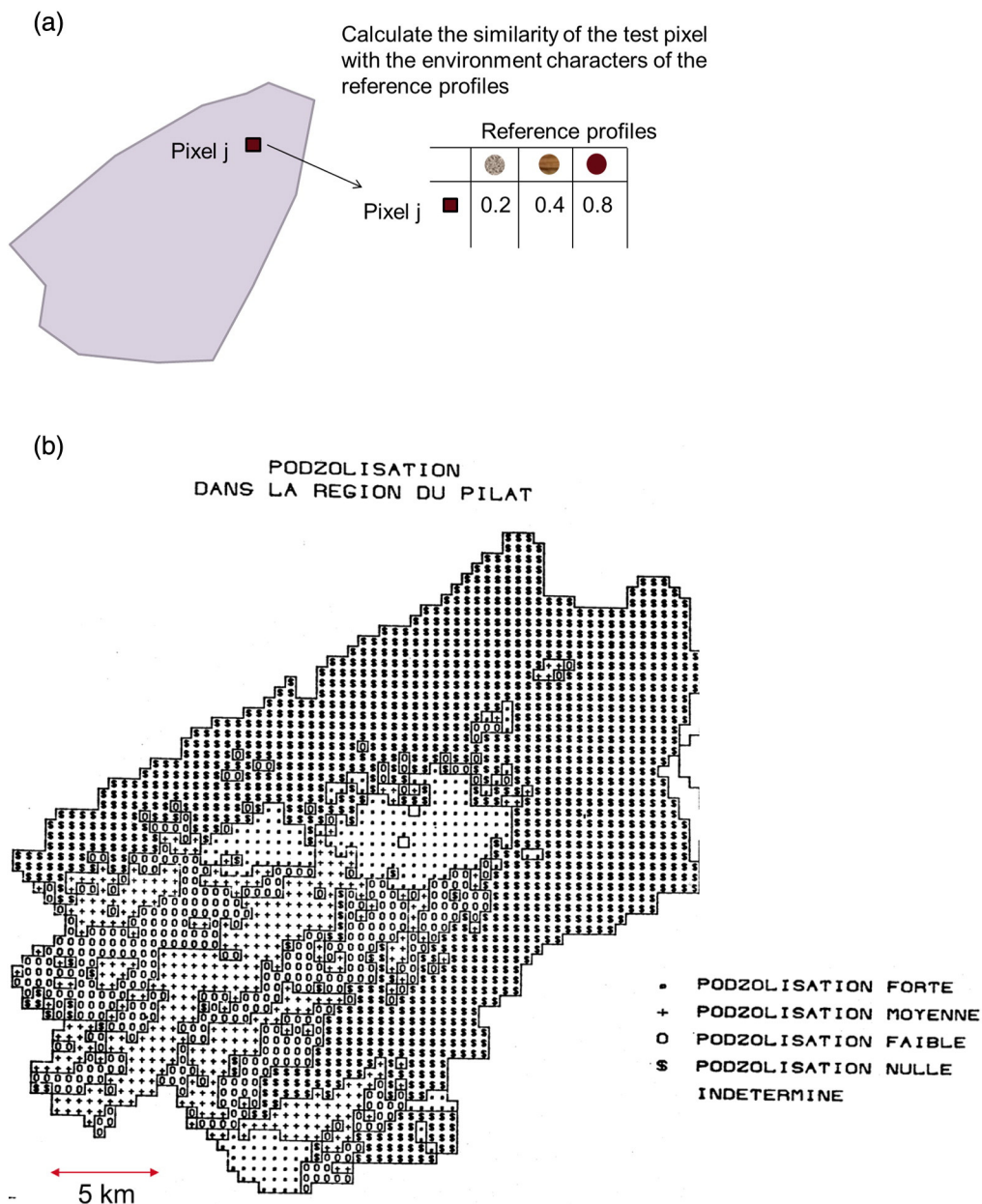


**Fig. 4.** (a) Illustration of the similarity method, and (b) the very first French digital soil map showing degree of podzolization in the region of Pilat, the map is represented as a raster with a pixel size of 500 × 500 m. (From Legros and Bonneric, 1979).

expert-system was the basis of the SoLIM (Soil Land Inference Model) model by (Zhu et al., 1997) who used a fuzzy inference algorithm.

#### 4.3. Digital soil cartography and soil information systems

With the advent of computers in the 1970s, Webster and Burrough (1972) attempted to map an area of north Berkshire in the UK. They selected a 1400 m × 600 m area that was sampled on a 100-m square grid, and seventeen properties of the soil profile were measured. The data (n = 84) were analysed by principal component analysis and the first component accounted for 40% of the total variance and well represented the field characters used for soil classification. An isarithmic map of the first component showed how the soil changes over the landscape and was in agreement with a soil-series map made by traditional soil survey.

Webster et al. (1979) demonstrated that soil parameters could be stored, and extracted with a rudimentary database system linked to software which printed out the map using a line-printer. In the report "Soils of the Ivybridge area in the West of England", an area of 60 km<sup>2</sup> was inspected from auger borings on a 100-m square grid. The area was partitioned into 60 columns and 101 rows. At each 'pixel', characteristics of the soil profile and site were recorded and the data were put into a database. Soil maps were made by a line printer, where each printed symbol represented one observation. A map of soil attributes such as the soil series, drainage class, soil colour value, texture class, depth to mottling, were made without the use of numerical interpolation techniques. This type of mapping based on grid survey has not persisted as it is too expensive and laborious to collect soil observations in a regular fine grid. There is no statistical inference involved, hence it is called digital soil cartography.

The International Soil Science Society (ISSS) Working Group on Soil Information Systems was formed in Wageningen, the Netherlands in 1975 with the aim to develop an information system for soil-related applications. Most of the work in this working group related to soil databases, data management, storing, retrieving and manipulating data and digitising conventional soil maps (Sadovski and Bie, 1977; Burrough and Bie, 1984). Legros and Hensel (1978) reported on computer-assisted soil cartography in France. Computer system approaches were proposed on digitised soil maps (Tomlinson, 1978). The storage and analysis of soil data in a grid format has long been considered desirable (Ragg, 1977). One way of making soil data in a regular grid is by interpolation and an example is by De Gruijter and Bie (1975) who used an early GIS package called SYMAP.

Efforts also been made to convert polygon maps into digital soil properties maps (Bliss and Reybold, 1989). However applying these "digital maps" that were derived from regional scale mapping (STATSGO) for landscape scale modelling was found to be problematic as the local variability was not captured (Lathrop et al., 1995).

