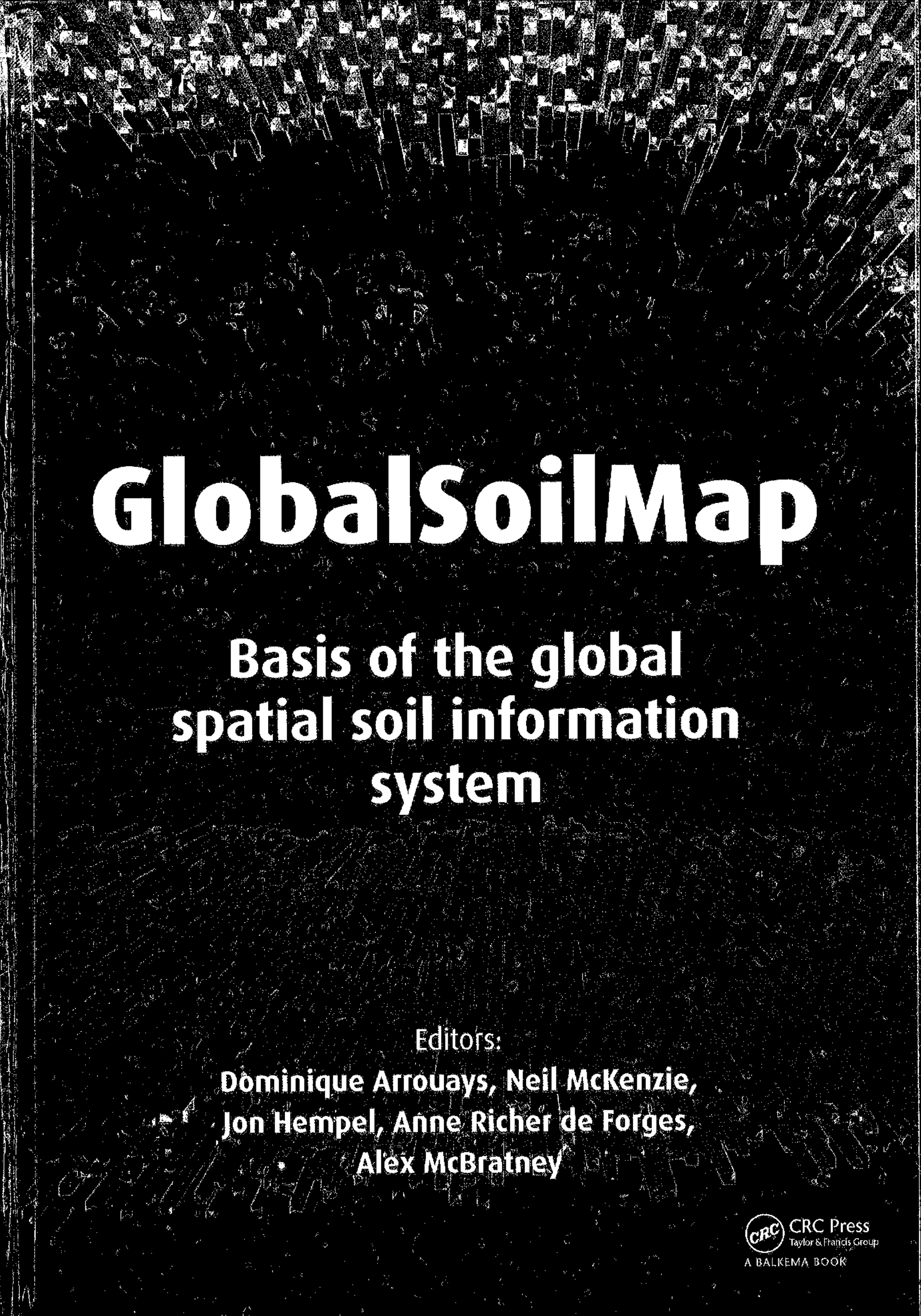

GlobalSoilMap

Basis of the global spatial soil information system

Editors:
Dominique Arrouays, Neil McKenzie,
Jon Hempel, Anne Richer de Forges,
Alex McBratney

