Geoderma 152 (2009) 195–207

Contents lists available at ScienceDirect

Geoderma

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

Review

Multi-criteria characterization of recent digital soil mapping and modeling approaches

S. Grunwald *

University of Florida, Soil and Water Science Department, 2169 McCarty Hall, PO Box 110290, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 30 December 2008
Received in revised form 24 April 2009
Accepted 10 June 2009
Available online 16 July 2009

Keywords:
Digital soil mapping
Digital soil modeling
Pedometrics
Quantitative methods
Soils

ABSTRACT

The history of digital soil mapping and modeling (DSMM) is marked by adoption of new mapping tools and techniques, data management systems, innovative delivery of soil data, and methods to analyze, integrate, and visualize soil and environmental datasets. DSMM studies are diverse with specialized, mathematical prototype models tested on limited geographic regions and/or datasets and simpler, operational DSMM used for routine mapping over large soil regions. Research-focused DSMM contrasts with need-driven DSMM and agency-operated soil surveys. Since there is no universal equation or digital soil prediction model that fits all regions and purposes the proposed strategy is to characterize recent DSMM approaches to provide recommendations for future needs at local, national and global scales. Such needs are not solely soil-centered, but consider broader issues such as land and water quality, carbon cycling and global climate change, sustainable land management, and more. A literature review was conducted to review 90 DSMM publications from two high-impact international soil science journals — Geoderma and Soil Science Society of America Journal. A selective approach was used to identify published studies that cover the multi-factorial DSMM space. The following criteria were used (i) soil properties, (ii) sampling setup, (iii) soil geographic region, (iv) spatial scale, (v) distribution of soil observations, (vi) incorporation of legacy/historic data, (vii) methods/model type, (viii) environmental covariates, (ix) quantitative and pedological knowledge, and (x) assessment method. Strengths and weaknesses of current DSMM, their potential to be operationalized in soil mapping/modeling programs, research gaps, and future trends are discussed. Modeling of soils in 3D space and through time will require synergistic strategies to converge environmental landscape data and denser soil datasets. There are needs for more sophisticated technologies to measure soil properties and processes at fine resolution and with accuracy. Although there are numerous quantitative models rooted in factorial models that predict soil properties with accuracy in select geographic regions they lack consistency in terms of environmental input data, soil properties, quantitative methods, and evaluation strategies. DSMM requires merging of quantitative, geographic and pedological expertise and all should be ideally in balance.

Contents

* Tel.: +1 352 392 1951x204; fax: +1 352 392 3902.
E-mail address: sabgru@ufl.edu.

0016-7061/$ – see front matter
doi:10.1016/j.geoderma.2009.06.003

1. Introduction

Soil properties are continuously modified by internal factors and anthropogenic impacts generating complex spatial soil patterns. Our understanding and ability to describe them have undergone tremendous change as outlined by Hartemink and McBratney (2008) who argued that we are in an age called "soil science renaissance". Introduction of digital technologies, such as remote and soil sensing, computer processing speed, management of spatial data, quantitative method to describe soil patterns and processes, and scientific visualization methods have provided new opportunities to predict soil properties and processes. The history of digital soil mapping and modeling (DSMM) is marked by adoption of new tools and techniques to analyze, integrate, and visualize soil and environmental datasets (Grunwald, 2006).

To quantify soil patterns mapping or modeling methods can be used. Modeling is defined as "use of mathematical equations to simulate and predict real events and processes", whereas mapping emphasizes to "make a map" or "to depict something on a map". Digital soil mapping (DSMa) [or predictive soil mapping] can be defined as "computer-assisted production of digital representations of soil type or soil properties, which involve the creation and population of spatially-explicit information by the use of field and laboratory methods, coupled with spatial and non-spatial soil inference systems". A comprehensive review of DSMa was provided in McBratney et al. (2003) and an overview of pedometric techniques used in DSMa by McBratney et al. (2000) and Grunwald (2006).

To predict soil properties and classes the SCORPAN approach (Eq. (1)) has been formulated by McBratney et al. (2003) rooted in earlier works by H. Jenny and V.V. Dokuchaev.

$$S_a\left[x,y,\sim t\right] \text{ or } S_c\left[x,y,\sim t\right] = f\left(s\left[x,y,\sim t\right],c\left[x,y,\sim t\right],o\left[x,y,\sim t\right],r\left[x,y,\sim t\right],p\left[x,y,\sim t\right],a\left[x,y,\sim t\right],n\right) \tag{1}$$

where

S_a	soil attribute
S_c	soil class
s	soils, other attributes of the soil at a point
c	climate factor
o	organisms, vegetation or fauna or human activity
r	relief (topography, landscape attributes)
p	parent material, lithology
a	age, the time factor
n	space, spatial position
t	time (where t is defined as an approximate time).

Digital soil modeling (DSMo) is more comprehensive than DSMa because it entails description and/or prediction of soils in space and through time usually implemented using mechanistic simulation models. Pedometrics integrates DSMa and DSMo into a spatially and temporally explicit framework merging soil science, environmental science, GIScience, cartography, statistics, geostatistics, and mathematics for the study of the distribution and genesis of soils.

Hartemink et al. (2001) noted a strong decline in soil mineralogy, soil morphology and soil genesis research in comparison to a strong increase in pedometrics applications from 1967 to 2001. Historically, soil science has been rooted in agriculture, geology, and chemistry. But the emphasis has shifted from classification and inventory to understanding and quantifying spatially and temporally soil patterns in relation to hydrologic cycle and ecosystems health. This environmental-centered approach views soils as integral part of an ecosystem interacting with environmental factors generating complex patterns and processes that co-evolve through time.

In this paper a comprehensive review of DSMM is provided covering a wide range of specialized, mathematical prototype models tested on limited geographic regions and/or datasets and simpler, operational DSMM used for routine mapping over large soil regions.

The aim of this study was to assess the usefulness of recent DSMM approaches to meet specific needs at local, national, and global scales. Such needs are not solely soil-centered, but consider broader issues such as soil and water quality, food security, carbon cycling and sequestration, global climate change, sustainable land management, and more. Strengths and weaknesses of current DSMM, their potential to be operationalized in soil mapping/modeling programs, research gaps, and future trends are discussed.

2. Materials and methods

Publications of two high-impact and international soil science journals – Geoderma and Soil Science Society of America Journal – were reviewed covering the period 2007–2008 to identify studies focused on DSMM. The Soil Sci. Soc. Am. J. Issues 2007 Vol. 71 No. 6 to 2008 Vol. 72 No. 5 and Geoderma 2008 Vol. 140 Issue 4 to Vol. 146 Issues 1–2 with a total number of 90 papers were considered in the analysis. A selective approach was used to identify published DSMM studies that cover the multi-factorial DSMM space. Only papers that focused on predictive modeling of soil properties or soil classes at unsampled locations and/or future time periods qualified, whereas publications focused on experimental and analytical laboratory/plot studies that described soils and soil–soil relationships were not considered in this investigation.

The following criteria to characterize DSMM studies were used:

(i) Predicted soil attributes/classes (SPred)
(ii) Sampling (Samp) total number of soil samples (N), soil profiles/number of observation sites (SP), number of soil layers/horizons (SL), and frequency/time period (F)
(iii) Soil geographic region (SGR)
(iv) Spatial scale (SS)
(v) Distribution of soil observations (sampling design) (Dist)
(vi) Incorporation of legacy/historic soil data (Leg)
(vii) Digital soil prediction method/model (DSMM)

(viii) Covariates (Cov) explicitly incorporated into DSMM considering the following landscape factors: soils (S), climate (C), organisms/vegetation (O), relief (R), and parent material (P); distance measures (D), and easting and northing (XY) coordinates; and type(s) of data collection
(ix) Knowledge (Know)/specialized expertise required to implement a DSMM schema. Knowledge was assessed using an ordinal scale (1–5) to distinguish knowledge on quantitative methods/modeling (Q) and knowledge in pedology and environmental science (interpretation of soil and environmental factor relationships and/or pedogenic and ecosystem processes) (PE):
- *Know-Q* with 1: very basic to 5: complex knowledge on quantitative methods/modeling
- *Know-PE* with 1: very basic to 5: comprehensive pedological and environmental interpretation of findings

(x) DSMM assessment mode (Ass) with: model development only (M), cross-validation (CV), and validation (V).

Predictive capabilities (considering accuracy, precision, and bias of predictions), automation capabilities, and costs of DSMM studies could not be considered due to the unavailability of quantitative data in reviewed publications.

3. Results

3.1. Prediction of soil properties and classes at plot/field and coarse landscape scales

Prediction of soil properties and classes at plot/field scale encompassed 35.6% (*N*: 32) of all reviewed studies, of which 87.5% (*N*: 28) predicted soil properties and 12.5% (*N*: 4) soil classes (Table 1). A total of 31.3% (*N*: 10) of plot/field scale studies used a mechanistic modeling approach (DSMo) to predict soil properties. For example, Bricklemyer et al. (2007) employed the CENTURY model and Causarano et al. (2007) EPIC to model soil organic carbon. At coarser landscape scales a total of 48.9% (*N*: 44) DSMM studies were reviewed covering study area extents of 0.82 km^2 to the world (Table 2). Out of the 44 studies 36.4% (*N*: 16) focused on prediction of standard soil properties, environmental quality and/or nutrient management, 34.1% (*N*: 15) on soil carbon, 20.5% (*N*: 9) on hydrologic properties, and 18.2% (*N*: 8) on soil classes. Fourteen DSMM could not be grouped into plot/field and landscape scale applications (Table 3) of which 3 were DSMo studies.

3.2. Spatial scale

DSMM research was conducted at various spatial scales from specific sites/fields (Table 1) (e.g. Allison et al., 2007; Bricklemyer et al., 2007; Oorts et al., 2007) up to large soil regions with thousands or millions of square kilometers (compare Batjes (2008), Corstanje et al. (2008), and Zhang et al. (2008a)) (Table 2). For example, Zhang et al. (2008a) predicted SOC using a sampling density of 9.89×10^{-4} sites km^{-2} (806,400 km^2), Eldeiry and Garcia (2008) predicted soil salinity using 0.51 sites km^{-2} (500 km^2), and Grunwald and Reddy (2008) predicted TN and TP using a sampling density of 5.54 sites km^{-2} (48 km^2). As the study area extent increased typically the sampling density decreased. McBratney and Pringle (1999) showed in a comprehensive review that spatial autocorrelations of numerous soil properties are smaller than 500 m (often < 100 m) for pH, clay, sand, SOC, and nitrate-nitrogen based on studies performed in various ecosystem types and soil-landscape settings. This fine-scale variability is often smoothed out as the spatial extent of DSMM studies increases whereby often only the top soil is analyzed. Only 27.7% (*N*: 25) of the investigated DSMM studies sampled multiple soil layers/horizons (soil databases are not considered in these numbers). Only one paper by Carrara et al. (2007) provided three-dimensional (3D) soil representations and one paper by Zhang et al. (2008a) used an explicit 3D Geographic Information System (GIS)-based tabulation of soil data.

