Geoderma 213 (2014) 29–35

Contents lists available at ScienceDirect

Geoderma

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

Interpolation of Mexican soil properties at a scale of 1:1,000,000

Gustavo Cruz-Cárdenas [a,*], Lauro López-Mata [b], Carlos Alberto Ortiz-Solorio [b], José Luis Villaseñor [c], Enrique Ortiz [c], José Teodoro Silva [a], Francisco Estrada-Godoy [a,d]

[a] Instituto Politécnico Nacional, CIIDIR-IPN-Michoacán, COFAA, Justo Sierra 28, 59510 Jiquilpan, Michoacán, México
[b] Colegio de Postgraduados, Campus Montecillo, Carretera México-Texcoco km 36.5, 56230 Texcoco, Estado de México, México
[c] Instituto de Biología, Departamento de Botánica, Universidad Nacional Autónoma de México, Tercer Circuito s/n, Ciudad Universitaria, 04510 Delegación Coyoacán, Mexico D.F., México
[d] Instituto de Horticultura, Departamento de Fitotecnia, Universidad Autónoma Chapingo, Carretera México-Texcoco km 38.5, 56230 Texcoco, Estado de México, México

ARTICLE INFO

Article history:
Received 21 June 2011
Received in revised form 2 July 2013
Accepted 21 July 2013
Available online xxxx

Keywords:
Geostatistics
Kriging
Soil properties
Spatial databases

ABSTRACT

México maintains several databases with biotic or abiotic information that enable large scale-studies (for example, at a resolution of 1 km^2); unfortunately, there is no information at this resolution for soil properties. The goal of this paper was to generate a set of soil variables to address this absence. We evaluated 4400 soil samples taken on Mexican territory. The following nine soil properties were evaluated for each sample: Ca, K, Mg, Na, organic C, organic matter, electrical conductivity, sodium absorption ratio, and pH. With the use of geostatistical methods, a layer was generated for each soil property after six different semivariance models were evaluated. The kriging model (simple, ordinary, or universal) was selected as the best semivariance method with a 10-fold cross-validation approach. Exponential, pentaspherical, and spherical models were selected with ordinary and universal kriging methods to spatially predict the soil properties. Five classes for each soil map were generated. The nine soil properties were classified primarily as slightly low or medium according to their distribution throughout Mexican territory.

1. Introduction

Knowledge of the distribution of soil properties is important because it provides information regarding which actions to take, including sampling strategies, specific handling methods, and studies of the relationships between soil properties and the distribution patterns of other parameters (Ayoubi et al., 2007; Cruz-Cárdenas et al., 2012; López-Mata et al., 2012). Classic statistical methods should not be used to analyze spatial distribution because they are based on the assumption of seasonality in space and time, independent data, and an identical distribution of parameters (Rossi et al., 1992). Instead, starting from the assumption that the value at any given point is not independent of the values of neighboring points, there is a spatial dependence that allows the use of geostatistical analysis techniques (Wagner, 2003). The first stage in geostatistical methods is to evaluate the spatial structure of the data, which is usually described using an experimental variogram. This is a graph of the semivariance between pairs and their distance in geographical space. A variogram is defined by permissible theoretical models (exponential, spherical, and logarithmic) and the following parameters: the sill—the difference of the mean squares of two independent observations; the range—the maximum distance at which pairs of observations can be influenced or are autocorrelated; and the nugget—the variance within a sampling unit. After the spatial autocorrelation structure has been determined with the variogram, a kriging interpolation is carried out (Mulla and McBratney, 1999). Several studies have been performed using geostatistical techniques at a plot or regional level to analyze the spatial distribution of soil properties (Hengl et al., 2007; Kravchenko and Bullock, 1999; Mueller et al., 2001; Robinson and Metternicht, 2006). However, at the national or global level, there have been only a few studies; nevertheless, the FAO's Harmonized World Soil Database has assumed the task of gathering existing information on soil properties to generate worldwide maps and, therefore, national maps.

Large-scale global or national databases (for example, at a resolution of 1 km^2) are important to create data entries that can be used by computer algorithms to model phenomena of interest. Such databases make it possible to determine the behavior of phenomena with greater confidence, either spatially or temporally. Several of these databases are accessible online; for example, of the 19 bioclimatic variables of WorldClim (http://www.worldclim.org/), 11 are temperature variables and the remaining 8 are precipitation variables (Hijmans et al., 2005). Remote sensing databases, such as the MODIS database, are also available and contain information on a daily, 8-, 16-, or 30-day basis. Digital elevation models are another dataset commonly used in environmental analyses. These variables are widely used, especially to model the potential distribution of species (Elith et al., 2006), to digitally map soils (McBratney et al., 2003), and to evaluate land-use changes (Hansen et al., 2000).

* Corresponding author. Tel.: +52 3535330083.
E-mail address: gustavo.cruz.cardenas@gmail.com (G. Cruz-Cárdenas).