CRC Press
Taylor & Francis Group
A BALKEMA BOOK

PROCEEDINGS OF THE 1ST GLOBALSOILMAP CONFERENCE, ORLÉANS, FRANCE, 7–9 OCTOBER 2013

GlobalSoilMap

Basis of the global spatial soil information system

Editors

Dominique Arrouays
INRA, France

Neil McKenzie
CSIRO, Australia

Jon Hempel
NSSC, USA

Anne C. Richer de Forges
INRA, France

Alex McBratney
University of Sydney, Australia

CRC Press
Taylor & Francis Group
Boca Raton London New York Leiden

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A BALKEMA BOOK

CRC Press/Balkema is an imprint of the Taylor & Francis Group, an informa business

Typeset by V Publishing Solutions Pvt Ltd., Chennai, India
Printed and bound in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY

Published by: CRC Press/Balkema
P.O. Box 11320, 2301 EH Leiden, The Netherlands
e-mail: Pub.NL@taylorandfrancis.com
www.crcpress.com – www.taylorandfrancis.com

ISBN: 978-1-138-00119-0 (Hbk)
ISBN: 978-1-315-77558-6 (eBook PDF)

GlobalSoilMap – Arrouays et al. (Eds)
© 2014 Taylor & Francis Group, London, ISBN 978-1-138-00119-0

Part II—Integration of data to work towards a Meta Soil Carbon Model in the U.S.

S. Grunwald, B. Cao, X. Xiong & C.W. Ross
University of Florida, FL, USA

R. Patarasuk
Arizona State University, AZ, USA
Former affiliation: Soil and Water Science Department, University of Florida, FL, USA

J.W. Hempel, L.T. West, S.S. Andrews & S. Wills
Natural Resources Conservation Service, Lincoln, NE, USA

T.D. Loecke
University of Nebraska, NE, USA

ABSTRACT: This paper presents the development of a soil-environmental data superstructure for the U.S. in context of a coherent Meta Soil Carbon Model. The Meta Soil Model provides a formalized approach to create digital soil realizations grounded in various philosophical worldviews and is described in detail in part I. We assembled a wide suite of environmental covariates from a variety of sources representing potential soil forming factors in consistent format at the finest spatial resolution available across the U.S. A total of 250+ geospatial environmental variables were acquired and many were mosaicked from tiles. These include 700+ spatial layers representing different time periods of environmental variables. In addition we created/retrieved various regional and national spatially-explicit soil carbon datasets. Soil and environmental data were harmonized to prepare digital modeling to create a multiplicity of soil realizations (maps and models) including those that meet the specifications of the Global Soil Map project.

1 INTRODUCTION

1.1 *Background*

We adopt the conceptual framework for creation of a Meta Soil Model (MSM) presented in part I combining ontology (i.e., building of a meta-data model that represents the total ecosystem), methodology (i.e., fusing/bundling of various methods to build digital soil models), and perspective-dimensions such as incorporating space-time dimensions to describe the evolution of soil systems. The MSM provides a formalized approach to create digital soil realizations grounded in various philosophical worldviews including axiology (*the Why*), motivations, needs, and purposes (*for What/Whom*), ontology (*the What*) and hermeneutics (interpretation), epistemology (*the Who*), and methodology (*the How*).

1.2 *Why do we need a Meta Soil Model?*

Many soil samples have been collected within the U.S. as part of national soil survey programs by the Natural Resources Conservation Service (NRCS)—U.S. Department of Agriculture, regional and State agencies, research and extension programs at academic institutions, private industry, consultants, and land owners. This has brought forth one of the largest national pedon databases. Despite these efforts sample collection that peaked in the 1980–1990 has steadily declined since the 1990s. Considering the approximate costs for soil surveying in the U.S. of US \$10.30 ha^{-1} and size of the U.S. (about 9.827 million km^2) (Grunwald et al., 2011) the need for efficient Digital Soil Mapping (DSM) is high.

Over the past decades budget limitations have impacted the capacity to update digital soil maps (e.g., the Soil Survey Geographic Database, SSURGO, and the U.S. General Soil Map, STATSGO) at the same speed as soil is changing. Digital soil mapping is critical to provide quantitative assessment methods to operationalize the process of production of digital soil maps.