3.3. Temporal scale

There are multiple dimensions to be considered to characterize soils including spatial space (*x*, *y* and *z* directions) and temporal. Out of 90 studies only 14.4% (*N*:13) collected soil properties at more than one time period (soil databases are not considered in this percentage) (compare Causarano et al. (2007), Douaik et al. (2007), Jiang et al. (2007), Lobell et al. (2007), Oorts et al. (2007), Shukla et al. (2007), Bauer et al. (2008), Bonilla et al. (2007), Gomez et al. (2008), Grunwald et al. (2008), Kazemi et al. (2008), Sadeghi et al. (2008), and Vasques et al. (2009)). In contrast, the spatial dimension was considered in 95.5% (*N*: 86) of all investigated DSMM studies (≥1 site sampled) [3 studies did not collect any point-specific soil samples; in 2 studies it was not clear how soil samples were collected]. These numbers suggest that soil variability was investigated in much more detail than soil development through time. This may be explained by the long time periods over which processes in soils operate (e.g. weathering) and difficulties to model and test temporal predictions of soils.

3.4. Prediction of soil properties and classes

Out of 90 investigated studies 40.0% (*N*: 36) focused on predictions of base soil properties such as texture, bulk density, structure, 31.1% (*N*: 28) on soil carbon/global warming, 24.4% (*N*: 22) on eutrophication/environmental quality assessment, 16.6% (*N*: 15) on hydrologic properties (such as soil moisture, saturated hydraulic conductivity, or soil water content), 8.9% (*N*: 8) on soil degradation (salinity, acidity, and erosion), and 15.6% (*N*: 14) on mapping of soil taxonomic/ecological classes. In particular, soil properties that focused on mapping of soil organic carbon (SOC) and soil organic matter (SOM) were prominently represented. These studies aim to provide supporting information to address issues related to carbon mitigation and global climate warming assessing soil carbon stored across a soil-landscape (e.g. Grimm et al., 2008, Zhang et al., 2008a). Soil carbon inventories are useful to assess the amount of soil carbon stored across a region for a given time, but fall short to describe biogeochemical processes, carbon fluxes and transformations, and carbon sequestration rates. Only one study by Bauer et al. (2008) investigated CO_2 fluxes using a mechanistic simulation model applied to one site. Carbon fractions (labile, recalcitrant, and total carbon) and mineralizable carbon were predicted from spectral data across a large subtropical watershed by Vasques et al. (2009). Such studies aim to link site-specific information on C pools and controlling processes to the landscape scale.

Soil degradation and environmental concerns were addressed in numerous DSMM studies focused on nutrient enrichment such as total phosphorus (TP) and total nitrogen (TN) (Grunwald et al., 2008; Grunwald and Reddy, 2008; Rivero et al., 2007) or heavy metal distribution patterns (Wu et al., 2007, 2008) in terrestrial and aquatic ecosystems. Salinity, soil acidity, and erosion were modeled in various studies (compare Eldeiry and Garcia (2008), Li et al. (2008), Reinds et al. (2008), and Sommer et al. (2008)) focused on soil degradation and quality. These environmental-centered DSMM research studies respond to critical societal needs including environmental quality assessment, soil degradation, soil quality, and health as outlined in Hartemink (2006).

Out of 90 reviewed studies 15.6% (*N*: 14) predicted soil taxonomic, drainage classes, soil quality classes or ecological sites types, whereas all others predicted soil properties measured on continuous scale. These numbers are lower when compared to McBratney et al. (2003) who found that 70% reviewed studies predicted soil properties, while 30% predicted soil classes. Soil taxonomic units have been the major focus of agencies responsible for developing, maintaining, and updating soil survey maps and databases. For example, in the U.S. the Natural Resource Conservation Service (NRCS) — United States Department of Agriculture (USDA) developed taxonomic oriented soil

Table 1
Summary of plot/field scale digital soil mapping and modeling papers (2007–2008) published in Geoderma and Soil Science Society of America journals.

References	SPred	Samp: *N*, SP, SL, *F*	Region	Spatial extent	Dist	Leg	DSMM	Cov	Know-Q Know-PE	Ass
Prediction of soil properties — empirical models										
Allison et al. (2007)	Soil microbial communities	SP—36 (?) SL—6 (100 cm)	US	7 plots	TA	N	PCA, SpT	–	Q-1 PE-3	M
Brenning et al. (2008)	ECa (global, pooled fields)	No soil prop. collected	E	1.2 km^2	–	–	H	–	Q-3 PE-1	M, V
Chen et al. (2008)	SOC	*N*—139 SL—1	US	10 fields	?	N	NN	O [RS]	Q-4 PE-1	M, V
Cockx et al. (2007)	Clay	SP—23 SL—4	E	0.02 km^2	TA	N	Fuz, IK	S [ECa]	Q-3 PE-3	M, V
De Gryze et al. (2008)	SM, TC, TN, TP	*N*—400 SL—1; and SP—25 (1–100 cm) SL—10	E	0.032 km^2	GR + R	N	A, OK	O, R	Q-1 PE-3	M
Douaik et al. (2007)	Salinity	*N*—13 to 20 SL—2 *F*—19 times (1994–2001)	E	0.025 km^2	GR pseudo	N	MLR, ST, TS	S [ECa]	Q-3 PE-1	M, V
Guerrero et al. (2007)	MTR	*N*—310 SL—1	E	5 sites	?	N	PLSR	S [NIR]	Q-2 PE-4	M, V
Jiang et al. (2007)	PAW, PAWc	*N*—47 SL—1 *F*—2 (10/2005; 3/2006)	US	2 fields: 0.028 and 0.013 km^2	?	N	MLR	S [ECa]	Q-2 PE-2	M
Kerry and Oliver (2007a)	Clay	*N*—205; 109; 70; 118 SL—1	E	4 fields: 0.11; 0.07; 0.44; 9,15 km^2	GR	N	OK; Var-MoM; Var-REML	O, S [AP, ECa]	Q-4 PE-1	M, CV
Kerry and Oliver (2007b)	Structure	*N*—205; 109; 294; 102 SL—2	E	4 fields: 0.11; 0.07; 0.44; 9,15 km^2	GR	N	IK	O [AP]	Q-2 PE-2	M
Lu et al. (2008)	24 soil properties	SP—200 SL—1	US	2 fields	T	N	PMF	S	Q-4 PE-3	M
Mabit et al. (2008)	BD, SOM, Tex	*N*—42	Ca	0.0216 km^2	T	Y (^{137}Cs)	GIS, MLR, OK	–	Q-1 PE-4	M
Marchant and Lark (2007)	K	*N*—434	E	0.84 km^2	GR	Y	Matv, Bay	–	Q-5 PE-1	M, V
Ouimet (2008)	Ca, Mg, K, Na, Ti, Zr, SOM, Tex	SP—3	Ca	3 forest sites	?	N	C, PMB	–	Q-5 PE-5	M
Sadeghi et al. (2008)	SS	*N*—162 *F*—6/2004 to 7/2006	SEA	0.048 km^2	M	N	MLR, MNR	–	Q-2 PE-2	M, V
Shukla et al. (2007)	Cphyf, Nphyf, Tex, WCA	*N*—45 SL—1 *F*—4	US	1 ref. site and 3 reclaimed sites	G	N	OK	–	Q-1 PE-4	M
Waiser et al. (2007)	Clay	SP—72 SL—m (105 cm)	US	6 fields in TX	?	N	PLSR	S [VNIR]	Q-2 PE-2	M, V
Winzeler et al. (2008)	ECEC, M3 K, PANK	*N*—384 SP—128 SL—3	US	0.036 km^2	T	N	PMix	S, R	Q-2 PE-5	M
Prediction of soil properties — mechanistic modeling approach										
Bauer et al. (2008)	CO_2 prod. and fluxes	SP—1 *F*—2	US	1 site	–	N	MM (6 C sim. models)	S, C, T	Q-5 PE-5	M, V
Bonilla et al. (2007)	Erosion	*N*—12 *F*—RR	US	Idealized field; erosion plots	?	N	MM	S, C, O, R, P, T	Q-5 PE-5	M, V
Bricklemyer et al. (2007)	SOC	SP—30 SL—3	US	5 fields	R	Y	MM, MLR (CENTURY)	S, (SSURGO, STATSGO), T	Q-4 PE-4	M, V
Causarano et al. (2007)	SOC, corn yield	*N*—106 SL—1 *F*—3 (3 yrs)	US	0.04 km^2	T	N	MM (EPIC)	S, O, R, T	Q-5 PE-4	M, V
Finke and Hutson (2008)	40 soil properties	SP—1 SL—2 *F*—thousands of years	E	2 sites	?	N	MM	S, C, O, R, P, T	Q-5 PE-5	M, V
Jasques et al. (2008)	Water flux, Ca, Cd, Cl, Na, K, Zn	SP—1 *F*—30 yrs (sim. runs)	E	1 Podzol	?	N	MM	S, C, O, R, P, T	Q-5 PE-5	M
Kazemi et al. (2008)	Br, Atrz	SP—20 (Atrz) SP—49 (Br) SL—5 *F*—5	US	0.001 km^2	GR	N	OK, MM	S, T	Q-2 PE-3	M, CV
Oorts et al. (2007)	C and N flux	SP—1 SL—3 *F*—33 yrs	E	1 site	–	N	MM	S, C, O, T	Q-5 PE-5	M, V (?)
Reinds et al. (2008)	Soil acidification	SP—182	E	Forest sites in Europe	M	Y	Bay, MM	S, T	Q-5 PE-3	M, V

Table 1 (*continued*)