0016-7061/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geoderma.2013.07.014

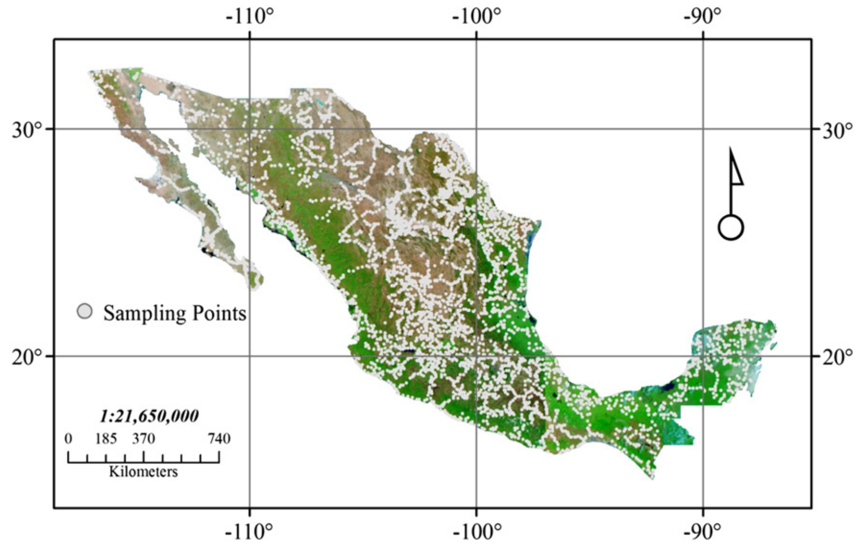

Fig. 1. Soil sampling point distribution in Mexico.

México is one country that has no available information on soil properties at a national level. This information would be useful because it would complement research on modeling, cartography, and soil-use change, among others. Therefore, the aim of this paper is to generate and present soil coverages for nine soil properties on a large scale,1 km^2 resolution (30" arc). To do so, we used a dataset obtained from soil sampling throughout Mexican continental territory.

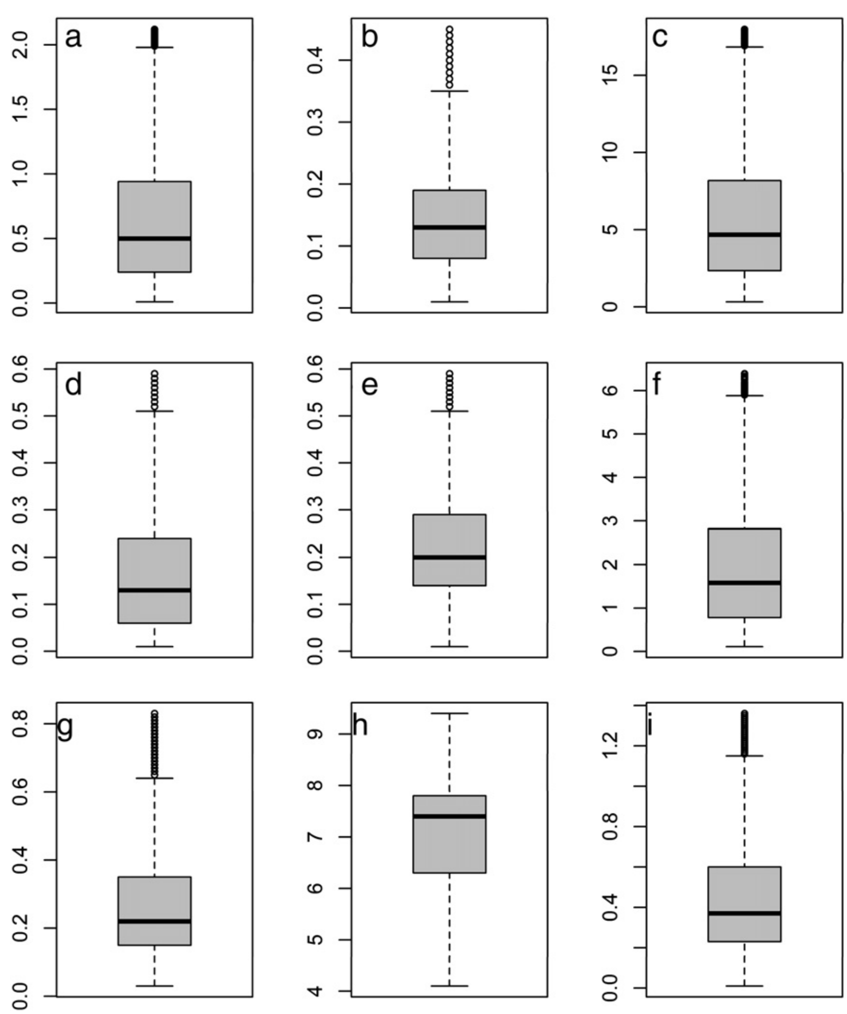

Fig 2. Boxplots of trimmed soil properties. a) Calcium (cmol L^{-1}); b) electrical conductivity (dS m^{-1}); c) organic carbon (kg m−2); d) potassium (cmol L^{-1}); e) magnesium (cmol L^{-1}); f) organic matter (%); g) sodium (cmol L^{-1}); h) pH; i) sodium absorption ratio.

Table 1
Descriptive statistics of the nine soil variables.