Environmental covariates are utilized in DSM and many spatial data layers representing historic soil conditions, topography, ecology, parent material, atmosphere, water/hydrology, biota, and human-induced activities/impacts. These data to represent the total ecosystems are readily available from local, state, tribal, and federal government, NGOs, academia, private companies, and volunteers/crowds due to the open sharing policy adopted in the U.S. The Federal Geographic Data Committee coordinates the development of the National Spatial Data Structure (NSDS). The Geospatial Platform of NSDS focuses on web applications that facilitate participatory information sharing, interoperability, user-centered design, and collaboration on the World Wide Web, including meta data standards. However, at the current time these geospatial data are not fully harmonized to a common map projection, scale, and platform that would also include all satellite based geospatial products from the National Aeronautics and Space Administration, NASA (see latest business plan of NSDS, Sept. 2012). Other national data integration efforts are in the planning phase, such as EarthCube (National Science Foundation). Hence, the needs for a coherent, integrated geospatial data structure that fuses soil and environmental geospatial data to support DSM are evident.

1.3 *Objectives*

In this paper we present the development of a soil-environmental data superstructure for the U.S. in context of a coherent Meta Soil Carbon Model that reflects the state of knowledge.

2 APPROACH

2.1 *Environmental covariates*

We assembled a wide suite of environmental covariates from a variety of sources representing potential soil forming factors in consistent format at the finest spatial resolution available across the U.S. A total of 250+ geospatial environmental variables were acquired and many were mosaicked from tiles. These include 700+ spatial layers representing different time periods of environmental variables. An overview of the data sources is provided in Table 1. Several space-time datasets were harmonized into a standard format (e.g., climate data such as temperature and precipitation covering monthly values for the past decades, or remote sensing derived data from the Moderate-Resolution Imaging Spectroradiometer (MODIS) at 16 day frequency for leaf area index, Normalized Difference Vegetation Index, etc.). Geospatial data included soil data from gridded Soil Survey Geographic Database (gSSURGO), soil moisture data (Soil Moisture and Salinity Data, SMOS), climate data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) and Idaho Geospatial, downscaled climate change projections and historic climate records (North American Regional Climate Change Assessment Program, NARCCAP), geology layers and gamma ray data to infer on geology, various layers of land use/land cover representing different time periods, canopy and vegetation data (Landscape Fire and Resource Management Planning Tools, LANDFIRE), biomass data (National Biomass and Carbon Data), ecoregions, National Elevation Dataset (NED) and Shuttle Radar Topographic Mission (SRTM) elevation data, and primary and secondary topographic derivatives. The U.S. environmental covariate set amounts to ~8 TeraBytes. Data integration and fusion was conducted using a combination of ArcGIS, Saga GIS, Structured Query Language (SQL) database and queries, R, customized code, and various other data source specific extractions and conversion.

2.2 *Soil properties*

We created/retrieved various regional and national spatially-explicit soil carbon datasets including (i) National Cooperative Soil Survey Characterization Data (Natural Resource Conservation Service, NRCS) (n = 258,094 samples), (ii) Rapid Carbon Assessment Project (RaCA) NRCS (n = 144,833 samples), (iii) Pine Integrated Network: Education, Mitigation, and Adaptation project (PINEMAP) covering the southeastern U.S. (tier 1 set with n = 750 [forest production data]; plus future tier 2 and 3 sets containing soil carbon data), (iv) Rapid Assessment and Trajectory Modeling of Changes in Soil Carbon across Florida (n = 1,014 topsoil samples), (v) Florida Historic Soil Characterization Data (n = 8,269 soil samples), (vi) National Soil Carbon Network, (vii) North American Carbon Program, and (viii) other university-based available soil carbon datasets.

2.3 *Data integration*

We used spatial extractions to delineate environmental covariates that match pedon locations. We projected the environmental covariates (GIS layers) to ensure consistency in geo-referencing. The covariate values were then extracted to the geo-referenced soil sample locations. Missing values in bulk density will be predicted by a pedotransfer function. Soil carbon data by genetic horizons will be converted to continuous values using a depth-weighted average and equal-area quadratic spline function (Bishop et al., 1999).

We (will) fuse(d) all soil carbon data (and associated bulk density and ancillary environmental attributes) into our web-facilitated data system: the

Table 1. Overview of soil-environmental variables assembled and harmonized for the United States.