#### 4.4. Geostatistics and pedometrics

Research in the 1980s experimented with the use of statistical models for analysing soil data or what was later termed pedometrics (Webster, 1994). In Japan, Kosaki et al. (1982) performed a numerical classification of soil observations, and devised a computer programme to map the point data of soil types into polygons. In the USA, Shovic and Montagne (1985) attempted a statistical model of the landscape and soils for use in mapping.

Researchers also started to use geostatistics that was developed in the mining industry in the 1960s for spatial optimal linear prediction. It has become an attractive tool for soil scientists as the deterministic model (such as Jenny's topofunctions) can only explain a portion of the variation in spatial data and the spatial random component is large and auto-correlated. Soil has been useful for the development of geostatistics as its properties generally vary continuously over space, and its spatial variation can be quantified using variograms.

Some of the first published applications of geostatistics in soil studies included salinity studies in southwest Iran (across several, 100 ha fields) (Hajrasulih et al., 1980) and three farms in the UK (Burgess and Webster, 1980). This was followed by Vieira et al. (1981) who quantified the spatial variation of field-measured infiltration rate in a field in California. Various forms of kriging have been developed to deal with continuous soil properties and classes, to give estimates for blocks or pixels of varying size, and to estimate uncertainty. This led Giltrap (1983) to propose converting randomly located soil survey observations to estimates of numerical properties located on a regular grid based on geostatistics. This is much what is currently been undertaken under the broad umbrella of digital soil mapping. However this idea was not taken up by soil surveyors until the digital soil mapping concept had been accepted. Thus most geostatistics were applied for soil chemical properties for agronomic purposes (Vieira et al., 1983).

In the development of soil geostatistics it was found that soil could be better predicted if denser sets of secondary variables (spatially cross-) correlated with the primary variable were available – this technique is called co-kriging. In the early co-kriging studies, such ancillary variables were other soil variables, indicating that other soil variables are useful predictors of soil (Vauclin et al., 1983). With improved technology, co-kriging was performed with detailed secondary data sets of environmental variables derived from digital elevation models and satellite images.

Pedometrics also explored the use of fuzzy set theory for numerical soil classification and mapping (Odeh et al., 1992a, 1992b; McBratney and Odeh, 1997). Fuzzy theory can be used to develop rules for delineating the landscape based on expert knowledge or based on numerical clustering of terrain attributes. Zhu et al. (1996) used fuzzy membership to extract knowledge from an expert pedologist on the occurrence of certain soil series from its topographic variables. These fuzzy rules were then used to map out the occurrence of soil series in an area. Irvin et al. (1997) used fuzzy k-means algorithm to group 6 terrain attributes in a landscape and demonstrated that the resulting classes were meaningful and also spatially coherent.

#### 4.5. Soil spatial models and environmental correlation

Soil spatial models where soil properties or soil classes are mapped based on mathematical functions started to appear with the advent of digital elevation models (DEMs) and terrain analysis in the beginning of the 1990s. Geographic positioning systems (GPS) became available, and allowed accurate spatial referencing of field samples and registration with environmental covariates in the GIS. Linear models were developed for spatial prediction of soil properties in a field, notably by the studies of Bell et al. (1992); Moore et al. (1993); and Slater (1994) in the USA and Odeh et al. (1992a, 1994); McKenzie and Austin (1993); and Gessler et al. (1995) in Australia. These studies demonstrated a pragmatic statistical method, where soil properties such as pH and carbon were related to topographic indices by linear regression models. Similarly Arrouays and Pelissier (1994) made a spatial prediction model of surface soil carbon pools in forest soils in France, using climatic, geomorphological, and pedological data in the form of a multiple linear regression model. These studies resumed from Jenny et al. (1968)'s seminal work, the difference is that the model is now used for mapping.

Digital terrain models and other raster layers provided complete enumeration of an area into a GIS. Soil observation points were intersected with the layers of environmental variables or covariates and a model was fitted by various means. Then the model was used to predict all other locations on the raster. For quantitative prediction purposes this has been called the 'clorpt' equation (McBratney et al., 2000). McKenzie and Austin (1993) and McKenzie and Ryan (1999) termed this approach "environmental correlation" which involves the development of explicit statistical or rule-based models that supersede the implicit mental models used by soil surveyors (Ryan et al., 2000). The statistical model is then applied to the environmental variables to

predict site observations across the complete region. Some would argue that environmental correlation is simply the natural, quantitative, evolution of clorpt. However, unlike the environmental similarity or expert knowledge approach which attempted to capture pedologist's mental model (Section 4.2), this soil spatial model is purely based on a statistical approach with known prediction uncertainty. As noted in Moore et al. (1993), it was also motivated by work in pedotransfer functions, where one soil property is predicted from other surrogate soil properties,  $S = f(s)$ .

This period also coincides with the development and availability of data-mining techniques that deal with nonlinearity of soil–landscape relationships (Pachepsky et al., 1996; McKenzie and Ryan, 1999). The success of many field-scale studies encouraged the development of mapping on larger extents, such as regions and across continents (Bui and Moran, 2003). Henderson et al. (2005) were among the first to employ a data-mining approach (selecting covariates to produce regression rules) for spatial prediction of soil properties over the entire area of Australia.

McBratney et al. (2000) suggested that the clorpt method forms a deterministic model, and that it should be combined with geostatistics to explain the random component of the spatial variation of the residuals of the model. They called it a hybrid model that offers powerful spatial prediction methods, especially for catchment and regional studies (Fig. 5).

## 5. The development of digital soil mapping

### 5.1. The rise in computer power

The ability to model and process data spatially requires significant computing power. Moore's law is an observation that the number of transistors in a dense integrated circuit doubles approximately every two years (Moore, 1998), which implies that the processing power of microchips doubles every eighteen months. Although this empirical law has attracted various criticisms, the prediction has proven to be accurate enough and well established. Because digital soil mapping is dependent on computer power and information technology, Lagacherie and McBratney (2006) speculated a relationship between the size of digital soil mapping projects that might be tackled in the future. Taking data on the number of pixels described or predicted from earlier projects they found an exponential growth with time:

$$\log_{10}(\text{no. of pixels}) = -272.7 + 0.139 * \text{Year AD.}$$

The fitted line describes a doubling in the number of pixels every 26 months, which is slightly slower than Moore's law. Recently large extent and high resolution digital soil maps have been produced globally, so we revisited the data and refitted the model and found:

$$\log_{10}(\text{no. of pixels}) = -294.8 + 0.151 * \text{Year AD.}$$

The rate of increase in the number of pixels is now approximately  $\sim 2^{0.5} \Delta \text{Year}$ , or doubling every 2 years (Fig. 6). The increase in the number of pixels is largely a response to the production of GlobalSoilMap products from countries with very large land areas such as the USA and Australia. The production of large extent digital maps has been accelerated with the availability of high performance computing, super computers, parallel processing, multicore processors such as Blue Waters and cloud computing such as the Google Earth Engine. Chaney et al. (2014) used an enhanced map disaggregation algorithm DSMART (Odgers et al., 2014) applied on the conterminous USA to produce POLARIS (Probabilistic Remapping of SSURGO) dataset with a 1 arc second spatial resolution (approximately 30 m). Soil map units were disaggregated using random forest models, producing the 50 most probable soil classes and their associated probabilities at each 1 arc second grid cell over the whole conterminous USA. The resulting data contains 9 billion grid cells (corresponding to 2 terabytes of data), and took only 5 h of computing time using the Blue Waters high performance computer. In another study, Padarian et al. (2015) used the Google Earth Engine facility to automate all tasks in digital soil mapping in one single operation for mapping global carbon distribution at a resolution of 500 m.

### 5.2. Digital soil mapping

The developments discussed in previous sections led to the formulation of the Digital Soil Mapping *scorpan* framework in 2003 by McBratney et al. (2003) (Fig. 2 & Table 1). The advance of digital soil mapping is a confluence of several factors: the availability of digital spatial data (DEM, Landsat images), the availability of computing power to process large data, the development of data mining tools and progress in GIS, applications beyond geostatistics, and institutional rejuvenation and a new generation of soil scientists that were attracted by the spatial analysis of soils. The production of digital soil maps coincided with renewed interest in agriculture (food, feed, fuel) by global fora that brought soils back onto the global research agenda (Hartemink and McBratney, 2008). The need for up-to-date and fine resolution soil information and the revival of soil research has been highlighted and

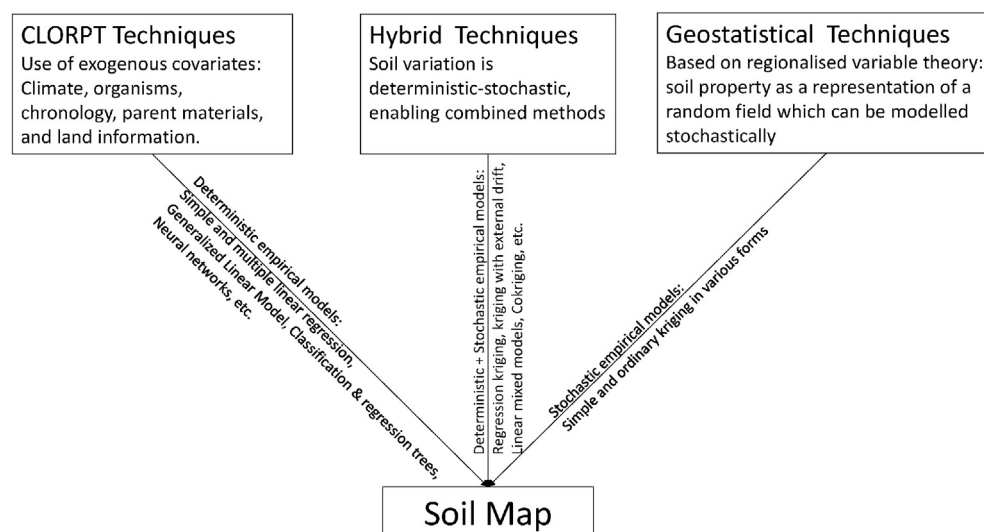


Fig. 5. The generic pedometric techniques, their hybrid and styles with respect to mode of spatial prediction, after McBratney et al. (2000).

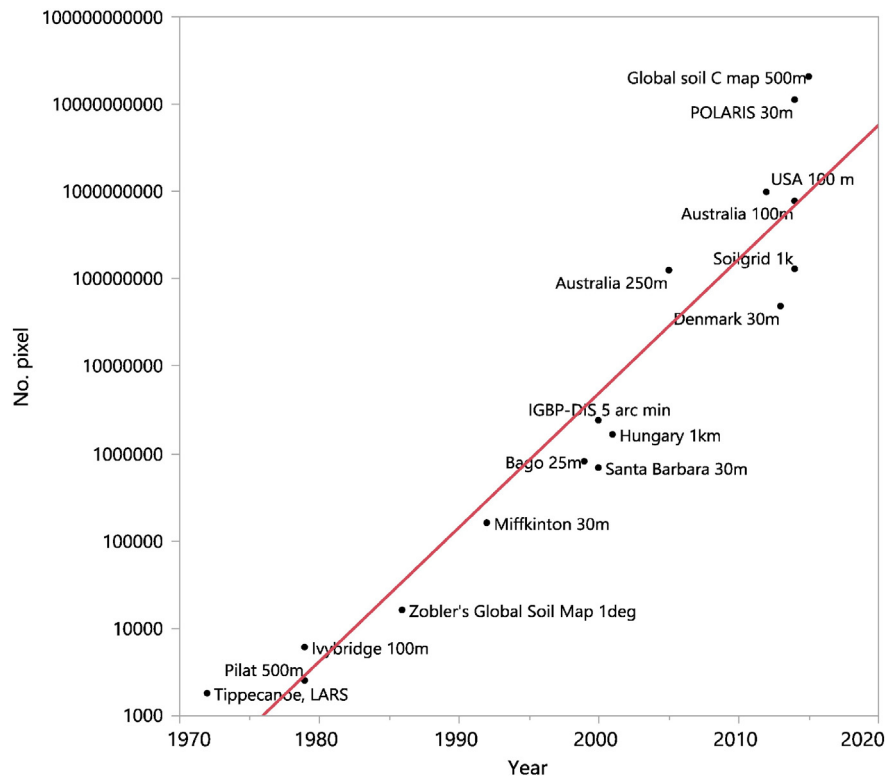


Fig. 6. The number of pixels as a function of year showing the exponential increase in the number of pixels in digital soil maps.

prioritised in several studies by the UN and other international organizations. Soil data is demanded to address contemporary natural resource management issues such as climate change, food security, water security and biodiversity management. In addition, many of the developments in digital soil mapping came from universities across the world that now were able to produce soil maps that previously was the sole domain of national soil survey centres (Adhikari et al., 2013, 2014b; Akpa et al., 2014).