Variable	Coefficient of variation (%)	Raw data		$^4\sqrt{}$-transformed data	
		Skewness	Kurtosis	Skewness	Kurtosis
Ca (cmol L^{-1})	86.30	1.29	1.12	−0.04	−0.75
EC (dS m^{-1})	158.33	4.08	18.37	−0.16	−0.36
OC (kg m^{-2})	76.60	0.99	0.23	−0.18	−0.62
K (cmol L^{-1})	95.24	1.59	2.2	0.07	−0.79
Mg (cmol L^{-1})	85.19	3.2	14.58	−0.16	0.07
OM (%)	80.93	1.15	0.7	−0.14	−0.62
Na (cmol L^{-1})	116.67	2.7	7.69	−0.11	−0.19
pH	14.65	−0.82	−0.14	−0.78	−0.22
SAR (%)	98.28	2.33	6.28	−0.25	−0.28

EC: electrical conductivity; OC: organic carbon; OM: organic matter: SAR: sodium absorption ratio.

2. Materials and methods

2.1. Sampling and soil analyses

Random soil samples were taken from 4400 points (Ortiz-Solorio, 2002) of México's continental surface (the total surface of México is approximately 1,949,359 km^2) (Fig. 1). The samples were obtained as a composite of the first 20 cm of topsoil. The following nine variables for each point were determined (Ortiz-Solorio, 2002): electrical conductivity (EC), organic carbon (OC) in kg m^{-2}, four soluble cations (Ca, K, Mg, Na), pH, the sodium absorption ratio (SAR), and organic matter (OM).

2.2. Exploratory analysis

An exploratory analysis was performed for each variable to eliminate outliers, reduce data skews and fit the data to a normal distribution. The skewness and kurtosis were used to assess the fit of the data in a normal distribution, and when they did not fit, an appropriate transformation was applied based on the distribution of the data.

2.3. Variogram and spatial dependence

The first step in a geostatistical analysis is to determine the spatial dependence among the data for a particular variable. To do this, the semivariance (γ) is calculated by employing Eq. (1) (Goovaerts, 1999).

$$\gamma(h) = \frac{1}{2}n\sum_{i=1}^{n}[Z(x_i+h)-Z(x_i)]^2 \tag{1}$$

where x_i and $x_i + h$ are the sampled localities separated by a distance h; and $Z(x_i)$ and $Z(x_i + h)$ are Z values measured from their corresponding localities. When the semivariance is graphed against distance (h), an experimental variogram is obtained; this depends on three axes, two independent variables (direction and distance h) and one dependent variable (observation $Z(x_i)$) (Gassner and Schnug, 2008). The variogram can then be fitted using several models (spherical, exponential, Gaussian, linear, and logarithmic) (Gallardo, 2006). The model that shows the best adjustment to the experimental variogram is preferred. For each variable, an estimation based directly on the graphs can be performed, but this requires enough personal experience from the study area or the automatic use of the addition of the sum of squared error (SSE; Cressie, 1985).

To evaluate the spatial dependence, a nugget and sill relationship was used. A nugget:sill of 0.25 or less was considered evidence of a strong spatial dependence, a relationship between 0.25 and 0.75 was considered evidence of a moderate spatial dependence, and a relationship greater than 0.75 was considered to be evidence of a weak spatial dependence (Cambardella et al., 1994).

2.4. Assessment of kriging

Once the theoretical *ad hoc* model for each soil property was selected, a spatial inference was performed using the kriging technique. In this study, the ordinary, simple, and universal kriging models were evaluated with a 10-fold cross-validation. The data were divided into 10 subgroups; one of the subgroups was used as test data, and the rest were used as training data. The cross-validation process was repeated 10-fold with each of the possible data subgroups, and prediction errors were obtained (Pebesma, 2004). Using these prediction errors, the mean error (ME, Eq. (2)) was calculated; this value must be close to zero, the root mean squared error (RMSE; Eq. (3)) should be lower than the sampling variance, and the mean standardized prediction error (MSPE) should be close to zero (Eq. (4)).

$$\mathrm{ME} = \frac{1}{N}\sum_{i=1}^{N}[z(x_i)-\hat{z}(x_i)] \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[z(x_i)-\hat{z}(x_i)]^2} \tag{3}$$

$$\mathrm{MSPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\mathrm{ME}}{\sigma^2(x_i)} \tag{4}$$

where $\hat{z}(x_i)$ is the estimated value, $z(x_i)$ is the known value, N is the sampling size, and σ is the kriging variance of x_i (Kravchenko and Bullock, 1999). We chose the theoretical model with the lowest RMSE, MSPE, and ME.

2.5. Pixel size

The pixel size of the map outcome is important to adequately represent the spatial variability of the soil properties. The inspection density method was used to determine the pixel size. It consists of dividing the sampling size (N) by the extent of the study area (A). The respective calculations were performed using the following equations (Hengl, 2006):

$$\text{Coarsest legible resolution} \leq 0.1 \times \sqrt{\frac{A}{N}} \tag{5}$$

Table 2
Sum of squared error of variogram models (bold lower errors).