Types	Soil-environmental variables	Data source *(Additional information)*
Soils/water	Soil order Soil suborder Soil group Soil subgroup Annual minimum water table depth Annual minimum water table depth from April to June Available water storage top 25 cm Available water storage top 50 cm Available water storage top 100 cm Available water storage top 150 cm Organic matter (low) Organic matter (representative) Organic matter (high) Soil texture Drainage class (dominant) Hydrologic group (dominant)	**Gridded Soil Survey Geographic Database (gSSURGO) (Natural Resources Conservation Service, NRCS)** Legacy soil data *10 m spatial resolution*
Soils/water	10-day average soil moisture 10-day variance soil moisture Monthly average soil moisture Monthly variance soil moisture Seasonal average soil moisture Seasonal variance soil moisture Yearly average soil moisture Yearly variance soil moisture	**Soil Moisture Data (SMOS) (European Space Agency, ESA)** *Time period: July 2010 to June 2011* *Spatial resolution: 15 km*
Atmosphere	Precipitation Minimum temperature (Tmin) Maximum temperature (Tmax) Dew point temperature Mean temperature (average of Tmin and Tmax) Vapor pressure	**Parameter-elevation Regressions on Independent Slopes Mode (PRISM) (Oregon State University)** *Period: 1970 to 2011* *Monthly climatic data;* *Yearly averages of climatic data;* *Long-term (30 year) average of climatic data.* *Spatial resolution: ~4 km; 2.5 arcmin* *Spatial resolution (SE U.S.): 800 m; 30 arcsec*
Atmosphere	Precipitation Maximum relative humidity Minimum relative humidity Mean specific humidity Mean downward shortwave radiation at surface Mean wind direction Minimum temperature Maximum temperature Mean wind speed	**Idaho Geospatial Data Center (University of Idaho)** *Period: 1971 to 2011* *Daily climatic data;* *Monthly climatic data;* *Yearly averages of climatic data;* *Long-term (30 year) average of climatic data.* *Spatial resolution: ~4 km; 2.5 arcmin*
Atmosphere	Mean monthly solar radiation—Jan Mean monthly solar radiation—Feb Mean monthly solar radiation—Mar Mean monthly solar radiation—Apr Mean monthly solar radiation—May Mean monthly solar radiation—Jun Mean monthly solar radiation—Jul Mean monthly solar radiation—Aug Mean monthly solar radiation—Sep Mean monthly solar radiation—Oct Mean monthly solar radiation—Nov Mean monthly solar radiation—Dec Mean solar radiation Standard deviation solar radiation	**North American Regional Reanalysis (NARR)—National Oceanic and Atmospheric Administration (NOAA)** *Temporal resolution: 30 yrs averages for each month* *Spatial resolution: 32 km*

(*Continued*)

Table 1. (*Continued*).

Types	Soil-environmental variables	Data source *(Additional information)*
Atmosphere	Surface air temperature Precipitation Surface downdwelling shortwave radiation Surface pressure Surface specific humidity Num. frost days Minimum monthly temperature Maximum monthly temperature Mean daily minimum temperature Mean daily maximum temperature	**North American Regional Climate Change Assessment Program (NARCAAP)** *Historic climatic data (1971–2000);* *Climate change projections (5 different scenarios) (downscaled to the finest resolution in North America) 2041–2070* *Spatial resolution: 50 km*
Geology	Geologic formations	**Mineral Resources Program (U.S. Geological Survey, USGS)** *Map scale: 1:100,000* *Shapefile (polygon)*
Geology	Aeroradiometric survey—Absorbed dose Aeroradiometric survey—Potassium Aeroradiometric survey—Thorium Aeroradiometric survey—Uranium Bouguer gravity Isostatic Magnetic anomaly—Magnetic anomaly Residual total intensity Magnetic anomaly—Filtered anomaly high-pass 500 km Magnetic anomaly—Residual using satellite baseline	**Gamma Ray (USGS)** *Spatial resolutions depending on layers with 1, 2 and 4 km*
Biota	Normalized Difference Vegetation Index (NDVI) 16-day frequency Enhanced Vegetation Index (EVI) 16-day frequency Leaf Area Index (LAI) 16-day frequency Fraction of Photosynthetically Active Radiation (FPAR) 16-day frequency Gross Primary Productivity (GPP) Net Primary Productivity (NPP)	**Moderate-resolution Imaging Spectro-radiometer (MODIS) (NASA)** *Time period: 2010* *Frequency: 16 days (layer of each variable every 16 days in 2010)* *Spatial resolution: 500 m (NDVI and EVI); all other variables: 1000 m*
Biota	Existing Vegetation Cover (EVC) Existing Vegetation Height (EVH) Existing Vegetation Type (EVT) BioPhysical Settings (BPS) Environmental Site Potential (ESP) Canopy Bulk Density (CBD) Canopy Base Height (CBH) Forest Canopy Cover (CC) Forest Canopy Height (CH) Fire Regime Groups (FRG) Mean Fire Return Interval (MFRI) Percent Low-severity Fire (PLS) Percent Mixed-severity Fire (PMS) Percent High-severity Fire (PRS) Succession Class (SCLASS) Disturbance 1999–2008	**LANDFIRE program (various agencies)** *Period: 2008* *Spatial resolution: 30 m*
Biota	Landuse	**National Land Characterization Data (NLCD) (USGS)** *Period: 2006* *Spatial resolution: 30 m*