References	SPred	Samp: *N*, SP, SL, *F*	Region	Spatial extent	Dist	Leg	DSMM	Cov	Know-Q / Know-PE	Ass
Sommer et al. (2008)	Erosion/deposition	SL—m *F*—10 yr time steps (long period)	E	2 sites	?	N	MM	S, C, O, R, P, T	Q-5 PE-5	M
Prediction of soil classes										
Liu et al. (2008)	Drainage class	SP—144	Ca	~1 km^2			CA, MLC, DA	S, O, R [ECa, RS (HR, SAR]	Q-3 PE-4	M, CV
Park et al. (2007)	Soil types	*N*—120 and *N*—600*	E	0.12 km^2	R	Y	GCMC, SIS	S [soil map]	Q-5 PE-1	M, V
Reynolds et al. (2008)	SPQ	SP—10	Ca	Brookston clay loam soil	?	N	SR	S	Q-1 PE-4	M
Vitharana et al. (2008)	SMclass	SP—110 SL—2	E	0.08 km^2	?	N	Fuz, PCA, OK	S, R [ECa, LIDAR]	Q-3 PE-5	M

databases at various scales including (i) Land Resource Region (LRR) 1:7,500,000, (ii) Major Land Resource Area (MLRA) 1:3,500,000, (iii) Land Resource Unit (LRU) 1:1,000,000, (iv) State Soil Geographic Database (STATSGO) 1:250,000, (v) Soil Survey Geographic Database (SSURGO — Soil Data Mart) 1:24,000, (v) Official Soil Series Descriptions at site-specific scale, and (vi) General Description for Soil Characterization Data (database of the Soil Survey Laboratory, National Soil Survey Center, which currently contains analytical data for more than 20,000 pedons of U.S. soils and about 1100 pedons from other countries). Similar taxonomic oriented databases and maps have been developed in numerous countries around the world. Most of them have been developed using traditional field mapping supported by aerial photographs that rely heavily on tacit knowledge of soil surveyors and mapping of genetic features. These polygon-based soil maps are now used in numerous DSMM studies to provide preliminary base data. A total of 14.4% (*N*: 13) of investigated DSMM studies derived their soil data from existing soil maps/soil databases, while 3.3% (*N*: 3) used them to compare, test and/or validate their new DSMM models.

Out of 90 reviewed studies 54.4% (*N*: 49) predicted only one specific soil property/class, whereas 45.6% (*N*: 41) predicted or modeled two or more soil properties/classes. For example, Lu et al. (2008) predicted 24 different soil properties across 2 fields, Finke and Hutson (2008) modeled 40 different soil properties, and Zhang et al. (2008b) predicted 45 geochemical properties over a large landscape (71,000 km^2).

Global models covering a wide range of soil characteristics are required to provide robust predictions elsewhere. For example, Vasques et al. (2008) developed SOC spectral-based prediction models covering SOC from 0.16 to 270 g kg^{-1} and Brown et al. (2006) used a global soil set ranging from ~0.0 to 537 g kg^{-1}. In contrast, other spectral-based DSMM applications showed a much more limited range in SOC values of 8 to 15.7 g kg^{-1} (Hill and Schütt, 2000) or 1 to 12.8 g kg^{-1} (Kooistra et al., 2003) limiting the transferability of these soil prediction models to other regions.

3.5. Factors used to predict soil properties and classes

Numerous DSMM studies adopted the SCORPAN framework as the underlying conceptual model to predict soil properties and classes. To predict a soil property or class the S factor was used in 84.0%, the C factor 10.7%, the O factor 38.7%, the R factor 29.3%, and P factor 10.7% out of all investigated DSMM studies that documented the explicit use of SCORP factors as covariates (*N*: 75) (Case A, Table 4). The wide geographic extent of these studies provided a good representation of different soil formation history spanning across different environmental factor combinations (e.g. flat to steep terrain, aquatic and terrestrial systems, arctic to tropical climate). Note that reviewed studies included site-specific mechanistic simulation models in which one or few sites were used to predict soil properties into the future. For some of these mechanistic studies the site description was limited to few factors. For instance, Bauer et al. (2008) provided only S and C factors; Bricklemyer et al. (2007), Kazemi et al. (2008), and Reinds et al. (2008) the S factor; Causarano et al. (2007) S, O, and R factors; and Oorts et al. (2007) S, C and O factors in their DSMo studies. Furthermore some of the reviewed DSMM studies used a pedo-transfer function approach either based on lab-based spectral data (e.g. Wu et al., 2007; Awiti et al., 2008; Vasques et al., 2008) or neural network models (e.g. Lamorski et al., 2008) to predict soil properties without considering SCORPAN factors as covariates. This may explain why percentages of SCORP factors used to predict soil properties or classes differ from other review studies. For example, McBratney et al. (2003) found that the key predictor factors were R (80% of studies) followed by S (35%), O and P (both 25%), *N* (20%), and C (5%) in a comprehensive review. Case B (Table 4) provides percent contributions of SCORP factors to predict soil properties/classes excluding those reviewed papers that used mechanistic models to predict soil properties/classes; and Case C (Table 4) that excluded both mechanistic modeling as well as pedo-transfer function studies. The contribution of S was 50.7%, C 6.0%, O 34.3%, R 23.9%, and P 6.0% (Case C).

Soil prediction models were implemented with different quantitative methods and different combination sets of SCORPAN factors, perhaps sometimes restricted by the lack in environmental input data, which further confound findings. Some DSMM studies only incorporated few SCORPAN factors in their models (e.g. S and O factors were utilized by West et al. (2008); and S and R factors by Santra and Das (2008)), while other studies (e.g. Jasques et al., 2008; Grinand et al., 2008) incorporated a more comprehensive set of SCORPAN factors. The S factor was often characterized by soil taxonomic data derived from soil surveys (e.g. Bockheim and McLeod, 2008; Li, 2007) or by intrinsic, site-specific soil physico-chemical mapping using apparent electrical conductivity or cone penetrometers (e.g. Jiang et al., 2007; Weller et al., 2007). Whereas R and O factors can be easily derived using high-resolution remote sensing images and digital elevation models (compare Grimm et al. (2008), Liu et al. (2008), and Rivero et al. (2007)) representing fine-scale spatial variability, the P factor shows longer spatial autocorrelation ranges and is often only available at coarser map scales. The dominance of the R factor in soil predictive models as outlined by Thompson et al. (2001), McBratney et al. (2003), and Bishop and Minasny (2006) could not be confirmed in this review study. Rivero et al. (2007) pointed out that the R factor has minor predictive power in areas with relatively flat terrain (e.g. southeastern U.S.). In these cases other factors show more predictive capabilities to infer on soil properties. In particular factors that can be characterized using remote or soil sensors to infer on O and S factors, respectively, were prominently presented in investigated DSMM studies. Furthermore, multiple of the reviewed DSMM studies used a rather simplistic approach to predict soils using GIS-based aggregation/grouping methods or ordinary kriging of field observations that did not explicitly incorporate any SCORPAN factors as covariates.

3.6. Incorporation of sensors into soil predictive models

Out of 90 DSMM studies 38.9% (*N*: 35) utilized soil or remote sensors, out of which 23.3% (*N*: 21) used soil sensors to complement analytical soil data which are more costly and labor-intensive to derive. In 16.7% (*N*: 15) of the studies VNIR, mid-infrared, near-infrared and Fourier-transform spectroscopy, in 8.9% (*N*: 8) apparent electrical conductivity (ECa), in 1.1% (*N*: 1) magnetometry, and in 1.1% (*N*: 1) cone penetrometers were incorporated. Out of 15 studies that utilized spectroscopy, 9 were conducted across larger landscapes, 4 at undefined spatial scale, and 2 at plot/field scales. In contrast, almost all ECa applications (*N*: 6) were conducted at plot or field scale, one within a small farm (Weller et al., 2007), and one at landscape scale (Lobell et al., 2007). Remote sensing was used in 16.7% (*N*: 15) of DSMM studies due to their superior capabilities to provide environmental information at high spatial resolution over large terrain. Out of the 15 studies that used remote sensing, 10 were conducted at landscape scale (Table 2), 4 at plot/field scale (Table 1), and 1 miscellaneous (Table 3).

3.7. Legacy soil data

The use of legacy data in DSMM studies has become a commonplace with 36.7% (*N*: 33) of investigated DSMM using such data. For the development of pedo-transfer functions on soil water properties large databases have been formed such as the Unsaturated Soil Hydraulic Property Database (UNSODA) or the database of hydraulic properties of European Soils (HYPRES) that span across the soil variability domain capturing different soil types in various geographic settings. For example HYPRES was used by Baker (2008) and Baker and Ellison (2008) using 2764 samples and by Zacharias and Wessolek (2007) using 823 samples to predict soil moisture retention.

3.8. Methods used to predict soil properties and classes

While some DSMM research studies are driven to improve our understanding of processes, others are need-driven, aiming to solve land resource or environmental problems resulting in an enormous diversity in terms of methodology, approach, intensity of sampling, and sampling design. Different pedometric methods are often compared with each other to predict soil properties across landscapes driven by the goal to identify the prediction method that performs best in a given soil-landscape setting. For example, Rivero et al. (2007) compared three different univariate and multivariate geostatistical methods incorporating remote sensing and environmental landscape properties to assess predictive capabilities for soil TP in a subtropical wetland. Vasques et al. (2008) compared five different methods to predict SOC using VNIR spectra and Minasny and McBratney (2007b) compared four different methods to predict clay, cotton yield, pH and Zn. These studies and others identified which factor combinations and method performed best in a given landscape setting. Out of 90 reviewed studies 40% (*N*: 36) presented soil prediction results derived from only one method, whereas 60% (*N*: 54) used two or more quantitative methods to predict or model soil properties/classes.

The selection of various quantitative methods ranging from statistical, geostatistical, and hybrid methods to mechanistic simulation models to predict soil properties or soil development through time was remarkable. The most popular method to predict soils at 41.1% (*N*: 37) were regressions followed by classification/discrimination methods (32.2%, *N*: 29), univariate kriging (18.9%, *N*: 17), and tree-based models (13.3%, *N*: 12) (Table 5). The latter group characterizes fitting methods that optimize relationships between inputs (e.g. environmental factors) and outputs (e.g. soil property or class) and these have become very popular because of their ease of use and excellent prediction capabilities (Breiman et al., 1984; Breiman, 1996; Hand et al., 2001). On the other hand tree-based optimization methods have been criticized for overfitting, lack of robustness if transferred to other geographic regions, and poor performance on small datasets. Mechanistic simulation models accounted for 14.4% (*N*: 13) of all investigated DSMM studies. Surprisingly, multivariate kriging methods that account for both deterministic trend component and spatial variation of soil properties (Grunwald, 2008) comprised only 7.8% (*N*: 7) of DSMM studies. Other methods such as GIS-based explicit modeling, neural networks, fuzzy logic based models, and stochastic simulations were less frequently used to predict soils (Table 5). Knowledge-based DSMM was rare (e.g. Bockheim and McLeod, 2008; MacMillan et al., 2007) when compared to stochastic or deterministic methods to predict soils.