Variogram model	Ca	CE	CO	K	Mg	MO	Na	pH	RAS
Spherical	0.0011	0.00015	0.0013	0.0023	**0.00002**	0.0011	0.0005	0.000008	0.0008
Exponential	0.0012	**0.00012**	**0.0005**	**0.0012**	0.00004	**0.0004**	0.0006	**0.000004**	**0.0007**
Gaussian	0.0024	0.00025	0.0018	0.0023	0.00006	0.0011	0.0015	0.000009	0.0009
Linear	0.0015	0.00039	0.0028	0.0031	0.00003	0.0022	0.0005	0.000015	0.0008
Matern	0.0012	0.00013	0.0006	0.0013	0.00004	0.0005	0.0006	0.000005	0.0008
Bessel	0.0011	0.00013	0.0007	0.0015	0.00004	0.0005	0.0005	0.000005	0.0008
Pentaspherical	**0.0010**	0.00015	0.0012	0.0021	0.00003	0.0009	**0.0004**	0.000007	0.0008

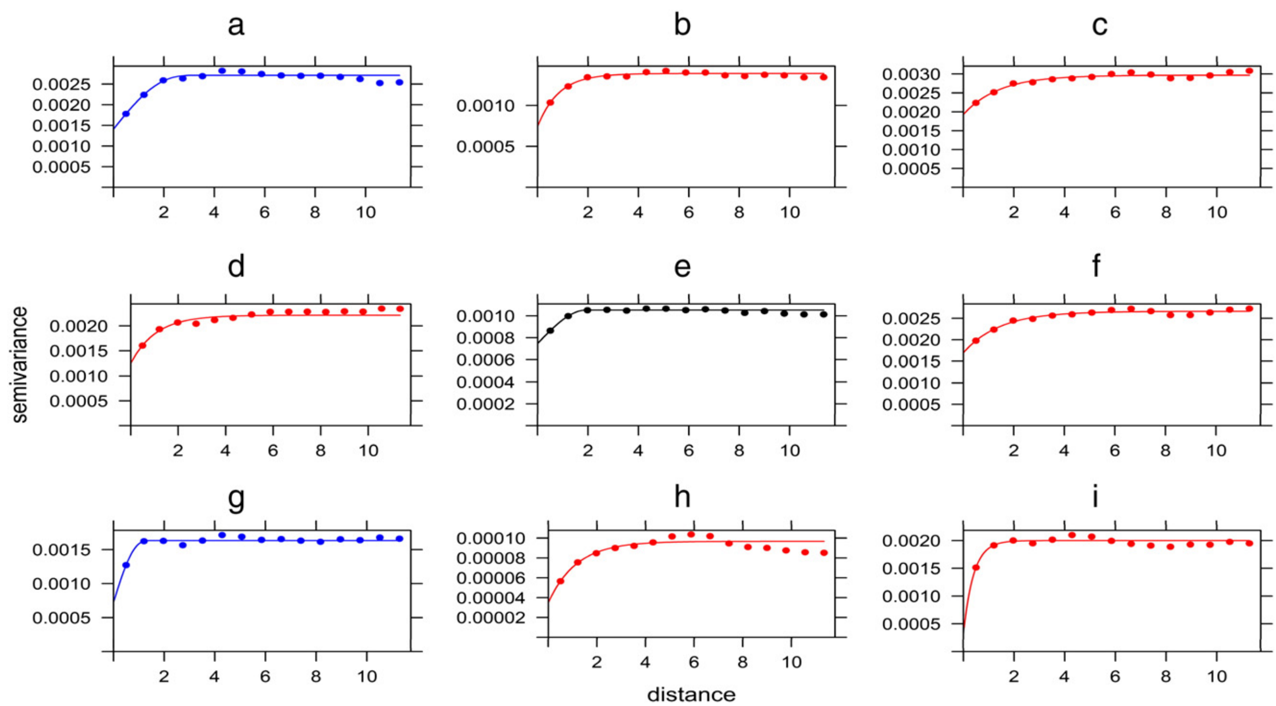

Fig. 3. Variograms selected for krigings. a) Calcium (cmol L^{-1}); b) electrical conductivity (dS m^{-1}); c) organic carbon (kg m–2); d) potassium (cmol L^{-1}); e) magnesium (cmol L^{-1}); f) organic matter (%); g) sodium (cmol L^{-1}); h) pH; i) sodium absorption ratio. Semivariance soil properties in fourth root and distance in degrees. Blue color, fit to pentaspherical model; red color, fit to the exponential model; black color, fit to the spherical model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\text{Finest legible resolution} \leq 0.05 \times \sqrt{\frac{A}{N}} \tag{6}$$

$$\text{Recommended compromise} \leq 0.0791 \times \sqrt{\frac{A}{N}}. \tag{7}$$

2.6. Map classes

To determine the interval and number of classes for each map (nine soil properties), a method proposed by Law et al. (2009) was used. The method considers the mean and standard deviation of the map's pixel values to determine the number of classes.

2.7. Software

A geostatistical analysis was performed using the GSTAT package (Pebesma, 2004) of the R software program (R Core Team, 2011). The final map editing was accomplished with ArcGIS 9™.

3. Results

3.1. Exploratory analysis

The exploratory analysis detected outliers in the soil properties data. All of the variables showed outliers beyond the upper limit of the interquartile range. The outliers were trimmed in each soil property datum and the results are shown in Fig. 2. The skewness and kurtosis were calculated with the trimmed data and showed non-normal distributions (Table 1). Four transformations were applied to the data, and the 4th root was the one that obtained the best fit for a normal distribution, as the skewness and kurtosis values were near zero. These previous analyses allowed us to reduce the nugget effect in the variograms of Ca, Mg, Na, and the SAR. The construction and adjustment of the variograms, as well as the validation tests and space prediction, were carried out with the transformed data.

3.2. Variograms and spatial dependence

The EC, OC, K, OM, pH, and SAR semivariances were fitted to an exponential model, the Ca and Na semivariances were fitted to a pentaspherical model, and the Mg semivariance was fitted to a spherical model (with the lower SSE of the variogram models; Table 2; Fig. 3). All of the soil property variables showed a moderate spatial dependence, with the exception of the SAR, which had a strong spatial dependence (Table 3; Fig. 3; Cambardella et al., 1994).