The concept of digital mapping is not only happening in soil science. For example, similar approaches are used in the mapping of species distribution or Species Distribution Modelling (SDM) (Franklin, 2010). Data mining techniques such as classification and regression trees were used to model the distribution of plant species. Inevitably, digital

mapping approaches using data mining and environmental covariates will happen in soil mapping. However, defining and formalising it as a formal discipline boosted the adoption of digital soil mapping. The development of the IUSS working group on Digital Soil Mapping and the biennial global workshops ensure a steady integration of new science with emerging technologies. The workshops are a meeting place for soil surveyors, practitioners, and pedometricians to discuss progress, research applications and operational gaps. As an example, scientists developed sampling techniques that optimise the coverage of covariate or feature space, while practitioners informed researchers on operational constraints, how the sampling approach needs to be flexible, efficient, and compatible with a project area's land use and terrain (Kidd et al., 2015).

Table 1

Some major developments in soil mapping leading to the establishment of digital soil mapping *scorpan* concept.

| Period | Main development                                                             | Examples                                                               |
|--------|------------------------------------------------------------------------------|------------------------------------------------------------------------|
| 1960s  | Topofunctions                                                                | Walker et al. (1968)                                                   |
|        | Integrated clorpt model                                                      | Jenny et al. (1968)                                                    |
| 1970s  | Soil information systems                                                     | IUSS Working Group on Soil Information System (Sadovski and Bie, 1977) |
|        | Digital soil cartography                                                     | Webster et al. (1979)                                                  |
|        | Remote sensing and soils                                                     | LARS (Cipra, 1973)                                                     |
|        | Digital application of environmental similarity method                       | Legros and Bonneric (1979)                                             |
| 1980s  | Soil geostatistics                                                           | Hajrasuliha et al. (1980), Burgess and Webster (1980)                  |
|        | Geographical information system                                              | Burrough (1986)                                                        |
| 1990s  | Pedometrics: nonlinear geostatistics, fuzzy clustering, uncertainty analysis | McBratney and Gruijter (1992), Odeh et al. (1992b)                     |
|        | Digital terrain modelling                                                    | Gallant and Wilson (1996)                                              |
|        | Environmental correlation & clorpt model                                     | Moore et al. (1993), McKenzie and Ryan (1999)                          |
|        | Mapping using reference area                                                 | Lagacherie et al. (1995)                                               |
|        | Mapping using expert knowledge and expert system                             | Skidmore et al. (1996), Zhu et al. (1997)                              |
|        | Use of data mining tools                                                     | Pachepsky et al. (1996), Bui and Moran (2003)                          |
| 2000s  | Pedotransfer functions                                                       | Wösten et al. (1995)                                                   |
|        | Digital soil mapping                                                         | McBratney et al. (2003)                                                |
|        | Digital soil assessment                                                      | Carré et al. (2007)                                                    |
|        | Globalsoilmap                                                                | Arrouays et al. (2014)                                                 |
|        | Proximal soil sensing                                                        | Viscarra Rossel et al. (2010)                                          |
|        | Digital soil morphometrics                                                   | Hartemink and Minasny (2014)                                           |

Digital soil mapping has transformed soil survey and cartography more than anything else in recent times, and it is being tested and routinely used in soil mapping programmes around the world. Several studies have shown that digital soil mapping produces comparable or more accurate maps compared to conventional mapping (Bazaglia Filho et al., 2013; Collard et al., 2014). Lorenzetti et al. (2015) compared 1:5,000,000 map of Italian soil regions and digital soil mapping at 1 km grid resolution in predicting the WRB reference soil groups (RSGs). They showed that the digital soil map had a higher accuracy compared to traditional maps. The mean of positive predictive value (the probability that an event occurs when the model estimates it to occur) was 0.44 for digital soil maps, while traditional maps only achieved a value of 0.22. In addition, a study by Brungard et al. (2015) showed that covariates selected by soil scientists familiar with the study area did not result in the most accurate models as compared to covariates which were automatically selected by machine learning algorithms.

### 5.3. Capacity building and training: an example from experiences in Australia

While Australia is one of the forefront in the development of digital soil mapping (Bui, 2006), its adoption by soil surveyors for “real world” application was not until 2010. After applying mapping techniques at the research level, we decided to disseminate the tools to end users. The first digital soil mapping course in Australia was presented at the University of Sydney in February 2011 in response to the request by ACLEP (Australian Collaborative Land Evaluation Program) to provide DSM skills across all states and territories in Australia. The course “Getting Operational – An introductory course of practical exercises for Digital Soil Mapping” was structured to develop a basic level of skills in operational DSM. The course has a lot of hands-on computer time, demonstrating how to apply DSM techniques developed in research for real-world application. The training has helped to start digital soil mapping activities throughout Australia.

At the same time, government agencies are requested to provide land resource information for land use intensification, and irrigation development (Kidd et al., 2012; Harms et al., 2015). Digital soil mapping techniques combined with a land suitability framework facilitated the rapid evaluation of regional-scale agricultural potential in remote area (Harms et al., 2015).

The Australian Soil and Landscape Grid, launched in December 2014 at the 2014 National Soil Science Conference in Melbourne, provides information on soil attributes across the continent according to the GlobalSoilMap specification (Arrouays et al., 2014). Working in partnership with CSIRO and all of the state agencies, methods developed from research were put into practice, including the flexible equal-area spline depth function which is combined with modern numerical methods and advanced geostatistical techniques for mapping continuous depth distribution of soil properties (Malone et al., 2009; Malone et al., 2011). These methods were successfully applied for the national scale mapping in the USA (Odgers et al., 2012), Denmark (Adhikari et al., 2014a), Korea (Hong et al., 2013), and Nigeria (Akpa et al., 2014) as part of the GlobalSoilMap project.

The experience in Australia tells us that in order for the new methods to be used and applied, we need to disseminate the tools for the users. By working together with state-agencies, we demonstrated the capability of new technologies and fulfilling their needs. DSM is an evidence-based technique that is part of the everyday job (operational), and it is the way to make soil maps. The experience in Australia demonstrates the need for training in digital soil mapping techniques for young scientists and staff involved in land resource management.