3.9. Calibration and validation of soil prediction models

Out of 90 investigated studies 21.1% (*N*: 19) used cross-validation, 46.7% (*N*: 42) used validation, and 35.6% (*N*: 32) did not use cross-validation, validation or any other performance test. Rigorous performance tests to evaluate soil predictions in various geographic soil-landscape settings are critical to minimize uncertainty in soil predictions at unsampled locations. DSMM studies that did not use cross-validation or validations are marked with “M” in Tables 1–3.

4. Final thoughts — discussion and recommendations

4.1. External and internal factors imparting control on soil properties and classes

Most DSMM studies focused on external drivers (controlling factors) such as climate, vegetation, and land use that modulate SOC, TP or other soil properties that were predicted (compare Rivero et al. (2007), Grimm et al. (2008), and Schulp and Veldkamp (2008)). Environmental factors such as land cover and terrain can be rapidly and cost-effectively mapped using remote sensing and digital elevation models covering large areas. But these models did not incorporate intrinsic soil properties such as mineralogy, microbial communities, structure, or soil aggregation that have shown strong correlations to soil formation at site-specific scales (Zibilske and Bradford, 2007; Zinn et al., 2007). To incorporate intrinsic, site-specific properties derived at fine spatial scales into coarser scale soil prediction models is more challenging due to analytical costs and labor required to collect these properties. However, to further improve existing prediction models focused on functional relationships between environmental factors and soil properties a crucial next step would require to incorporate biogeochemical process-knowledge into the modeling process (Grunwald et al., 2007).

4.2. Boundary conditions

The reviewed research-oriented DSMM studies are in contrast to operational surveys that require standardization to produce a soil survey product following a specific protocol at a predefined map scale. Given the tremendous variability in soil properties across landscapes we are challenged to adapt soil prediction models to specific landscapes and a unifying DSMM method that would fit all needs and landscapes seems out of reach. Thus, defining boundary conditions for soil prediction models is of paramount importance but is often not clear (compare Awiti et al. (2008), Lamorski et al. (2008), and Lilly et al. (2008)). The potential to employ existing DSMM models in other regions is constrained by the lack of definition of soil boundary conditions and by the failure to explicitly characterize the soil variability domain. This raises concern about the transferability of soil prediction models developed in one region to similar

Table 2
Summary of coarse scale digital soil mapping and modeling papers (2007–2008) published in Geoderma and Soil Science Society of America journals.

References	SPred	Samp: *N*, SP, SL, *F*	Region	Spatial extent	Dist	Leg	DSMM	Cov	Know-Q Know-PE	Ass
Prediction of soil properties focused on environmental quality and/or nutrient management and/or basic properties										
Awiti et al. (2008)	Ca, ECEC, K, Mg, pH, SC, TC, TN	*N*—582 SL—1	Af	~50 km^2	SR	N	DA, PCA, PLSR	S [NIR]	Q-4 PE-2	M, V
Borůvka et al. (2007)	Al, C, Ca, Mg, N, pH, S	SP—98 SL—2	E	196 km^2	iGR	N	PCK	S, O, R	Q-3 PE-3	M
Brown (2007)	Clay, 2:1 Clay, SOC	*N*—4184; *N*—418 SL—m	Af, E, US, SEA	World	?	Y	CT	S, R [VNIR]	Q-2 PE-2	M, CV
Chaudhari et al. (2008)	Tex	*N*—100	I	Large region in India	?	N	MLR	–	Q-1 PE-2	M
Corstanje et al. (2008)	Ca, TN, TP, TPI	*N*—1341 SL—1	US	8248 km^2	SR	N	MCS, USGRF, WLV	–	Q-5 PE-1	M
Eldeiry and Garcia (2008)	Salinity (EM-38)	*N*—256 SL—1 and 68 wells	US	500 km^2	?	N	LR, OK, RK	O [RS]	Q-3 PE-3	M, V
Grunwald et al. (2008)	BD, TP, TPi	SP—62 (1998) SL—2 (1998) SP—111 (2003) SL—2 (2003) *F*—2	US	432.8 km^2	GR + R; SR	N	Fuz, OK	–	Q-3 PE-5	M, CV
Grunwald and Reddy (2008)	TN, TP	SP—266 SL—1	US	48 km^2	GR + R	N	LNOK	–	Q-3 PE-3	M, CV
Jordanova et al. (2008)	11 heavy metals	*N*—20 SL—1	E	2800 km^2	GR	N	C, Fuz	S [Mag]	Q-1 PE-4	M
Lobell et al. (2007)	Soil salinity, wheat yield	*N*—122; SL—2 *N*—60 SL—2 *F*—6 yrs	SA	270 km^2	TA	N	MLR	S, O [RS, ECa]	Q-2 PE-4	M, V
Lufafa et al. (2008)	BD, SOC	45 plots × 8 locations SL—1	Af	44,000 km^2	?	N	OK, MLR	O	Q-1 PE-4	M, CV
Minasny and McBratney (2007b)	Clay, cotton yield, pH, Zn	*N*—155; *N*—399; *N*—341; *N*—248 *F*—SD	E; Au; (4 diff. areas)	10 km^2: ?; ?; 0.83 km^2	?; LH;	Y	Matv, REML-EBLUP, OK, RK	S, O, R [RS]	Q-5 PE-1	M, V
Rivero et al. (2007)	TP	SP—111 SL—2	US	432.8 km^2	SR	N	CK, OK, RK	O [D, RS, XY]	Q-4 PE-4	M, CV
Wu et al. (2007)	8 heavy metals	*N*—61 SL—1	Ch	~16 km^2	G	N	C, PLSR	S [VNIR]	Q-2 PE-4	M
Wu et al. (2008)	4 heavy metals	*N*—94 SL—1	Ch	10.5 km^2	SR	N	A, BK	S	Q-2 PE-3	M, CV
Zhang et al. (2008b)	45 geochemical properties	*N*—1310 SL—1	E	71,000 km^2	G	N	G	S, O, P	Q-1 PE-3	M
Prediction of soil carbon										
Awiti et al. (2008)	Ca, ECEC, K, Mg, pH, SC, TC, TN	*N*—582 SL—1	Af	~50 km^2	SR	N	DA, PCA, PLSR	S [NIR]	Q-4 PE-2	M, V
Batjes (2008)	SOC	SP—5760 *F*—SD (SOTER)	Af	2.4 mill km^2	?	Y	GIS, PTF	C	Q-3 PE-4	M
Borůvka et al. (2007)	Al, C, Ca, Mg, N, pH, S	SP—98 SL—2	E	196 km^2	iGR	N	PCK	S, O, R	Q-3 PE-3	M
Brown (2007)	Clay, 2:1 Clay, SOC	*N*—4184; N—418 SL—m	Af, E, US, SEA	World	?	Y	CT	S, R [VNIR]	Q-2 PE-2	M, CV
Genxu et al. (2008)	LFOC, SOC	SP—40 (?) SL—4	Ch	1.212×10^5 km^2	T	Y	GIS, G	S, O, R, [RS]	Q-2 PE-4	M
Gomez et al. (2008)	SOC	*N*—146 SL—1 *F*—2 (10 and 12/2006)	Au	7.6 km corrid.	T	N	PLSR	S [VNIR RS–HR]	Q-4 PE-2	M, CV
Grimm et al. (2008)	SOC	SP—165 SL—4	SA	~100 km^2	?	N	RF	S, R, P	Q-2 PE-2	M, CV
Lufafa et al. (2008)	BD, SOC	45 plots × 8 locations SL—1	Af	44,000 km^2	?	N	OK, MLR	O	Q-1 PE-4	M, CV

(continued on next page)

Table 2 (*continued*)

References	SPred	Samp: *N*, SP, SL, *F*	Region	Spatial extent	Dist	Leg	DSMM	Cov	Know-Q Know-PE	Ass
Meersman et al. (2008)	SOC	*N*—6900 *F*—SD (Aardewerk)	E	~30,510 km^2	?	Y	A, G, MLR	S, O	Q-3 PE-4	M
Nyssen et al. (2008)	SOC	*N*—25 SL—1	Af	637 km^2	SR	Y	C, MLR	O, R [RS]	Q-3 PE-3	M
Schulp and Veldkamp (2008)	SOM	*N*—4000 (?) SL—1 *F*—SD	D	60 km^2	?	Y	C, G, MLR	S, O	Q-1 PE-4	M
Vasques et al. (2008)	SOC	*N*—554 SL—4 *F*—9/2003–1/2005	US	3585 km^2	SR	N	CT, MLR, PLSR, PCR, RT	S [VNIR]	Q-4 PE-2	M, V
Vasques et al. (2009)	HC, HWSC, MC, RC, SOC	*N*—141 SL—1	US	3,585 km^2	SR	N	CT, PLSR, PCR, RT, MLR	S [VNIR]	Q-4 PE-3	M, V
West et al. (2008)	SOC	*N*—? SL—3 m depth *F*—SD (SSURGO, STATSGO)	US	Large Midwestern region (679 counties)	?	Y	GIS	S, O [RS]	Q-3 PE-4	M
Zhang et al. (2008a)	SOC	SP—798 *F*—SD	Ch	806,400 km^2	R	Y	GIS, PTP	R	Q-1 PE-4	M
Prediction of hydrological soil properties										
Baker (2008)	SWR	*N*—2764 *F*—SD (HYPRES)	E	Large region	?	Y	Class-PTF	S	Q-5 PE-1	M, CV
Baker and Ellison (2008)	SWR	*N*—2764 *F*—SD (HYPRES)	E	Large region	?	Y	NN, PTF	S	Q-5 PE-1	M, CV
Lamorski et al. (2008)	SMP, SWC	*N*—806 SP—290 *F*—SD	E	Large soil region (Poland)	?	Y	NN, PTF, SVM	S	Q-4 PE-1	M, V
Lilly et al. (2008)	Ks	*N*—502	E	Large region	?	Y	PTF, RT	S	Q-2 PE-2	M, V
Minasny and McBratney (2007a)	WRSP	*N*—630 *F*—SD (GRIZZLY; UNSODA; Australian SD)	Au, E	Large, but no information how large	?	Y	NN	S	Q-3 PE-3	M, V
Parasuraman et al. (2008)	Ks	*N*—314 *F*—SD (UNSODA)	World	World	SR	Y	GP, NN	S	Q-5 PE-2	M, V
Santra and Das (2008)	Ks	*N*—646 *F*—SD	I	Benchmark soils E India	R	Y	MLR	S, R	Q-1 PE-3	M, V
Tranter et al. (2008)	MR, SMR	*N*—521	Au	Large region in New South Wales and Victoria	?	Y	NN, PLSR, PTF	S [MIR]	Q-4 PE-1	M, V
Zacharias and Wessolek (2007)	SMR	*N*—823 *F*—SD (IGBP-DIS; UNSODA)	World	–	?	Y	MLR, MNLR, PTF	S	Q-1 PE-1	M, V
Prediction of soil classes										
Bockheim and McLeod (2008)	Soil types	SP—550 SL—m (1–100 cm)	US	6,692 km^2	?	Y	GIS, Smap, KB	S, R	Q-1 PE-5	M
Grinand et al. (2008)	Soil units	SP—1800	E	900 km^2	?	Y	ClT, CT	S, O, R, P	Q-3 PE-3	M, V
Hengl et al. (2007)	Tex, soil group	SP—4250	ME	1.6 mill. km^2	?	Y	G, MLR, PCA, RK, SC	S, O, R,	Q-3 PE-1	M, V
Li and Zhang (2007)	Soil types	*N*—646; *N*—179; *N*—50*	US	9 km^2	R	Y	IK, MCRF, SIS, TG	S (SSURGO)	Q-4 PE-1	M, V
Li (2007)	Soil types	*N*—12936 and *N*—136*	US	92.2 km^2	?	Y	MCRF, TG	S (SSURGO)	Q-4 PE-1	M, V
MacMillan et al. (2007)	Eco	*N*—115 (field traverses)	Ca	82,000 km^2	R	N	GIS, KB, fuzzy	S, C, O, R, P	Q-2 PE-5	M, V
Minasny and McBratney (2007c)	Soil class	SP—718	Au	Large regions	?	Y	ClT, DT	S, O, R [RS]	Q-3 PE-3	M, CV
Nield et al. (2007)	Gypsic and natric soil areas (salt-affected soils)	*N*—93	US	220 km^2	TA	N	SR	O [RS]	Q-3 PE-5	M, V