3.3. Kriging

Based on the cross-validation 10-fold analysis, the ordinary and universal kriging models were selected to perform the spatial predictions (Table 4). Ordinary kriging fitted better with the variables Ca, EC, OC, K, and pH than universal kriging (Table 4). Universal kriging was used

Table 3
Spatial dependence of soil properties.

Variable	Nugget:Sill	Spatial class
Ca	0.52	M
EC	0.54	M
OC	0.66	M
K	0.57	M
Mg	0.70	M
OM	0.65	M
Na	0.44	M
pH	0.33	M
SAR	0.16	S

EC = electrical conductivity; OC = organic carbon; OM = organic matter; SAR = sodium absorption ratio; M = moderate spatially dependent; S = strong spatially dependent.

Table 4
Cross-validation 10-fold statistics of krigings (bold lower errors) in nine soil variables.

Variable	Universal			Ordinary			Simple		
	ME	RMSE	MSPE	ME	RMSE	MSPE	ME	RMSE	MSPE
Ca	−0.0004	01351	0.0183	**-4.6×10^{-5}**	**0.1352**	**0.0183**	−0.0590	0.1685	0.0284
EC	0.0001	0.2347	0.0551	**0.0004**	**0.2342**	**0.0548**	−0.0645	0.2586	0.0669
OC	-7.5×10^{-5}	0.0796	0.0063	**0.0002**	**0.0794**	**0.0063**	−0.1152	0.1855	0.0344
K	2.1×10^{-5}	0.1015	0.0103	**0.0003**	**0.1015**	**0.0103**	−0.0708	0.1521	0.0231
Mg	**−0.0001**	**0.0828**	**0.0068**	0.0001	0.0830	0.0069	−0.3159	0.4097	0.1679
OM	**−0.0002**	**0.1806**	**0.0326**	0.0004	0.1802	0.0325	−0.0633	0.2088	0.0436
Na	**−0.0001**	**0.0952**	**0.0091**	0.0002	0.0955	0.0091	−0.2910	0.3994	0.1595
pH	-7.7×10^{-5}	0.0382	0.0015	**-6.2×10^{-5}**	**0.0382**	**0.0015**	−0.0237	0.0614	0.0038
SAR	**0.0003**	**0.1156**	**0.0134**	0.0005	0.1156	0.0134	−0.1604	0.2653	0.0704

EC: electrical conductivity; OC: organic carbon; OM: organic matter; SAR: sodium absorption ratio; ME: mean error; RSME: root mean squared error; MSPE: mean standardized prediction error.

for the spatial prediction of Mg, OM, Na, and the SAR because it showed the lowest error in the ME and MSPE (Table 4).

3.4. Pixel size definition

Based on Eqs. (5), (6), and (7), the recommended pixel size needed to generate layers for the nine analyzed soil properties is 1–2 km^2 (Table 5). We decided to use 1 km^2 pixels because most available online databases use this pixel size at the global, continental, regional, or national scales.

3.5. Soil property descriptions

Based on the mean and standard deviations of the pixel-layer values of the soil properties, maps were generated using the five interval classes. The first interval corresponds to the lowest category, the second one to the slightly lower category, the third one to the intermediate category, the fourth one to the slightly higher category, and the fifth one to the highest category. This classification was based on statistical approaches that do not always agree with the established classifications at a national or global level, except for the organic carbon and pH. The lowest and highest categories don't mean they are faulty or abundant in the content of a particular soil property; the aim was to have categories based in statistical criteria. The soils with a slightly low or intermediate Ca content occupy more than 80% of the Mexican territory sampled (Fig. 4a). The soils with a high Ca content are located primarily in arid and semi-arid zones, where the dominant parent material is limestone (Fig. 5a).

In general, the soils were not considered saline because the EC is lower than 2 dS m^{-1}; however, based on the proposed categories, the soils were classified as slightly low in EC, and only 8% of the soils showed high EC values (Fig. 4b). These latter soils were found in the arid and semi-arid climates where limestone was the parent material, and their distribution is restricted to northern México (Fig. 5b).

Soils with slightly low and medium OC density occupy 67% of Mexican territory (Fig. 4c). The soils with medium or slightly high OC density were distributed in areas where dense vegetation cover is present (Fig. 5c), which causes organic waste material accumulation.

The soils with slightly low and medium K content were distributed over more than 60% of the country (Fig. 4d). The soils with high K content were found in the arid zones of northern México (Fig. 5d). In contrast, the Mg distribution patterns in México showed no defined tendency (Fig. 5e); therefore, slightly low and medium were the main classes for this cation (Fig. 4e).

The organic matter amount (Fig. 4f) and distribution (Fig. 5f) were similar to the OC; however, their classes are less continuous than those of OC because their distribution is broken up by other classes. For example, inside the slightly high class, several high-class areas can be found.

The Na slightly low and medium classes were distributed throughout the country and showed no definite pattern (Fig. 5g). On the contrary, the high class was found mostly in the Baja California and Chihuahua deserts and along the coastal plains of Sinaloa and Sonora.