### 5.4. Where is it going?

Currently most digital soil mapping is focussed at medium resolution (D3 according to McBratney et al. (2003), grid spacing of 20 m to

200 m) for catchment and regional studies (van Zijl et al., 2014). High resolution digital soil mapping is a focus of proximal soil sensing (Viscarra Rossel et al., 2010). Meanwhile low resolution digital maps seem to be the focus of some national and global studies (Dai et al., 2013). Digital soil mapping has moved from a bright new research topic to a normal subdiscipline in soil science which is a universal process (Krasilnikov, 2014). Here we foresee some of the future:

- Global collaboration project, rather than top down approach as initiated by the GlobalSoilMap project. A complete global coverage of soil properties according to GlobalSoilMap specification spacing by 2025, and at 30 m resolution by 2030.
- Global space-time coverage of dynamic soil properties (such as soil organic carbon (OC) and pH) at 100 m spatial resolution and 5 year temporal spacing.
- The availability of a finer resolution and diverse covariates, such as microwave, radar and infrared spectrometers.
- A more-efficient time and cost-effective soil sampling and analysis using proximal sensors.
- New algorithms to deal with big data.
- Soil property spatial now casting, predicting current soil moisture, water erosion or soil OC status based on new remotely-sensed information.

## 6. Conclusions

Finally, we draw some lessons from this brief history:

- Digital soil mapping has shifted from a research phase into operational use.
- Confluence of emerging topics and the right timing enables the success of a new paradigm.
- Research that is too precocious is left unrecognised and unknown. Many Sleeping Beauties in digital soil mapping remained unnoticed and uncited.
- As shown in several examples, such work warrants rediscovery.
- Researchers in search of new ideas should, perhaps paradoxically, peruse old and obscure scientific literature.
- Research outputs need to be disseminated not only as journal publications, but they need to be made available for practical use. Distributing computer codes and “how to” manual ensures that the output can be applied in real-world situation.
- Training builds capacity; as has been demonstrated within Australia.

## Acknowledgements

The authors thank Prof. Alfred Hartemink from the University of Wisconsin, and Dr. Brendan Malone from the University of Sydney for their useful suggestions and comments on the draft. Budiman Minasny is funded through an ARC Future Fellowship.

## 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.
- Adhikari, K., Hartemink, A.E., Minasny, B., Kheir, R.B., Greve, M.B., Greve, M.H., 2014a. Digital mapping of soil organic carbon contents and stocks in Denmark. *PLoS One* 9, e105519.
- Adhikari, K., Minasny, B., Greve, M.B., Greve, M.H., 2014b. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma* 214–215, 101–113.
- Agbu, P.A., Fehrenbacher, D.J., Jansen, I.J., 1990. Soil property relationships with SPOT satellite digital data in east central Illinois. *Soil Sci. Soc. Am. J.* 54, 807–812.
- Akpa, S.I., Odeh, I.O., Bishop, T.F., Hartemink, A.E., 2014. Digital mapping of soil particle-size fractions for Nigeria. *Soil Sci. Soc. Am. J.* 78, 1953–1966.
- Al-Abbas, A., Swain, P., Baumgardner, M., 1972. Relating organic matter and clay content to the multispectral radiance of soils. *Soil Sci.* 114, 477–485.