Table 3
Summary of remaining reviewed digital soil mapping and modeling papers (2007–2008) published in Geoderma and Soil Science Society of America journals.

References	SPred	Samp: *N*, SP, SL, *F*	Region	Spatial extent	Dist	Leg	DSMM	Cov	Know-Q / Know-PE	Ass
Prediction of soil properties										
Bartholomeus et al. (2008)	SOC	*N*—40	Soils diverse C range	?	?	Y	PLSR	–	Q-3 PE-3	M, CV
Ben-Dor et al. (2008)	Fed, SC, SM, SOM, SSA	*N*—160 SL—m	Israel	?	T	N	PLSR	S [VNIR in-situ]	Q-2 PE-2	M, V
Carrara et al. (2007)	PR	*N*—6714 SL—m	E	?	T	N	SS	S [CP-R; 3D]	Q-3 PE-2	M
Farifteh et al. (2008)	EC, Mg, Ca, Na, K, pH	*N*—147	E	lab	?	N	C, MLR	–	Q-2 PE-4	M, CV
Janik et al. (2007)	SWR	N—96_c; N—916_v	Au	New South Wales, Victoria and SE Australia	?	Y	PCA, PLSR	S [MIR]	Q-2 PE-2	M, CV, V
Li et al. (2008)	Erosion	SP—19	US	4 hypoth. landscapes; and 0.027 km^2	–	N	MM	S, O, R, P, T	Q-5 PE-5	M, V
Rasmussen and Tabor (2007)	EEMT	SP—21	US	Across US (diff. bedrock material; scale unclear)	T	N	A, MLR, MNLR	S, C	Q-3 PE-5	M, V
Stevens et al. (2008)	SOC	*N*—117 (plots)	E	?	?	N	PLSR	S [VNIR - lab, ground, RS]	Q-2 PE-2	M, V
Weller et al. (2007)	BD, SOC, Tex	*N*—46	E	Small farm	?	N	MLR, NeN	S [ECa]	Q-3 PE-3	M
Wills et al. (2007)	SOC	SP—125 SL—m	US	parts of Iowa	GR	N	MLR	S [soil color]	Q-1 PE-3	M
Wöhling et al. (2008)	Ks	SP—100 SL—5	Au	Small catchment	?	N	invM, MM	S	Q-5 PE-5	M, V
Yoo and Mudd (2008)	Geochem. soil formation	–	–	–	–	N	MM	–	Q-5 PE-5	M
Prediction of soil classes										
Connolly et al. (2007)	Peat covg and type	*N*—300	E	Sites in Ireland	TA	Y	DT, GIS	S, C, O	Q-3 PE-2	M, V
Du et al. (2008)	Soil types	*N*—166 SL—1	ME	?	?	Y	DA, NN, PCA	S [FTIR-PAS]	Q-4 PE-2	M, V

Description of properties in Table 1 to 3:
SPred — predicted soil attributes/classes: atrazine (Atrz), bromide (Br), bulk density (BD), cadmium (Cd), carbon physically fractionated into size classes (Cphyf), ecological site types (Eco), effective cation exchange capacity (ECEC), free iron oxides (Fed), hot water soluble carbon (HWSC), hydrolysable carbon (HC), light fraction organic carbon (LFOC), loss on ignition (LOI), maximum temperatures reached in soils (MTR), Mehlich-3 K (M3K), mineralizable carbon (MC), moisture retention (MR), nitrogen physically fractionated into size classes (Nphyf), penetration resistance (PR), plant available non-exchangeable K (PANK), plant available water (PAW), plant available water capacity (PAWc), rate of effective energy and mass transfer (EEMT), recalcitrant carbon (RC), saturated hydraulic conductivity (Ks), specific surface area (SSA), soil carbonates (SC), soil condition ("soil quality"), soil management class (SMclass), soil matric potential (SMP), soil moisture (SM), soil moisture retention (SMR), soil organic carbon (SOC), soil organic matter (SOM), soil water content (SWC), soil water retention (SWR), soil physical quality (SPQ), suspended sediment (SS), texture (Tex), total aluminum (TAl), total carbon (TC), total calcium (TCa), total inorganic phosphorus (TPi), total iron (TFe), total magnesium (TMg), total nitrogen (TN), total phosphorus (TP), water retention shape parameter (WRSP), and water stability of aggregates (WCA).
Samp — sampling: total number of soil samples (*N*) [*samples = pixels randomly resampled from available soil map], soil profiles/number of observation sites (SP), number of soil layers/horizons (SL) [m: multiple], and frequency/time period (*F*) [if *F* is not listed the frequency was "1"]; SD — soil database was used where time period of samples included was unclear; Aardewerk — Belgian National Soil Survey, Australian soil database, Grenoble soil catalog (GRIZZLY), database of hydraulic properties of European Soils (HYPRES), international soil dataset for pedo-transfer functions (IGBP-DIS); Soil and Topographic Database (SOTER); Soil Survey Geographic Database (SSURGO); State Soil Geographic Database (STATSGO); Unsaturated Soil Hydraulic Database (UNSODA); rainfall-runoff events (RR).
Region — soil geographic region: Africa (Af), Australia (Au) [and New Zealand], Canada (Ca), China (Ch), India (I), Europe (E), Middle East (ME), South America (SA), South-east Asia (SEA), United States (US).
Spatial extent: size of study area (in km^2) or description of region.
Dist — distribution of soil observations (sampling design): Grid (GR), irregular grid (iGR), monitoring network (M), Latin-hypercube (LH), random (R), stratified-random (SR), transect (T), targeted sampling (i.e., specific sites were strategically selected) (TA).
Leg — incorporation of legacy/historic soil data: No (N), yes (Y).
DSMM — digital soil mapping/modeling techniques: Analysis of Variance (A), Bayesian approach (Bay), block kriging (BK), canonical analysis (CA), co-kriging (CK), classification trees (ClT), committee trees (CT) [or boosted regression trees], correlations (C), decision tree (DT), fuzzy c-means clustering (Fuz), fuzzy rule-based system (fuzzy), discriminant analysis (DA), generalized coupled Markov chain (GCMC), genetic programming (GP), Geographic Information System (GIS), grouping/clustering of soil properties (G), homogenization of ECa on multiple fields (H), indicator kriging (IK), inverse modeling (invM), knowledge-based soil mapping (KB), linear regression (LR), log-normal ordinary kriging (LNOK), Markov Chain Random Field (MCRF), Matérn variogram (Matv), maximum likelihood classifier (MLC), mechanistic model (MM), Monte Carlo simulation (MCS), multi-Gaussian kriging (MGK), multinomial logistic regression or multivariate linear regression (MLR), multivariate non-linear regression (MNR), nearest neighbor (NeN), neural network (NN), ordinary kriging (OK), partial least squares regression (PLSR), pedological mass balance (PMB), pedo-transfer function (PTF), positive matrix factorization (PMF), principal component analysis (PCA), principal component kriging (PCK), principal component regression (PCR), procmix linear model (PMix), random forest (RF), regression kriging (RK), regression trees (RT), residual maximum likelihood-empirical best linear unbiased predictor (REML-EBLUP), sequential Gaussian simulation (SGS), shift test (temporal mean and spatial) (ST), sequential indicator simulation (SIS), soil map (Smap), spatial trend (SpT), spectral ratios (SR), stepwise discriminant analysis (SDA), stochastic simulation (SS), supervised classification (SC), support vector machines (SVM), structured regression (SR), temporal stability test (TS), transiograms (TG), unconditioned simulation of a Gaussian random field (USGRF), variogram-method of moments (Var-MoM); variogram-residual maximum likelihood (Var-REML), and window-based local estimation of variograms (WLV).
Cov — covariates: soils (S), climate (C), organisms (O), relief (R), parent material (P), distance measures (D), and easting and northing (XY) coordinates, and time (T). Types of data collection: aerial photographs (AP), apparent electrical conductivity (ECa), Cesium (Cs), cone penetrometer-resistance (CP-R), Fourier-transform infrared photoacoustic spectroscopy (FTIR-PAS), hyperspectral remote sensing data (HR), light detection and ranging-airborne laser-based scanning (LIDAR), magnetometry (Mag), mid-infrared spectroscopy (MIR), near-infrared spectroscopy (NIR), remote sensing (RS), synthetic aperture radar (SAR), visible/near-infrared spectroscopy (VNIR).
Know — knowledge: Know-Q with 1: very basic to 5: complex knowledge on quantitative methods/modeling; Know-PE with 1: very basic to 5: comprehensive pedological and environmental interpretation of findings.
Ass — DSMM assessment mode with Model development only (M), cross-validation (CV) and validation (V).

and/or other regions. In particular, empirical DSMM models developed at a local scale that capture only some of the soil variability of a larger landscape are likely to limit the transferability of such models to other regions. Upscaling from plot/field to landscape scales effectively increases both grain and extent and the heterogeneity in soil and environmental factors (McBratney, 1998). An increase in scale means an increase in parameter variance, and this may cause problems by interacting with a non-linearity in the process or model (Addiscott et al., 1995). Models developed in one region that are transferred to other regions face similar issues.