The soil pH in México is primarily slightly alkaline and neutral (Fig. 4h). Soils with a slightly high pH or alkaline soils are found in arid zones (Fig. 5h), whereas acid soils with a low pH are mostly found in temperate and tropical forests.

In general, Mexican soils can be classified as showing a medium sodium absorption ratio (SAR) content (Fig. 4i). The soils with high SAR content are found mostly in the Baja California peninsula (Fig. 5i).

4. Discussion

In México, the pH coefficient of variation (CV) is classified as low, but the CV for electrical conductivity and K are high. The pH CV fluctuates between 2 and 15%, mostly as a result of its intrinsic genesis characteristic (Cristobal et al., 2008). The electrical conductivity CV interval varies from 91 to 263%, and that of K oscillates between 39 and 157%. These ranges coincide with those found by Mulla and McBratney (2002). The moderate spatial dependence ratio class of pH, EC, OM, and SAR coincide with the results found by Ayoubi et al. (2007). They mention that this spatial dependence might be influenced by the mobility of nutrients. Those that have a strong spatial dependence are more mobile than those with a moderate spatial dependence. Based on this assumption, Mg would be the least mobile, and Na would have the greatest mobility.

The selected theoretical models for the EC, K, OM, and pH agree with those found in previous research (Kravchenko and Bullock, 1999; Mueller et al., 2001; Robinson and Metternicht, 2006). The range parameter of the variogram can be used to determine an optimum sampling space distribution with the aim of obtaining an adequate spatial

Table 5
Pixel sizes used to generate soil property maps.

Variable	Sampling size	Pixel size (km^2)		
		Coarse	Finest	Best
Ca	4215	2.20	1.07	1.70
EC	4485	2.08	1.04	1.65
OC	4249	2.14	1.07	1.69
K	4432	2.09	1.04	1.65
Mg	4400	2.10	1.05	1.66
OM	4373	2.11	1.05	1.67
Na	4342	2.11	1.05	1.67
pH	4576	2.06	1.03	1.63
SAR	4301	2.12	1.06	1.68

EC: electrical conductivity; OC: organic carbon; OM: organic matter: SAR: sodium absorption ratio.

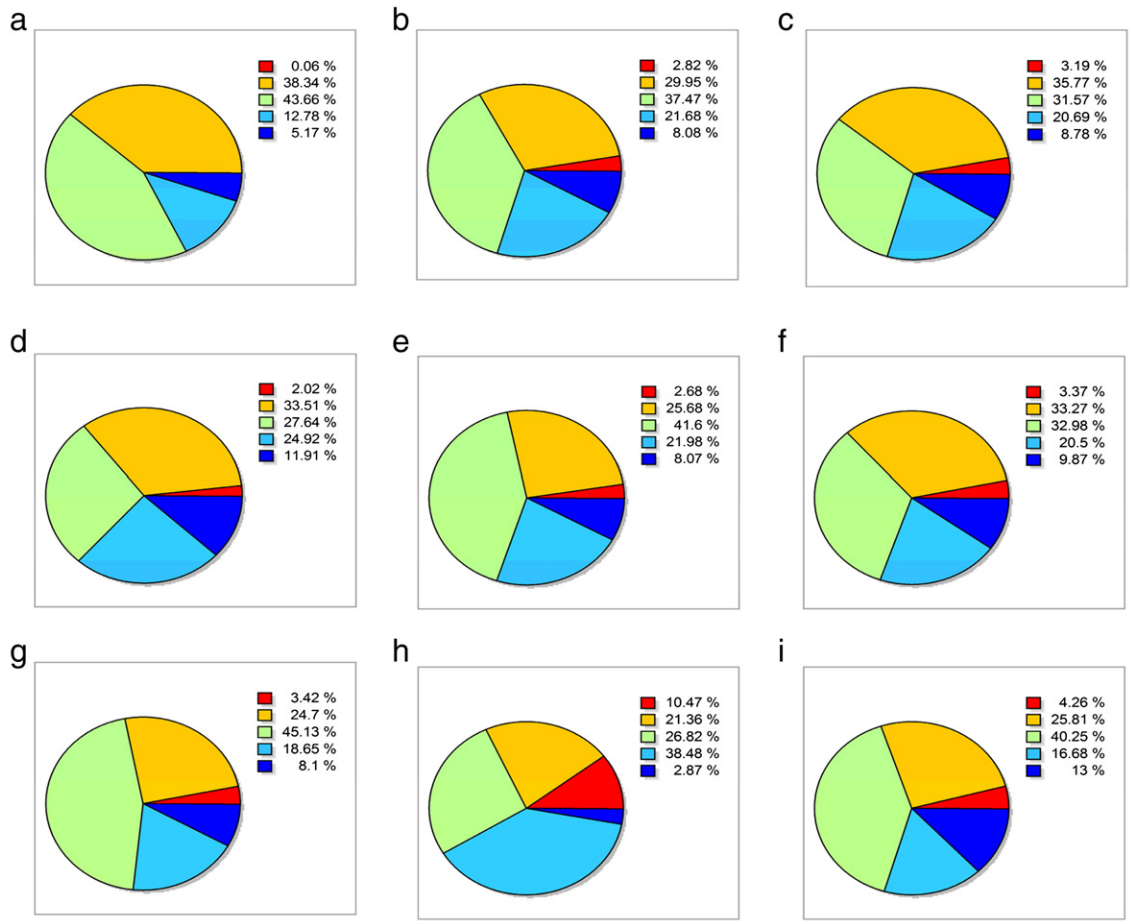

Fig. 4. Pie graph of soil properties. a) Calcium (cmol L^{-1}); b) electrical conductivity (dS m^{-1}); c) organic carbon (kg m–2); d) potassium (cmol L^{-1}); e) magnesium (cmol L^{-1}); f) organic matter (%); g) sodium (cmol L^{-1}); h) pH; i) sodium absorption ratio.