- Arrouays, D., Pelissier, P., 1994. Modeling carbon storage profiles in temperate forest humic loamy soils of France. *Soil Sci.* 157, 185–192.
- Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M., Hong, S.Y., Lagacherie, P., Lelyk, G., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.D., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A., Zhang, G.L., 2014. GlobalSoilMap. Toward a fine-resolution global grid of soil properties. *Adv. Agron.* 93–134.
- Bazaglia Filho, O., Rizzo, R., Lepsh, I.F., Prado, H.d., Gomes, F.H., Mazza, J.A., Demattê, J.A.M., 2013. Comparison between detailed digital and conventional soil maps of an area with complex geology. *Rev. Bras. Ciênc. Solo* 37, 1136–1148.
- Bell, J.C., Cunningham, R.L., Havens, M.W., 1992. Calibration and validation of a soil-landscape model for predicting soil drainage class. *Soil Sci. Soc. Am. J.* 56, 1860–1866.
- Bellon-Maurel, V., McBratney, A., 2011. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – critical review and research perspectives. *Soil Biol. Biochem.* 43, 1398–1410.
- Ben-Dor, E., Inbar, Y., Chen, Y., 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* 61, 1–15.
- Bliss, N.B., Reay, W.U., 1989. Small-scale digital soil maps for interpreting natural resources. *J. Soil Water Conserv.* 44, 30–34.
- Bliss, N.B., Waltman, S.W., Petersen, G.W., 1995. Preparing a soil carbon inventory for the United States using geographic information systems. *Soil Glob. Chang.* 275–295.
- Boettinger, J.L., 2010. Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. Springer Science & Business Media.
- Bou Kheir, R., Greve, M.H., Bocher, P.K., Greve, M.B., Larsen, R., McCloy, K., 2010. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: the case study of Denmark. *J. Environ. Manag.* 91, 1150–1160.
- Brown, D.J., 2006. A historical perspective on soil-landscape modeling. In: Grunwald, S. (Ed.), *Environmental Soil-Landscape Modeling: Geographic Information Technologies and Pedometrics*. Taylor & Francis, Boca Raton, p. 61.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132, 273–290.
- Brungard, C.W., Boettinger, J.L., Duniway, M.C., Wills, S.A., Edwards, T.C., 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma* 239, 68–83.
- Bui, E., 2006. A review of digital soil mapping in Australia. In: Lagacherie, A.B.M., Voltz, M. (Eds.), *Developments in Soil Science*. Elsevier, pp. 25–37.
- 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.
- Burgess, T., Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. *J. Soil Sci.* 31, 315–331.
- Burrough, P.A., 1986. Principles of Geographical Information Systems for Land Resources Assessment. Taylor & Francis.
- Burrough, P., Bie, S.W., 1984. Soil information systems technology, Proceedings of the Sixth Meeting of the ISSS Working Group on Soil Information Systems. Centre for Agricultural Publishing and Documentation, Bolkesjö, Norway.
- Bushnell, T., 1929. Aerial photography and soil survey. *Am. Assoc. Soil Surv. Bull.* 10, 23–28.
- Bushnell, T., 1943. Some aspects of the soil catena concept. *Soil Sci. Soc. Am. J.* 7, 466–476.
- Butler, B.E., 1959. Periodic Phenomena in Landscapes as a Basis for Soil Studies. CSIRO Australia.
- Carré, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007. Digital soil assessments: beyond DSM. *Geoderma* 142, 69–79.
- Chaney, N.W., Hempel, J.W., Odgers, N., McBratney, A.B., Wood, E.F., 2014. Spatial disaggregation and harmonization of gSSURGO, ASA, CSSA, & SSSA International Annual Meeting. American Society of Agronomy, Long Beach, CA.
- Cipra, J.E., 1973. Mapping soil associations using ERTS MSS data. LARS Technical Reports, p. 117.
- Collard, F., Kempen, B., Heuvelink, G.B.M., Saby, N.P.A., Richer de Forges, A.C., Lehmann, S., Nehlig, P., Arrouays, D., 2014. Refining a reconnaissance soil map by calibrating regression models with data from the same map (Normandy, France). *Geoderma Reg.* 1, 21–30.
- Cook, S., Corner, R., Groves, P., Grealish, G., 1996a. Use of airborne gamma radiometric data for soil mapping. *Soil Res.* 34, 183–194.
- Cook, S.E., Corner, R.J., Grealish, G., Gessler, P.E., Chartres, C.J., 1996b. A rule-based system to map soil properties. *Soil Sci. Soc. Am. J.* 60, 1893–1900.
- Dai, Y., Shangquan, W., Duan, Q., Liu, B., Fu, S., Niu, G., 2013. Development of a China dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling. *J. Hydrometeorol.* 14, 869–887.
- De Gruijter, J.J., Bie, S.W., 1975. A discrete approach to automated mapping of multivariate systems. *Trans. Comm. III Inter. Cartogr. Assoc. Enschede*, pp. 17–28.
- Dobos, E., Norman, B., Worstell, B., et al., 2002. The use of DEM and satellite data for regional scale soil databases. *Agrokém. Talajt.* 51, 263–272.
- Field, D.J., Minasny, B., 2008. Comments on “modeling energy inputs to predict pedogenic environments using regional environmental databases”. *Soil Sci. Soc. Am. J.* 72, 858–859.
- Florinsky, I.V., 2012. The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the 125th anniversary of its publication). *Eurasian Soil Sci.* 45, 445–451.
- Franklin, J., 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Prog. Phys. Geogr.* 19, 474–499.
- Franklin, J., 2010. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press.
- Frazier, B.E., Cheng, Y., 1989. Remote sensing of soils in the Eastern Palouse region with Landsat Thematic Mapper. *Remote Sens. Environ.* 28, 317–325.
- Gallant, J.C., Wilson, J.P., 1996. TAPES-G: a grid-based terrain analysis program for the environmental sciences. *Comput. Geosci.* 22, 713–722.
- Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J., 1995. Soil-landscape modelling and spatial prediction of soil attributes. *Int. J. Geogr. Inf. Syst.* 9, 421–432.