(*Continued*)

Table 1. (*Continued*).

Types	Soil-environmental variables	Data source *(Additional information)*
Biota	Cropland	**National Agricultural Statistics Service (NASS) (United States Department of Agriculture, USDA)** *Time period: 2002–present* *Map scale: 1:100,000*
Biota	Above-ground wooden biomass carbon Basal area weighted canopy height	**Wood Hole Research Center** *Time period: 2000* *Spatial resolution: 30 m*
Biota	Ecoregion	**Environmental Protection Agency (EPA)** *Map scale: 1:250,000* *Shapefile (polygon)*
Human	Population density	**U.S. Census Bureau**
Topography	Digital surface model Elevation Slope Plan curvature Profile curvature Aspect Flow accumulation Compound topographic index	**Shuttle Radar Topography Mission (SRTM)** GAP-filled SRTM-DEM from the Consultative Group on International Agricultural Research (CGIAR) *Spatial resolution: 90 m (3 arc sec)*
Topography	Digital elevation model Elevation Slope Plan curvature Profile curvature Aspect Flow accumulation Compound topographic index	**National Elevation Dataset (NED) (USGS)** *Spatial resolution: ~30 m (1 arc sec)*

Terrestrial Carbon Information System (TerraC) (http://TerraC.ifas.ufl.edu). TerraC allows public and private data sharing and invites collaborative research through a Creative Commons oriented data-sharing model. We envision that this soil carbon data superstructure supports the application and testing of various mixed methodologies, paradigms, and perspectives to build a Meta Soil Carbon Model for the U.S. and feed into the Global Soil Map and other terrestrial carbon mapping initiatives.

In TerraC data are stored in form of project sets that characterize (i) soil, (ii) vegetation, (iii) atmosphere, (iv) water, and (v) whole ecosystems. The overall data structure includes: (i) individual project datasets; umbrella projects, (ii) core/base (mandatory) data for each project, (iii) attribute data, and (iv) meta data. The TerraC system allows running queries across projects and integrating data into pooled sets facilitating synthesis across large areas (e.g., portions of the U.S. and conterminous U.S.), ecological or climatic trajectories, thematic foci (e.g., soil or terrestrial carbon), and hypothesis driven research.

The underlying idea of the Meta Soil Carbon Model is to integrate soil and environmental data, and applications of multiple DSM methods and paradigms guided by values and needs of end-users. It supports efforts of the Global Soil Map to produce a digital soil carbon map according to specifications. Beyond that the Meta Soil Carbon Model framework supports multiple other purposes, such as historic-current soil carbon assessment, cross-comparison of soil carbon in different regions, pooling effects of soil carbon data across escalating spatial and temporal scales, and more.

3 FINAL REMARKS

The massive data amount of space-time environmental covariates provides opportunities for DSM, but also technical/computational challenges for coherent integration with soil data. We envision that the Meta Soil Carbon Model can fulfill multiple objective that go beyond contribution to the Global Soil Map project. It will provide insights to address a variety of needs to create soil data, models, and realizations of the future to address complex socio-economic, environmental, and ecological issues of concern at regional, national, and global scales.

REFERENCES

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

Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: Finding solutions for global issues. Soil Sci. Soc. Am. J. 75, 1201–1213.