pattern for a particular property, and values from 1/4 to 1/2 of the range should be used as the reference (Mulla and McBratney, 1999). For this study, a rank of 1/4 was selected, due to the extremely heterogeneous topography of the country. Based on these decisions, a spatial sampling separation of 88, 23, 39, 29, 53, 39, 39, 32, and 11 km was recommended for Ca, EC, K, Mg, OC, OM, Na, pH, and the SAR, respectively, at the national level.

The selected kriging models for spatial prediction were different, depending on the soil property to be modeled. The errors obtained from the cross-validation test are similar to those found within the ranges reported in previous studies. These values were below 1 (Robinson and Metternicht, 2006; Scholoeder et al., 2001).

Five out of nine soil properties (Ca, EC, K, OM, and pH) showed distribution patterns that were determined primarily by two factors influencing soil formation: climate and parent material. In general, the soil properties evaluated were classified as slightly low and medium.

According to the results reported by the Global Soil Data Task Group (2000), the organic carbon density in México is similar to that obtained in this study, where the soils have an organic carbon density from low to slight. With respect to the pH, a similar distribution pattern is shown, with predominantly medium alkaline to neutral soils. However, based on the proposed classification, the soils in México have a slightly low to medium SAR with the agronomic criteria considered as a low SAR level (<6; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012).

The nine soil property layers generated in this study are applicable in different scientific fields, for example, agricultural or environmental sciences, or as additional variables to be used in species-distribution modeling. Their usefulness for modeling the geographic distribution of the cloud forest in México, the most important biodiversity biome, has already been evaluated (Cruz-Cárdenas et al., 2012; López-Mata et al., 2012; Villaseñor, 2010). For example, organic carbon in the soil, together with precipitation and temperature, defines the richness and potential distribution of species in this biome. Coudun et al. (2006) found that soil properties improved a logistical model of *Acer campestre* by reducing the "explained deviance"; among these, the pH remained in the final model. These two soil variables have also been used to model flora hotspots with the aim of outlining priority zones for the conservation and planning of ecosystems (Zhang et al., 2012).

Future work should evaluate the semivariance anisotropy in addition to those evaluated in this paper. Likewise, additional kriging techniques, such as block kriging, indicator kriging, and co-kriging should be evaluated because they could better adjust the data for spatial predictions.

Acknowledgments

The first author thanks the UNAM (Universidad Nacional Autónoma de México) for the fellowship granted in its DGAPA (Dirección General de Asuntos del Personal Académico) 2010 Postdoctoral Fellowships Program. The Biology Institute of UNAM provided human and material resources to carry out this research.