- Giltrap, D.J., 1983. Computer production of soil maps, I. Production of grid maps by interpolation. *Geoderma* 29, 295–311.
- Gray, J.M., Bishop, T.F., Yang, X., 2015. Pragmatic models for the prediction and digital mapping of soil properties in eastern Australia. *Soil Res.* 53 (1), 24–42.
- 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., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152, 195–207.
- Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. *Soil Sci. Soc. Am. J.* 75, 1201–1213.
- Haines, W.B., Keen, B.A., 1925a. Studies in soil cultivation. II. A test of soil uniformity by means of dynamometer and plough. *J. Agric. Sci.* 15, 387–394.
- Haines, W.B., Keen, B.A., 1925b. Studies in soil cultivation. III. Measurements on the Rothamsted classical plots by means of dynamometer and plough. *J. Agric. Sci.* 15, 395–406.
- Hajrasulhi, S., Baniabbassi, N., Metthey, J., Nielsen, D.R., 1980. Spatial variability of soil sampling for salinity studies in Southwest Iran. *Irrig. Sci.* 1, 197–208.
- Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., Schröder, B., 2012. Spatial disaggregation of complex soil map units: a decision-tree based approach in Bavarian forest soils. *Geoderma* 185–186, 37–47.
- Harms, B., Brough, D., Philip, S., Bartley, R., Clifford, D., Thomas, M., Willis, R., Gregory, L., 2015. Digital soil assessment for regional agricultural land evaluation. *Glob. Food Secur.* 5, 25–36.
- Hartemink, A.E., McBratney, A., 2008. A soil science renaissance. *Geoderma* 148, 123–129.
- Hartemink, A.E., Minasny, B., 2014. Towards digital soil morphometrics. *Geoderma* 230–231, 305–317.
- Hartemink, A.E., McBratney, A., de Lourdes Mendonça-Santos, M., 2008. Digital Soil Mapping with Limited Data. Springer Science & Business Media.
- Hartemink, A.E., Krasilnikov, P., Bockheim, J.G., 2013. Soil maps of the world. *Geoderma* 207–208, 256–267.
- Hempel, J., McBratney, A., Arrouays, D., McKenzie, N., Hartemink, A., 2014. GlobalSoilMap project history. In: Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A.C., McBratney, A.B. (Eds.), *GlobalSoilMap: Basis of the global spatial soil information system*. Taylor & Francis, London, pp. 3–8.
- Henderson, B.L., Bui, E.N., Moran, C.J., Simon, D.A.P., 2005. Australia-wide predictions of soil properties using decision trees. *Geoderma* 124, 383–398.
- Hong, S.Y., Minasny, B., Han, K.H., Kim, Y., Lee, K., 2013. Predicting and mapping soil available water capacity in Korea. *Peer J.* 1, e71. <http://dx.doi.org/10.7717/peerj.71>.
- Hudson, B.D., 1992. The soil survey as paradigm-based science. *Soil Sci. Soc. Am. J.* 56, 836–841.
- Irvin, B.J., Ventura, S.J., Slater, B.K., 1997. Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. *Geoderma* 77, 137–154.
- Jenny, H., Salem, A.E., Wallis, J.R., 1968. Interplay of soil organic matter and soil fertility with state factors and soil properties. “Organic matter and soil fertility”. *Pontif. Acad. Sci. Scr. varia* 32, 5–36.
- Karale, R., Bali, Y., Rao, K.S., 1983. Soil mapping using remote sensing techniques. *Proc. Indian Acad. Sci. Chem. Sci.* 6, 197–208.
- Ke, Q., Ferrara, E., Radicchi, F., Flammini, A., 2015. Defining and identifying sleeping beauties in science. *Proc. Natl. Acad. Sci.* 112, 7426–7431.
- Keen, B.A., Haines, W.B., 1925. Studies in soil cultivation. I. The evolution of a reliable dynamometer technique for use in soil cultivation experiments. *J. Agric. Sci.* 15, 375–386.
- Kempen, B., Heuvelink, G., Brus, D., Stoorvogel, J., 2010. Pedometric mapping of soil organic matter using a soil map with quantified uncertainty. *Eur. J. Soil Sci.* 61, 333–347.
- Kidd, D.B., Webb, M.A., Grose, C.J., Moreton, R.M., Malone, B.P., McBratney, A.B., Minasny, B., Viscarra-Rossel, R.A., Cotching, W.E., Sparrow, L.A., Smith, R., 2012. Digital soil assessment: guiding irrigation expansion in Tasmania, Australia. *Digital Soil Assessments and Beyond – Proceedings of the Fifth Global Workshop on Digital Soil Mapping*, pp. 3–8.
- Kidd, D., Malone, B., McBratney, A., Minasny, B., Webb, M., 2015. Operational sampling challenges to digital soil mapping in Tasmania, Australia. *Geoderma Reg.* 4, 1–10.
- Kornblau, M., Cipra, J., 1983. Investigation of digital Landsat data for mapping soils under range vegetation. *Remote Sens. Environ.* 13, 103–112.
- Kosaki, T., Torii, K., Kyuma, K., 1982. Automated soil map compilation. *Soil Sci. Plant Nutr.* 28, 389–399.
- Krasilnikov, P., 2014. Digital soil assessments and beyond, by Budiman Minasny. Proceedings of the Fifth Global Workshop on Digital Soil Mapping, Sydney, Australia, 10–13 April 2012. In: Malone, Brendan P., McBratney, Alex B. (Eds.), *Geoderma* 213. CRC Press, Boca Raton, pp. 131–132 (2012).
- Kristof, S., Baumgardner, M., Johansen, C., 1973. Spectral Mapping of Soil Organic Matter. LARS Technical Reports, p. 26.
- Lagacherie, P., 1992. Formalisation des lois de distribution des sols pour automatiser la cartographie pédologique à partir d'un secteur pris comme référence.
- Lagacherie, P., McBratney, A.B., 2006. Chapter 1 spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. *Dev. Soil Sci.* 3–22.
- Lagacherie, P., Legros, J.P., Burrough, P.A., 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma* 65, 283–301.
- Lagacherie, P., McBratney, A., Voltz, M., 2006. Digital Soil Mapping: An Introductory Perspective. Elsevier.
- Lathrop Jr., R.G., Aber, J.D., Bognar, J.A., 1995. Spatial variability of digital soil maps and its impact on regional ecosystem modeling. *Ecol. Model.* 82, 1–10.
- Legros, J.-P., 2006. Mapping of the Soil. Science Publishers.