In the future, pooling of different soil datasets will offer opportunities to enhance DSMM predictive capabilities to avoid extrapolations beyond the model development range. The Global Soil Spectroscopy Working Group has begun to pool spectral data collected in different regions across the globe covering 27 different countries with the aim of developing global spectral-based soil prediction models (Viscarra Rossel et al., 2008). Research gaps exist that test the robustness of soil predictions over different soil regions to assess their transferability. Few studies compared local (i.e., region-specific) and global soil prediction models. For example, Brown (2007) compared local and global soil spectral libraries using VNIR spectroscopy and X-ray diffraction and assessed their performance to predict clay and SOC. More such studies are needed to assess the performance of local (customized) soil prediction models compared to global ones developed at larger spatial scales. Scaling behavior and performance of soil prediction models across different spatial and temporal scales are still poorly understood. To unravel environmental correlations across pedo-geographic, climatic, topographic, geologic, and biologic boundaries will be critical in future global DSMM projects.

Collating soil data into open-access online databases, such as HYPRES or the international soil dataset for pedo-transfer functions (IGBP-DIS), has potential to develop global DSMM applications in the future encouraging reusability of the same datasets. However, caution is needed because historic data may be poorly georeferenced, mismatches among data collection and analytical methods may cause inconsistencies, and often uncertainty measures have not been incorporated in historic soil datasets.

Table 4
Percentages of soil (S), climate (C), organism (O), relief (R), and parent material (P) factors used in reviewed papers.

Cases	Review methods	Factors				
		S	C	O	R	P
Case A	Reviewed studies including predictions of soil properties/classes at spatially distributed unsampled locations and into the future using mechanistic models applied at one or few sites and pedo-transfer functions[a] (*N*: 75)[b]	84.0	10.7	38.7	29.3	10.7
Case B	Reviewed studies including predictions of soil properties/classes at spatially distributed unsampled locations and pedo-transfer functions[a], but excluding mechanistic models applied at one or few sites that make prediction into the future (*N*: 67)[b]	84.1	6.3	36.5	27.0	6.3
Case C	Reviewed studies — only spatially distributed predictions of soil properties/classes, but excluding mechanistic models applied at one or few sites that make predictions into the future and excluding pedo-transfer functions[a] (*N*: 44)[b]	50.7	6.0	34.3	23.9	6.0

[a] Pedo-transfer functions (e.g. neural network models, visible/near-infrared spectral prediction models).
[b] Fifteen studies were excluded because either no data were provided or no SCORPAN factors were explicitly incorporated into prediction models as covariates.

Table 5
Overview of quantitative methods used in 90 investigated digital soil mapping/modeling studies (note: some studies used more than one method).

Quantitative methods	Counts of methods (*n*)	Percent of 90 investigated DSMM studies	Specific methods and counts
Regression variants	37	41.1	LR = 1, MLR = 19, MNR = 2, PCR = 2, PLSR, 12, SR = 1
Classification/ discrimination	29	32.2	A = 4, CA = 1, C = 6, DA = 3, G = 5, MLC = 1, NeN = 1, PCA = 6, PMix = 1, SC = 1
Univariate kriging	17	18.9	BK = 1, IK = 3, LNOK = 1, MGK = 1, OK = 11
Tree-based models	12	13.3	CIT = 2, CT = 4, DT = 2, RF = 1, RT = 3
Mechanistic models	12	13.3	Quasi and complete process-based models
Geographic Information Systems (GIS)[a]	8	8.9	–
Multivariate kriging	7	7.8	CK = 1, PCK = 1, RK = 5
Neural networks	7	7.8	–
Fuzzy logic	5	5.6	Fuz (c-means) = 4, fuzzy (rule-based) = 1
Stochastic simulations	4	4.4	SIS = 2, SS = 1, UCGRF = 1

For acronym explanations, see Table 3 legend.
[a] Only those studies were counted in which GIS was explicitly used to implement a DSMM study. Numerous other papers used GIS to prepare environmental and soil input datasets.

4.3. Spatial and temporal scales

Review results showed that soil predictions or classes across space were more often implemented than predictions through time. There are needs to implement soil monitoring schemes to better address temporal fluctuations of soil properties due to short spatial and temporal autocorrelation ranges of dynamic soil properties (e.g. labile soil carbon, nitrate-nitrogen). To address global climate warming or evaluate soil ecosystem service it is pertinent to articulate how soil properties change with time (e.g. SOC sequestration rate; rate between labile and total carbon through time) and across a region. The dichotomy between either modeling soils in space or in time is still prevalent in many DSMM studies. Mechanistic-oriented models that simulate soil processes through time are focused on select sites, whereas empirical DSMM studies tend to adopt a spatially-explicit approach to map soil properties at one specific time over a large region. Soil sampling campaigns that fuse data to develop space–time models to predict soil properties and classes are critical to improve future DSMM applications.

4.4. Knowledge rankings

DSMM studies that showed a high Know-Q rank (≥3) used complex pedo-transfer functions based on neural networks and/or genetic programming (Parasuraman et al., 2008), Monte Carlo Simulation (MCS) and Unconditioned Simulation of a Gaussian Random Field (USGRF) and window-based local estimation of variograms (WLV) (Corstanje et al., 2008), transiograms (TG) and Markov Chain Random Field (MCRF) (Li, 2007), or mechanistic simulation modeling of soil formation (Bricklemyer et al., 2007; Finke and Hutson, 2008) which required profound mathematical expertise. These DSMM papers often showed negative correlations between Know-Q and Know-PE ranking, except for mechanistic modeling studies (Know-Q ≥ 4 and Know-PE ≥ 4) that required linking of process-based knowledge on soil formation into algorithms (Fig. 1). Other DSMM studies that demonstrated mastery of quantitative expertise often lacked pedological and environmental interpretations of results (i.e., showed low Know-PE scores < 3). Although the methodology sections of such papers were detailed and assessment of prediction performances was thorough hardly any pedological context

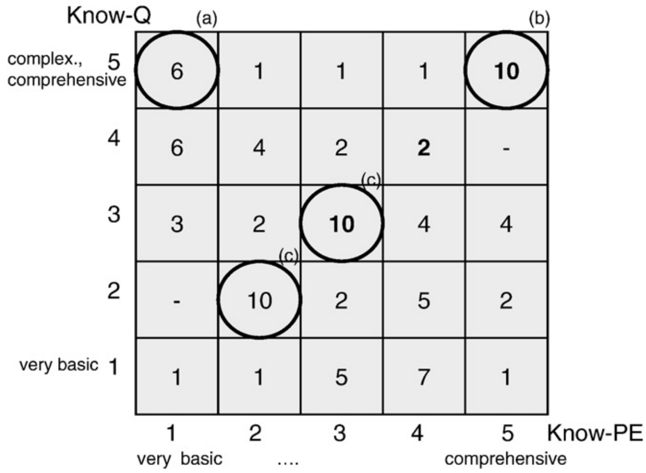

Fig. 1. Counts of DSM studies are shown in the matrix contrasting quantitative knowledge (Know-Q) and pedological/environmental knowledge (Know-PE) required to implement. (a) Mathematical oriented studies; (b) pedological simulation studies; and (c) regression/optimization methods.

was provided (e.g. Marchant and Lark, 2007; Chen et al., 2008; Corstanje et al., 2008; Gomez et al., 2008; Tranter et al., 2008), which may limit their widespread adoption by practitioners. Minasny and McBratney (2007b) suggested that for practical applications digital soil prediction methods, such as regression kriging, may be mathematically biased, however, they appear robust to predict soil properties in various soil regions. Both authors conclude that improvement in the prediction of soil properties does not rely on more sophisticated quantitative methods, but rather on gathering more useful and higher quality data. In Fig. 1 in the center of the matrix several DSMM applications that balance quantitative and pedological expertise are shown. These balanced approaches have the potential to be readily integrated into operational soil survey programs to generate new soil prediction maps at higher resolutions than historic products and provide associated prediction errors. Other studies that used less complex DSMM, for example, GIS-based pedo-transfer function (Batjes, 2008; Zhang et al., 2008a); GIS-based soil mapping (Bockheim and McLeod, 2008); partial least square regression modeling (Guerrero et al., 2007; Wu et al., 2008); or ProcMix linear models (Winzeler et al., 2008) have value to provide soil descriptions across larger landscapes, but are less rigorous in terms of their quantitative approach. A limitation of the latter suite of DSMM approaches is the lack in assessment of uncertainty in predicting soils.

4.5. Costs and prediction quality of digital soil mapping and modeling studies

Most DSMM studies published in Geoderma and Soil Science Society of America Journal are research-oriented and do not provide any information on total costs or on costs per ha mapped. However, analytical, sampling, and computational cost assessment is critical for operational soil survey programs. Thus, closer linkages between research and applied DSMM could be achieved by providing information about sampling, analytical, and computational costs. Mechanistic modeling studies are often associated with high costs due to the large amount of input data required to develop and test models in addition to high computational costs. Sensor-based methods aim to reduce costs to predict soil properties and classes, however, such studies often require costly instrumentation and are computationally expensive.

An assessment of soil prediction quality and uncertainty was not feasible in this review study. The variety of measures used, such as coefficient of determination, mean prediction error, root mean square prediction error, bias, residual prediction deviation, etc., was too large to put them onto a common scale. A standardization to report a minimum set of error measures, extent, and resolution would enhance comparability of different DSMM studies in the future.

5. Conclusion and outlook

A holistic understanding of soil patterns and processes is critical because soils are the matrix which all transport and transformation processes must pass. Scaling of soil properties and processes to an appropriate scale to address societal needs and evaluate ecosystem services will be critical in the future. DSMM can contribute to enhancing the decision making process to resolve local, national, and global environmental problems of significance by providing accurate soil predictions that capture the variability of soil-landscapes. Modeling of soils in 3D space and through time will require synergistic strategies to converge environmental landscape data and denser soil datasets. There is a need for more sophisticated technologies to measure soil properties and processes at fine resolution and with high accuracy. There are numerous quantitative models rooted in SCORPAN that predict soil properties with high accuracy in select geographic regions, but they lack consistency in terms of environmental input data, soil properties, quantitative methods, and evaluation strategies ("Wild West of DSMM"). Although standardization of DSMM into a generic, global model that predicts soils is appealing the variability of soil and environmental landscape conditions is daunting. Adaptations of DSMM to different local or regional conditions and knowledge on upscaling and downscaling of DSMM will close existing research gaps.