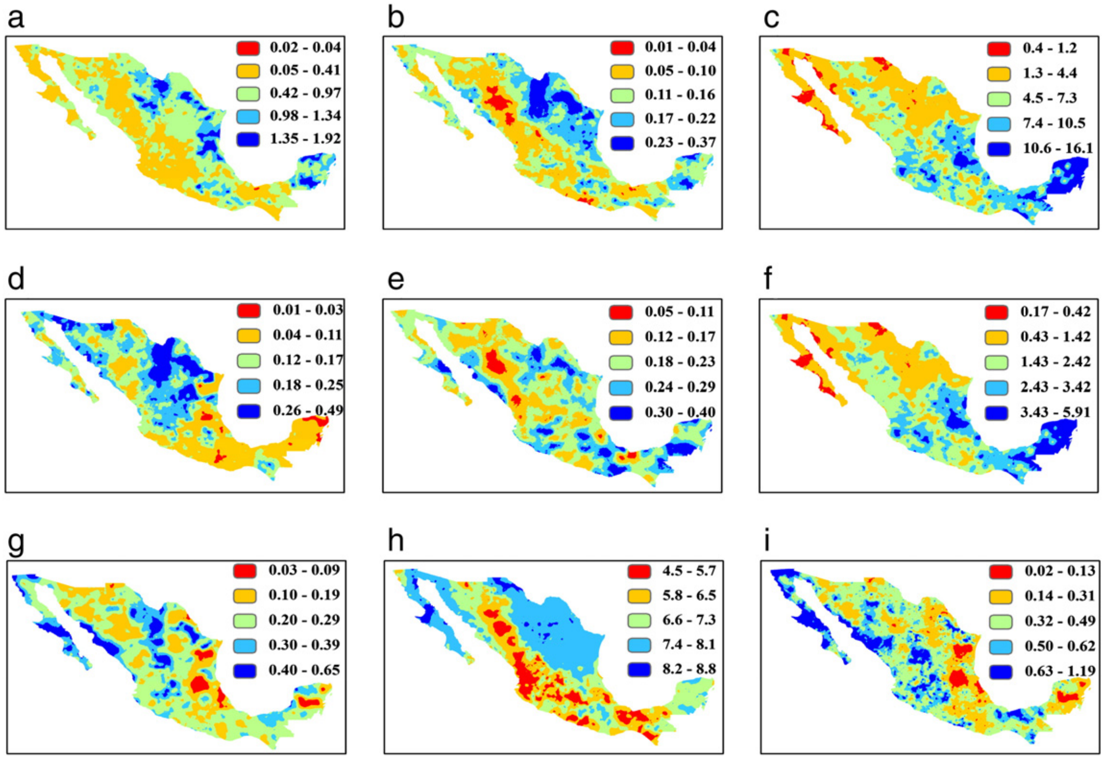

Fig. 5. Soil map properties. a) Calcium (cmol L^{-1}); b) electrical conductivity (dS m^{-1}); c) organic carbon; d) potassium (cmol L^{-1}); e) magnesium (cmol L^{-1}); f) organic matter (%); g) sodium (cmol L^{-1}); h) pH; i) sodium absorption ratio.

References

Ayoubi, Sh., Zamani, S.M., Khormali, F., 2007. Spatial variability of some soil properties for site specific farming in northern Iran. International Journal of Plant Production 1, 225–236.

Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F., Konopka, A.E., 1994. Field-scale variability of soil properties in Central Iowa soils. Soil Science Society of America Journal 58, 1501–1511.

Coudun, C., Gégout, J.C., Piedallu, C., Rameau, J.C., 2006. Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. Journal of Biogeography 33, 1750–1763.

Cressie, N., 1985. Fitting variogram models by weighted least squares. Journal of the International Association for Mathematical Geology 17, 563–586.

Cristobal, A.D., Álvarez, S.M.E., Hernandez, A.E., Maldonado, T.R., Pérez, G.M., Castro, B.R., 2008. Variabilidad espacial de propiedades químicas del suelo y su uso en el diseño de experimentos. Terra Latinoamericana 26, 317–324.

Cruz-Cárdenas, G., Villaseñor, J.L., López-Mata, L., Ortiz, E., 2012. Potential distribution of the Humid Mountain Forest in Mexico. Botanical Sciences 90, 331–340.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, G., Nakamura, M., Nakazawa, Y., Overton, J.M.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012. Harmonized World Soil Database (Version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Gallardo, A., 2006. Geoestadística. Ecosistemas 3, 49–59.

Gassner, A., Schnug, E., 2008. Geostatistics for soil science. In: Rattan, L. (Ed.), Encyclopedia of Soil Science. CRC, Boca Raton, FL, pp. 760–764.

Global Soil Data Task Group, 2000. Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS). [Global Gridded Surfaces of Selected Soil Characteristics (International Geosphere–Biosphere Programme—Data and Information System)]. Data set. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. http://dx.doi.org/10.3334/ORNLDAAC/569.

Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89, 1–45.

Hansen, M., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km resolution using a decision tree classifier. International Journal of Remote Sensing 21, 1331–1365.

Hengl, T., 2006. Finding the right pixel size. Computers and Geosciences 32, 1283–1298.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-Kriging: from equations to case studies. Computers and Geosciences 33, 1301–1315.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965–1978.

Kravchenko, A., Bullock, D.G., 1999. A comparative study of interpolation methods for mapping soil properties. Journal of Agronomy 91, 393–400.

Law, M.C., Balasundram, S.K., Husni, M.H.A., Ahmed, O.H., Harun, M.H., 2009. Spatial variability of soil organic carbon in oil palm. International Journal of Soil Science 4, 93–103.

López-Mata, L., Villaseñor, J.L., Cruz-Cárdenas, G., Ortiz, E., Ortiz-Solorio, C.A., 2012. Predictores ambientales de la riqueza de especies de plantas del bosque húmedo de montaña en México. Botanical Sciences 90, 27–36.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52.

Mueller, T.G., Pierce, F.J., Schabenberger, O., Warncke, D.D., 2001. Map quality for site-specific fertility management. Soil Science Society of America Journal 65, 1547–1558.

Mulla, D.J., McBratney, A.B., 1999. Soil spatial variability. In: Summer, M. (Ed.), Handbook of Soil Science. CRC, Boca Raton, FL, pp. A321–A352.

Mulla, D.J., McBratney, A.B., 2002. Soil spatial variability. In: Warrick, A. (Ed.), Soil Physics Companion. CRC, Boca Raton, FL, pp. 343–373.

Ortiz-Solorio, C.A., 2002. Evaluación de la degradación del suelo causada por el hombre en la República Mexicana, escala 1:250 000. Secretaría del Medio Ambiente y Recursos Naturales, Colegio de Postgraduados, México.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers and Geosciences 30, 683–691.

R Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, T.P., Metternicht, G., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. Computers and Electronics in Agriculture 50, 97–108.

Rossi, R.E., Mula, D.J., Journel, A.G., Franz, E.H., 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 2, 277–314.

Scholoeder, C.A., Zimmerman, N.E., Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Science Society of America Journal 65, 470–479.

Villaseñor, J.L., 2010. El bosque húmedo de montaña en México y sus plantas vasculares: catálogo florístico-taxonómico. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Universidad Nacional Autónoma de México, México, D.F.

Wagner, H.H., 2003. Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. Ecology 84, 1045–1057.

Zhang, M., Zhou, Z., Chen, W., Slik, J.W.F., Cannon, C.H., Raes, N., 2012. Using species distribution modeling to improve conservation and land use planning of Yunnan, China. Biological Conservation 153, 257–264.