- Legros, J., Bonneric, P., 1979. Modelisation informatique de la repartition des sols dans le Parc Naturel Régional du Pilat 4. Annales de l'Université de Savoie, Tome, pp. 63–68.
- Legros, J., Hensel, E., 1978. INRA 6-computer assisted cartographic system. In: Sadovski, A.N., Bie, S. (Eds.), *Developments in Soil Information Systems 2. Meeting of the ISSS Working Group on Soil Information Systems*. Centre for Agricultural Publishing and Documentation, Varna/Sofia, Bulgaria, pp. 95–96.
- Lorenzetti, R., Barbetti, R., Fantappiè, M., L'Abate, G., Costantini, E.A., 2015. Comparing data mining and deterministic pedology to assess the frequency of WRB reference soil groups in the legend of small scale maps. *Geoderma* 237, 237–245.
- 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., Minasny, B., Odgers, N.P., McBratney, A.B., 2014. Using model averaging to combine soil property rasters from legacy soil maps and from point data. *Geoderma* 232, 34–44.
- Mathews, H., Cunningham, R., Petersen, G., 1973. Spectral reflectance of selected Pennsylvania soils. *Soil Sci. Soc. Am. J.* 37, 421–424.
- McBratney, A., Gruijter, J.d., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *J. Soil Sci.* 43, 159–175.
- McBratney, A.B., Minasny, B., 2010. The sun has shone here antecedently. In: Viscarra Rossel, R.A., McBratney, A.B., Minasny, B. (Eds.), *Proximal Soil Sensing*. Springer, Netherlands, pp. 67–75.
- McBratney, A.B., Odeh, I.O.A., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma* 77, 85–113.
- McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. *Geoderma* 97, 293–327.
- McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52.
- McBratney, A.B., Minasny, B., Whelan, B., 2011. Defining proximal soil sensing. The Second Global Workshop on Proximal Soil Sensing Montreal.
- McKenzie, N.J., Austin, M.P., 1993. A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. *Geoderma* 57, 329–355.
- McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89, 67–94.
- Mendonça-Santos, M., dos Santos, H., 2006. The state of the art of Brazilian soil mapping and prospects for digital soil mapping. *Dev. Soil Sci.* 31, 39–601.
- Miller, B.A., Schaezel, R.J., 2014. The historical role of base maps in soil geography. *Geoderma* 230–231, 329–339.
- Minasny, B., Malone, B.P., McBratney, A.B., 2012. Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping 2012. Sydney. CRC Press, Australia.
- Moore, G.E., 1998. Cramming more components onto integrated circuits. *Proc. IEEE* 86, 82–85.
- Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. *Soil Sci. Soc. Am. J.* 57, 443–452.
- Moran, C.J., Bui, E.N., 2002. Spatial data mining for enhanced soil map modelling. *Int. J. Geogr. Inf. Sci.* 16, 533–549.
- Mulder, V.L., de Bruin, S., Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping – a review. *Geoderma* 162, 1–19.
- Odeh, I.O., Chittleborough, D.J., McBratney, A.B., 1992a. Fuzzy-c-means and kriging for mapping soil as a continuous system. *Soil Sci. Soc. Am. J.* 56, 1848–1854.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1992b. Soil pattern recognition with fuzzy-c-means: application to classification and soil-landform interrelationships. *Soil Sci. Soc. Am. J.* 56, 505–516.
- 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, 197–214.
- 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* 189–190, 153–163.
- Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., Clifford, D., 2014. Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma* 214–215, 91–100.
- Pachepsky, Y.A., Timlin, D., Várallyay, G., 1996. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Sci. Soc. Am. J.* 60, 727–733.
- Padarian, J., Minasny, B., McBratney, A.B., 2015w. Using Google's web-based platform for digital soil mapping. *Comput. Geosci.* 83, 80–88.
- Prescott, J.A., Taylor, J.K., 1930. The value of aerial photography in relation to soil surveys and classification. *CSIR Aust. J.* 3, 229–230.
- Ragg, J.M., 1977. The recording and organization of soil field data for computer areal mapping. *Geoderma* 19, 81–89.
- Rasmussen, C., Tabor, N.J., 2007. Applying a quantitative pedogenic energy model across a range of environmental gradients. *Soil Sci. Soc. Am. J.* 71, 1719–1729.
- Rasmussen, C., Southard, R.J., Horwath, W.R., 2005. Modeling energy inputs to predict pedogenic environments using regional environmental databases. *Soil Sci. Soc. Am. J.* 69, 1266–1274.
- Ruhe, R.V., 1956. Geomorphic surfaces and the nature of soils. *Soil Sci.* 82, 441–456.
- Runge, E.C.A., 1973. Soil development sequences and energy models. *Soil Sci.* 115, 183–193.
- Ryan, P.J., McKenzie, N.J., O'Connell, D., Loughhead, A.N., Leppert, P.M., Jacquier, D., Ashton, L., 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. *For. Ecol. Manag.* 138, 139–157.
- Sadovski, A.N., Bie, S.W., 1977. Developments in soil information system. Proceedings of the Second Meeting of the ISSS Working Groups on Soil Information System. Centre for Agricultural Publishing & Documentation, Wageningen.
- Sarathjith, M.C., Das, B.S., Wani, S.P., Sahrawat, K.L., 2014. Dependency measures for assessing the covariation of spectrally active and inactive soil properties in diffuse reflectance spectroscopy. *Soil Sci. Soc. Am. J.* 78, 1522–1530.
- Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: a review. *Prog. Phys. Geogr.* 27, 171–197.
- Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* 66, 988–998.
- Shi, X., Zhu, A.X., Burt, J.E., Qi, F., Simonson, D., 2004. A case-based reasoning approach to fuzzy soil mapping. *Soil Sci. Soc. Am. J.* 68, 885–894.
- Shovic, H.F., Montagne, C., 1985. Application of a statistical soil-landscape model to an order III wildland soil survey. *Soil Sci. Soc. Am. J.* 49, 961–968.
- Singh, A., 1980. Comparison of reconnaissance soil maps prepared by conventional method and Landsat imagery interpretation. *J. Indian Soc. Photo Interpretation Remote. Sens.* 8, 1–8.
- Skidmore, A.K., Ryan, P.J., Dawes, W., Short, D., Emmett, O.L., 1991. Use of an expert system to map forest soils from a geographical information system. *Int. J. Geogr. Inf. Syst.* 5, 431–445.
- Skidmore, A.K., Watford, F., Luckananurug, P., Ryan, P., 1996. An operational GIS expert system for mapping forest soils. *Photogramm. Eng. Remote. Sens.* 62, 501–511.
- Slater, B.K., 1994. Continuous Classification and Visualization of Soil layers: A Soil-Landscape Model of Pleasant Valley Wisconsin. University of Wisconsin-Madison, Madison.
- Stafford, J., Hendrick, J., 1985. Dynamic control of pan rupturing tines. *Trans. ASABE* 31, 9–13.
- Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. *Adv. Agron.* 107, 163–215.
- Stoner, E.R., Baumgardner, M., 1981. Characteristic variations in reflectance of surface soils. *Soil Sci. Soc. Am. J.* 45, 1161–1165.
- Stoner, E.R., Biehl, L.L., 1980. Development of a digital data base for reflectance-related soil information. LARS Technical Reports, p. 60.
- Subburayalu, S.K., Slater, B.K., 2013. Soil series mapping by knowledge discovery from an Ohio county soil map. *Soil Sci. Soc. Am. J.* 77, 1254–1268.
- Sun, Y., Ma, D., Schulze Lammers, P., Schmittmann, O., Rose, M., 2006. On-the-go measurement of soil water content and mechanical resistance by a combined horizontal penetrometer. *Soil Tillage Res.* 86, 209–217.
- Tomlinson, R., 1978. Design considerations for digital soil map systems. 11th Congress of Soil Science. ISSS, Edmonton, Canada.
- Troeh, F.R., 1964. Landform parameters correlated to soil drainage. *Soil Sci. Soc. Am. J.* 28, 808–812.
- Van Raan, A.F., 2004. Sleeping beauties in science. *Scientometrics* 59, 467–472.
- van Zijl, G.M., Bouwer, D., van Tol, J.J., le Roux, P.A.L., 2014. Functional digital soil mapping: a case study from Namarroi, Mozambique. *Geoderma* 219–220, 155–161.
- Vaudin, M., Vieira, S.R., Vachaud, G., Nielsen, D.R., 1983. The use of cokriging with limited field soil observations. *Soil Sci. Soc. Am. J.* 47, 175–184.
- Vieira, S.R., Nielsen, D., Biggar, J., 1981. Spatial variability of field-measured infiltration rate. *Soil Sci. Soc. Am. J.* 45, 1040–1048.
- Vieira, S., Hatfield, J., Nielsen, D., Biggar, J., 1983. Geostatistical theory and application to variability of some agronomical properties. *Hilgardia* 51, 1–73.
- Viscarra Rossel, R.A., McBratney, A.B., Minasny, B., 2010. *Proximal Soil Sensing*. Springer Science & Business Media.
- Volobuyev, V., 1964. Ecology of Soils Academy of Sciences of the Azerbaijan SSR. Israel Program for Scientific Translations, Jerusalem.
- Walker, P., Hall, G., Protz, R., 1968. Relation between landform parameters and soil properties. *Soil Sci. Soc. Am. J.* 32, 101–104.
- Webster, R., 1994. The development of pedometrics. *Geoderma* 62, 1–15.
- Webster, R., Burrough, P.A., 1972. Computer-based soil mapping of small areas from sample data. *J. Soil Sci.* 23, 222–234.
- Webster, R., Harrod, T., Staines, S., Hogan, D., 1979. Grid sampling and computer mapping of the Ivybridge area, Devon. Technical Monograph. Soil Survey of England and Wales, Harpenden, p. 64.
- Wösten, J., Finke, P., Jansen, M., 1995. Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics. *Geoderma* 66, 227–237.
- Yaalon, D.H., 1989. The Earliest Soil Maps and Their Logic. *Bulletin of the International Society of Soil Science*. International Society of Soil Science, Wageningen p. 24.
- Zhu, A.-X., Band, L.E., Dutton, B., Nimlos, T.J., 1996. Automated soil inference under fuzzy logic. *Ecol. Model.* 90, 123–145.
- Zhu, A.X., Band, L., Vertessy, R., Dutton, B., 1997. Derivation of soil properties using a soil land inference model (SoLIM). *Soil Sci. Soc. Am. J.* 61, 523–533.