Notes

The ranking of DSMM studies into different categories was done to the best knowledge and understanding of the author. High, average or low scores were intended to classify a given study relative to other DSMM studies and there was no intention to reflect on the quality of a specific paper.

Acknowledgements

I would like to thank Dr. Janis Boettinger (organizer and host) and the program committee who invited me to give a keynote talk at the 3rd Global Conference on Digital Soil Mapping organized by the International Union of Soil Sciences, Soil Science Society of America and Utah State University, Sept. 30–Oct. 3, 2008, Logan, UT. This provided the stimulus to conduct the review analysis on DSMM presented in this manuscript. I am very grateful for their encouragement. I would also like to thank Bob MacMillan and David Rossiter for their excellent suggestions.

References

Addiscott, T., Smith, J.M., Bradbury, N., 1995. Critical evaluation of models and their parameters. J. Environ. Qual. 24, 803–807.
Allison, V.J., Yermakov, Z., Miller, R.M., Jastrow, J.D., Matamala, R., 2007. Assessing soil microbial community composition across landscapes: do surface soils reveal patterns? Soil Sci. Soc. Am. J. 71, 730–734.
Awiti, A.O., Walsh, M.G., Shepherd, K.D., Kinyamario, J., 2008. Soil condition classification using infrared spectroscopy: a proposition for assessment of soil condition along a tropical forest-cropland chronosequence. Geoderma 143, 73–84.
Baker, L., 2008. Development of class pedotransfer functions of soil water retention — a refinement. Geoderma 144, 225–230.
Baker, L., Ellison, D., 2008. Optimization of pedotransfer functions using an artificial neural network ensemble method. Geoderma 144, 212–224.
Batjes, N.H., 2008. Mapping soil carbon stocks of Central Africa using SOTER. Geoderma 146, 58–65.
Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma 145, 28–36.
Bauer, J., Herbst, M., Huisman, J.A., Weihermüller, L., Vereecken, H., 2008. Sensitivity of simulated soil heterotrophic respiration to temperature and moisture reduction functions. Geoderma 145, 17–27.
Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A novel method of classifying soil profiles in the field using optical means. Soil Sci. Soc. Am. J. 72, 1113–1123.

Bishop, T.F.A., Minasny, B., 2006. Digital soil-terrain modeling: the predictive potential and uncertainty. In: Grunwald, S. (Ed.), Environmental Soil-Landscape Modeling — Geographic Information Technologies and Pedometrics. CRC Press, New York, pp. 185–214.

Bockheim, J.G., McLeod, M., 2008. Soil distribution in the McMurdo dry valleys, Antarctica. Geoderma 144, 43–49.

Bonilla, C.A., Norman, J.M., Molling, C.C., 2007. Water erosion estimation in topographically complex landscapes: model description and first verifications. Soil Sci. Soc. Am. J. 71, 1524–1537.

Borůvka, L., Mládková, L., Penížek, V., Drábek, O., Vasat, R., 2007. Forest soil acidification assessment using principal component analysis and geostatistics. Geoderma 140, 374–382.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth International Group, CA.

Brenning, A., Koszinski, S., Sommer, M., 2008. Geostatistical homogenization of soil conductivity across field boundaries. Geoderma 143, 254–260.

Bricklemyer, R.S., Miller, P.R., Turk, P.J., Paustian, K., Keck, T., Nielsen, G.A., 2007. Sensitivity of the Century model to scale-related soil texture variability. Soil Sci. Soc. Am. J. 71, 784–792.

Brown, D.J., 2007. Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. Geoderma 140, 444–453.

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.

Carrara, M., Castrignanó, Comparetti, A., Febo, P., Orlando, S., 2007. Mapping of penetrometer resistance in relation to tractor traffic using multivariate geostatistics. Geoderma 142, 294–307.

Causarano, H.J., Shaw, J.N., Franzluebbers, A.J., Reeves, D.W., Raper, R.L., Balkcom, K.S., Norfleet, M.L., Isaurralde, R.C., 2007. Simulating field-scale soil organic carbon dynamics using EPIC. Soil Sci. Soc. Am. J. 71, 1174–1185.

Chaudhari, S.K., Singh, R., Kundu, D.K., 2008. Rapid textural analysis for saline and alkaline soils with different physical and chemical properties. Soil Sci. Soc. Am. J. 72, 431–441.

Chen, F., Kissel, D.E., West, L.T., Adkins, W., Rickman, D., Luvall, J.C., 2008. Mapping soil organic carbon concentrations for multiple fields with image similarity analysis. Soil Sci. Soc. Am. J. 72, 186–193.

Connolly, J., Holden, N.M., Ward, S.M., 2007. Mapping peatland in Ireland using a rule-based methodology and digital data. Soil Sci. Soc. Am. J. 71, 492–499.

Corstanje, R., Grunwald, S., Lark, R.M., 2008. Inferences from fluctuations in the local variogram about the assumptions of stationarity in the variance. Geoderma 143, 123–132.

Cockx, L., van Meirvenne, M., De Vos, B., 2007. Using the EM38DD soil sensor to delineate clay lenses in a sandy forest soil. Soil Sci. Soc. Am. J. 71, 1314–1322.

Douaik, A., van Meirvenne, M., Toth, T., 2007. Statistical methods for evaluating soil salinity spatial and temporal variability. Soil Sci. Soc. Am. J. 71, 1629–1635.

Du, C., Linker, R., Shaviv, A., 2008. Identification of agricultural soils using mid-infrared photoacoustic spectroscopy. Geoderma 143, 85–90.

Eldeiry, A.A., Garcia, L.A., 2008. Detecting soil salinity in alfalfa fields using spatial modeling and remote sensing. Soil Sci. Soc. Am. J. 72, 201–211.

Farifteh, J., van der Meer, F., van der Meijde, M., Atzberger, C., 2008. Spectral characteristics of salt-affected soils: a laboratory experiment. Geoderma 145, 196–206.

Finke, P.A., Hutson, J.L., 2008. Modelling soil genesis in calcareous loess. Geoderma 145, 462–479.

Genxu, W., Yuanshou, L., Yibo, W., Qingbo, W., 2008. Effects of permafrost thawing on vegetation and soil carbon pool losses on the Qinghai — Tibet Plateau, China. Geoderma 143, 143–152.

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008. Soil organic prediction by hyperspectral remote sensing and field VIS-NIR spectroscopy: an Australian case study. Geoderma 146, 403–411.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island — digital soil mapping using Random Forest analysis. Geoderma 146, 102–113.

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. (Ed.), 2006. Environmental Soil-Landscape Modeling — Geographic Information Technologies and Pedometrics. CRC Press, New York.

Grunwald, S., 2008. Reconstruction and scientific visualization of earthscapes considering trends and spatial dependence structures. New J. Phys. 10 (No. 125011) 15 pp.

Grunwald, S., Reddy, K.R., 2008. Spatial behavior of phosphorus and nitrogen in a subtropical wetland. Soil Sci. Soc. Am. J. 72, 1174–1183.

Grunwald, S., Rivero, R.G., Reddy, K.R., 2007. Understanding spatial variability and its application to biogeochemistry analysis (chapter 20). In: Sarkar, D., Datta, R., Hannigan, R. (Eds.), Environmental Biogeochemistry — Concepts and Case Studies. Elsevier, pp. 435–462.

Grunwald, S., Osborne, T.Z., Reddy, K.R., 2008. Temporal trajectories of phosphorus and pedo-patterns mapped in Water Conservation Area 2, Everglades, Florida, USA. Geoderma 146, 1–13.

Guerrero, C., Mataix-Solera, J., Arcenegui, V., Mataix-Beneyto, J., Gomez, I., 2007. Near-infrared spectroscopy to estimate the maximum temperatures reached on burned soils. Soil Sci. Soc. Am. J. 71, 1029–1037.

De Gryze, S., Six, J., Bossuyt, H., Van Oost, K., Merckx, R., 2008. The relationship between landform and the distribution of soil C, N and P under conventional and minimum tillage. Geoderma 144, 180–188.

Hand, D., Mannila, H., Smyth, P., 2001. Principles of Data Mining. Massachusetts Institute of Technology. MIT Press, Cambridge, Massachusetts.

Hartemink, A.E. (Ed.), 2006. The Future of Soil Science. Int. Union of Soil Sciences.

Hartemink, A.E., McBratney, A.B., 2008. A soil science renaissance. Geoderma 148, 123–129.

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

Hengl, T., Toormanian, N., Reuter, H.I., Malakouti, M.J., 2007. Methods to interpolate soil categorical variables from profile observations: lessons from Iran. Geoderma 140, 417–427.

Hill, J., Schütt, B., 2000. Mapping complex patterns of erosion and stability in dry Mediterranean ecosystems. Remote Sens. Environ. 75, 557–569.

Janik, L.J., Merry, R.H., Forrester, S.T., Lanyon, D.M., Rawson, A., 2007. Rapid prediction of soil water retention using mid-infrared spectroscopy. Soil Sci. Soc. Am. J. 71, 507–514.

Jasques, D., Šimůnek, J., Mallants, D., van Genuchten, M.Th., 2008. Modelling coupled water flow, solute transport and geochemical reactions affecting heavy metal migration in a podzol soil. Geoderma 145, 449–461.

Jiang, P., Anderson, S.H., Kitchen, N.R., Dudduth, K.A., Sadler, E.J., 2007. Estimating plant-available water capacity for claypan landscapes using apparent electrical conductivity. Soil Sci. Soc. Am. J. 71, 1902–1908.

Jordanova, N., Jordanova, D., Tsacheva, T., 2008. Application of magnetometry for delineation of anthropogenic pollution in areas covered by various soil types. Geoderma 144, 557–571.

Kazemi, H.V., Anderson, S.H., Goyne, K.W., Gantzer, C.J., 2008. Spatial variability of bromide and atrazine transport parameters for a Udipsamment. Geoderma 144, 545–556.

Kerry, P., Oliver, M.A., 2007a. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma 140, 383–396.

Kerry, P., Oliver, M.A., 2007b. The analysis of ranked observations of soil structure using indicator geostatistics. Geoderma 140, 397–416.

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R., Buydens, L.M.C., 2003. The potential of field spectroscopy for the assessment of sediment properties in river floodplains. Anal. Chim. Acta 484, 189–200.

Lamorski, K., Pachepsky, Y., Slawiński, C., Walczak, R.T., 2008. Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. Soil Sci. Soc. Am. J. 72, 1243–1247.

