Geoderma 264 (2016) 301–311

Contents lists available at ScienceDirect

Geoderma

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

Digital soil mapping: A brief history and some lessons

Budiman Minasny *, 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 6 August 2015

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.

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 Reybold (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

* Corresponding author.
E-mail address: Budiman.minasny@sydney.edu.au (B. Minasny).

http://dx.doi.org/10.1016/j.geoderma.2015.07.017
0016-7061/© 2015 Published by Elsevier B.V.

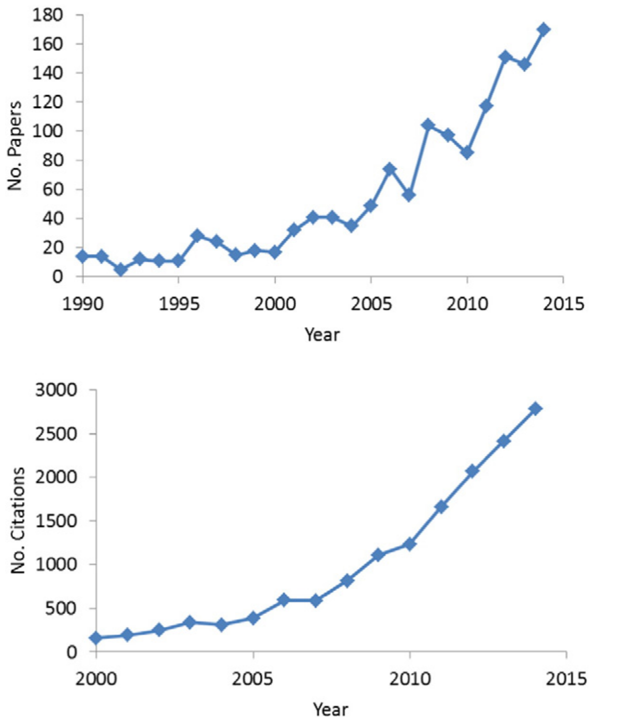

Fig. 1. Number of papers and number of citations from keyword search "digital" "soil" "mapping" from the Scopus database (data extracted in February 2015).

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.

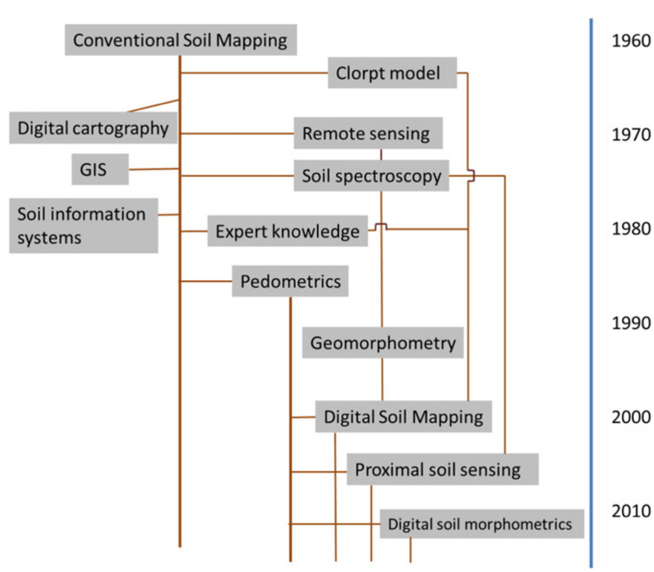

Fig. 2. A diagram illustrating the evolution and confluence of different concepts in digital soil mapping used in this paper.

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