Li, W., 2007. Transiograms for characterizing spatial variability of soil classes. Soil Sci. Soc. Am. J. 71, 881–893.

Li, W., Zhang, C., 2007. A random-path Markov chain algorithm for simulating categorical soil variables from random point samples. Soil Sci. Soc. Am. J. 71, 656–668.

Li, S., Lobb, D.A., Lindstrom, M.J., Papiernik, S.K., Farenhorst, A., 2008. Modeling tillage-induced redistribution of soil mass and its constituents within different landscapes. Soil Sci. Soc. Am. J. 72, 167–179.

Lilly, A., Nemes, A., Rawls, W.J., Pachepsky, Y.A., 2008. Probabilistic approach to the identification of input variables to estimate hydraulic conductivity. Soil Sci. Soc. Am. J. 72, 16–24.

Liu, J., Pattey, E., Nolin, M.C., Miller, J.R., Ka, O., 2008. Mapping within-field soil drainage using remote sensing, DEM and apparent soil electrical conductivity. Geoderma 143, 261–272.

Lobell, D.B., Ortiz-Monasterio, J.I., Gurrola, F.C., Valenzuela, L., 2007. Identification of saline soils with multiyear remote sensing of crop yields. Soil Sci. Soc. Am. J. 71, 777–783.

Lu, J., Jiang, P., Wu, L., Chang, A.C., 2008. Assessing soil quality data by positive matrix factorization. Geoderma 145, 259–266.

Lufafa, A., Diédhiou, I., Samba, S.A.N., Séné, M., Khouma, M., Kizito, F., Dick, R.P., Dossa, E., Noller, J.S., 2008. Carbon stocks and patterns in native shrub communities of Senegal's Peanut Basin. Geoderma 146, 75–82.

Mabit, L., Bernard, C., Makhlouf, M., Laverdière, M.R., 2008. Spatial variability of erosion and soil organic matter content estimated from ^{137}Cs measurements and geostatistics. Geoderma 146, 245–251.

MacMillan, R.A., Moon, D.E., Coupé, R.A., 2007. Automated predictive ecological mapping in a forest region of B.C., Canada, 2001–2005. Geoderma 140, 353–373.

Marchant, B.P., Lark, R.M., 2007. The Matérn variogram model: implications for uncertainty propagation and sampling in geostatistical surveys. Geoderma 140, 337–345.

McBratney, A.B., 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. Nutr. Cycl. Agroecosyst. 50, 51–62.

McBratney, A.B., Pringle, M.J., 1999. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. Precision Agric. 1, 125–152.

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.

Meersman, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143, 1–13.

Minasny, B., McBratney, A.B., 2007a. Estimating the water retention shape parameter from sand and clay content. Soil Sci. Soc. Am. J. 71, 1105–1110.

Minasny, B., McBratney, A.B., 2007b. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. Geoderma 140, 324–336.

Minasny, B., McBratney, A.B., 2007c. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. Geoderma 142, 285–293.

Nield, S.J., Boettinger, J.L., Ramsey, R.D., 2007. Digitally mapping gypsic and natric soil areas using Landsat ETM data. Soil Sci. Soc. Am. J. 71, 245–252.

Nyssen, J., Tmesgen, H., Lemenih, M., Zenebe, A., Haregeweyn, N., Haile, M., 2008. Spatial and temporal variation of soil organic carbon stocks in a lake retreat area of the Ethiopian Rift Valley. Geoderma 146, 261–268.

Oorts, K., Garnier, P., Findeling, A., Mary, B., Richard, G., Nicolardot, B., 2007. Modeling soil carbon and nitrogen dynamics in no-till and conventional tillage using PASTIS model. Soil Sci. Soc. Am. J. 71, 336–346.
Ouimet, R., 2008. Using compositional change within soil profiles for modeling base cation transport and chemical weathering. Geoderma 145, 410–418.
Parasuraman, K., Elshorbagym, A., Si, S.C., 2008. Estimating saturated hydraulic conductivity using genetic programming. Soil Sci. Soc. Am. J. 71, 1676–1684.
Park, E., Elfeki, A.M.M., Song, Y., Kim, K., 2007. Generalized coupled Markov chain model for characterizing categorical variables in soil mapping. Soil Sci. Soc. Am. J. 71, 909–917.
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.
Reinds, G.J., van Oijen, M., Heuvelink, G.B.M., Kros, H., 2008. Bayesian calibration of the VSD soil acidification model using European forest monitoring data. Geoderma 146, 475–488.
Reynolds, W.D., Drury, C.F., Yang, X.M., Tan, C.S., 2008. Optimal soil physical quality inferred through structured regression and parameter interactions. Geoderma 146, 466–474.
Rivero, R.G., Grunwald, S., Bruland, G.L., 2007. Incorporation of spectral data into multivariate geostatistical models to map soil phosphorus variability in a Florida wetland. Geoderma 140, 428–443.
Sadeghi, S.H.R., Mizuyama, T., Miyata, S., Gomi, T., Kosugi, K., Fukushima, T., Mizugaki, S., Onda, Y., 2008. Development, evaluation and interpretation of sediment rating curves for a Japanese small mountainous reforested watershed. Geoderma 144, 198–211.
Santra, P., Das, B.S., 2008. Pedotransfer functions for soil hydraulic properties developed from a hilly watershed of Eastern India. Geoderma 146, 439–448.
Schulp, C.J.E., Veldkamp, A., 2008. Long-term landscape–land use interactions as explaining factor for soil organic matter variability in Dutch agricultural landscapes. Geoderma 146, 457–465.
Shukla, M.K., Lal, R., VanLeeuwen, D., 2007. Spatial variability of aggregate-associated carbon and nitrogen contents in the reclaimed mine soils of eastern Ohio. Soil Sci. Soc. Am. J. 71, 1748–1757.
Sommer, M., Gerke, H.H., Deumlich, D., 2008. Modelling soil landscape genesis — a “time split” approach for hummocky agricultural landscapes. Geoderma 145, 480–493.
Stevens, A., van Wesemael, B., Bartholomeus, H., Rossillon, D., Tychon, B., Ben-Dor, E., 2008. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. Geoderma 144, 395–404.
Thompson, J.A., Bell, J.C., Butler, C.A., 2001. Digital elevation model resolution: effects on terrain attribute calculations and quantitative soil-landscape modeling. Geoderma 100, 67–89.
Tranter, G., Minasny, B., McBratney, A.B., Viscarra Rossel, R.A., Murphy, B.W., 2008. Comparing spectral soil inference systems and mid-infrared spectroscopic predictions of soil moisture retention. Soil Sci. Soc. Am. J. 72, 1394–1400.
Vasques, G.M., Grunwald, S., Sickman, J.O., 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma 146, 14–25.
Vasques, G.M., Grunwald, S., Sickman, J.O., 2009. Modeling of soil organic carbon fractions using visible/near-infrared spectroscopy. Soil Sci. Soc. Am. J. 73, 176–184.
Viscarra Rossel, R., on behalf of the Global Soil Spectroscopy Working Group (40 collaborators from 27 countries), 2008. The soil spectroscopy group and the development of a global spectral library. 3rd Global Conference on Digital Soil Mapping organized by the International Union of Soil Sciences, Soil Science Society of America and Utah State University, Sept. 30–Oct. 3, 2008, Logan, UT.
Vitharana, U.W.A., Van Meirvenne, M., Simpson, D., Cockx, L., De Baerdemaeker, J., 2008. Key soil and topographic properties to delineate potential management classes for precision agriculture in the European loess area. Geoderma 143, 206–215.
Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Sci. Soc. Am. J. 71, 389–396.
Weller, A., Zipprich, M., Sommer, M., Zu Castell, W., Wehrhan, M., 2007. Mapping clay content across boundaries at the landscape scale with electromagnetic induction. Soil Sci. Soc. Am. J. 71, 1740–1747.
West, T.O., Wilson, B.S., Hellwinckel, C.M., Tyler, D.D., Marland, G., De La Torre Ugarte, D., Larson, J.A., Nelson, R.G., 2008. Estimating regional changes in soil carbon with high spatial resolution. Soil Sci. Soc. Am. J. 72, 285–294.
Wills, S.A., Burras, C.L., Sandor, J.A., 2007. Prediction of soil organic carbon content using field and laboratory measurements of soil color. Soil Sci. Soc. Am. J. 71, 380–388.
Winzeler, H.E., Owens, P.R., Joern, B.C., Camberato, J.J., Lee, B.D., Anderson, D.E., Smith, D.R., 2008. Potassium fertility and terrain attributes in a Fragiudalf drainage catena. Soil Sci. Soc. Am. J. 72, 1311–1320.
Wöhling, T., Vrugt, J.A., Barkle, G.F., 2008. Comparison of three multiobjective optimization algorithms for inverse modeling of vadose zone hydraulic properties. Soil Sci. Soc. Am. J. 72, 305–319.
Wu, Y., Chen, J., Ji, J., Gong, P., Liao, Q., Tian, Q., Ma, H., 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. Soil Sci. Soc. Am. J. 71, 918–926.
Wu, C., Wu, J., Luo, Y., Zhang, H., Teng, Y., 2008. Statistical and geostatistical characterization of heavy metal concentrations in a contaminated area taking into account soil map units. Geoderma 144, 171–179.
Yoo, K., Mudd, S.M., 2008. Toward process-based modeling of geochemical soil formation across diverse landforms: a new mathematical framework. Geoderma 146, 248–260.
Zacharias, S., Wessolek, G., 2007. Excluding organic matter content from pedotransfer predictors of soil water retention. Soil Sci. Soc. Am. J. 71, 43–50.
Zhang, Y., Zhao, Y.C., Shi, X.Z., Lu, X.X., Yu, D.S., Wang, H.J., Sun, W.X., Darilek, J.L., 2008a. Variation of soil organic carbon estimates in mountain regions: a case study from Southwest China. Geoderma 146, 449–456.
Zhang, C., Fay, D., McGrath, D., Grennan, E., Carton, O.T., 2008b. Statistical analysis of geochemical variables in soils of Ireland. Geoderma 146, 378–390.
Zibilske, L.M., Bradford, J.M., 2007. Soil aggregation, aggregate carbon and nitrogen, and moisture retention induced by conservation tillage. Soil Sci. Soc. Am. J. 71, 793–802.
Zinn, Y.L., Lal, R., Bigham, J.M., Resck, D.V.S., 2007. Edaphic controls on soil organic retention in the Brazilian Cerrado: texture and mineralogy. Soil Sci. Soc. Am. J. 71, 1204–1214.