Geoderma 290 (2017) 75–82

Contents lists available at ScienceDirect

Geoderma

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

Sampling scheme optimization to map soil depth to petrocalcic horizon at field scale

Marisa Beatriz Domenech [a,*], Mauricio Castro-Franco [b], José Luis Costa [c], Nilda Mabel Amiotti [d]

[a] *Instituto Nacional de Tecnología Agropecuaria INTA, CEI Barrow. Ruta 3 Km 488, 7500 Tres Arroyos, Argentina*
[b] *Consejo Nacional de Investigaciones Científicas y Técnicas CONICET. Av. Rivadavia 1917, C1033AAJ Buenos Aires, Argentina*
[c] *Instituto Nacional de Tecnología Agropecuaria INTA, EEA Balcarce. Ruta 226. Km 73.5, Balcarce, Argentina*
[d] *Departamento de Agronomía, Universidad Nacional del Sur, Bahía Blanca. CERZOS-CONICET, Argentina*

ARTICLE INFO

Article history:
Received 14 April 2016
Received in revised form 2 December 2016
Accepted 13 December 2016
Available online xxxx

Keywords:
Conditioned Latin hypercube
Digital soil mapping
Precision agriculture
Argentina
Ordinary cokriging

ABSTRACT

Soil depth has played a key role in the development of soil survey, implementation of soil-specific management and validation of hydrological models. Generally, soil depth at field scale is difficult to map due to complex interactions of factors of soil formation at field scale. As a result, the conventional sampling schemes to map soil depth are generally laborious, time consuming and expensive. In this study, we presented, tested and evaluated a method to optimize the sampling scheme to map soil depth to petrocalcic horizon at field scale. The method was tested with real data at four agricultural fields localized in the southeast Pampas plain of Argentina. The purpose of the method was to minimize the sample dataset size to map soil depth to petrocalcic horizon based on ordinary cokriging, five calibration sample sizes (returned by Conditioned Latin hypercube –cLHS-), and apparent electrical conductivity (ECa) or elevation as variables of auxiliary information.

The results suggest that (i) only 30% of samples collected on a 30-m grid are required to provide high prediction accuracy ($R^2 > 0.95$) to map soil depth to petrocalcic horizon; (ii) an independent validation dataset based on 50% of the samples on a 30-m grid is adequate to validate the most realistic accuracy estimate; and (iii) ECa and elevation, as variables of auxiliary information, are sufficient to map soil depth to petrocalcic horizon. The method proposed provides a significant improvement over conventional to map soil depth and allows reducing cost, time and field labour. Extrapolation of the results to other areas needs to be tested.

1. Introduction

The spatial distribution of soil depth affects the spatial dynamic of water storage capacity, runoff generation, subsurface flow, nutrients availability and crops yield (Stieglitz et al., 2003). For that reason, soil depth has played a key role in the development of soil survey, implementation of soil-specific management and validation of hydrological models (Tesfa et al., 2009). However, the spatial dynamic of soil depth at field scale is difficult to predict due to that the complex interactions of factors of soil formation at field scale such as topography, climate, parent material and land use (Jenny, 1941; Tesfa et al., 2010). As a result, the conventional sampling schemes to map soil depth are generally laborious, time consuming and expensive. Evidently, accurate and inexpensive sampling schemes are needed to map soil depth at field scale.

Terrain attributes obtained from digital elevation models and proximal soil sensors data are sources of inexpensive auxiliary information that have been used to map soil depth. For example, Tesfa et al. (2009) reported statistical models to map soil depth based upon the relationship between soil depth and terrain attributes. Ziadat (2010) reported that the modelling depth soil-landscape relationships using terrain attributes was a promising approach to map soil depth. On the other hand, Boettinger et al. (1997) and Bork et al. (1998) determined that electromagnetic induction data are potentially a powerful, inexpensive and quick tool to map soil depth. These examples suggest that terrain attributes and proximal soil sensor data are optimal sources of auxiliary information to map soil depth. However, there is little consensus on the optimal sampling scheme to map soil depth, especially where spatial soil depth pattern is highly variable.

The availability of auxiliary information is important to optimize sampling schemes (Hengl et al., 2004; Minasny and McBratney, 2006; Shaner et al., 2008) and to serve as ancillary variable in the local prediction of a soil property when using hybrid interpolation techniques such as cokriging (Vašát et al., 2010). This interpolation technique is used in cases where there are two or more spatially interdependent variables and incorporates those interdependent variables into spatial interpolation to obtain high prediction accuracy with limited sample data (Wang et al., 2013). Generally, cokriging needs two previous processes to improve prediction accuracy. The first process is a selection of the most important variables of auxiliary information characterized by high

* Corresponding author at: Ruta 3 Km 488 CC 50, 7500 Tres Arroyos, Argentina.
E-mail address: domenech.marisa@inta.gob.ar (M.B. Domenech).

http://dx.doi.org/10.1016/j.geoderma.2016.12.012
0016-7061/

interdependence with the variable to predict. At this respect, Behrens et al. (2010) proposed that using a variables selection technique based on Random Forest (RF), could help to reduce prediction model complexity while decreasing computation time and improving prediction accuracy. The second process is a selection of a model-based sampling scheme that allows quantifying of the spatial dependence and provide good area coverage for reliable prediction (Simbahan and Dobermann, 2006). According to that, several studies of digital soil mapping (DSM) have demonstrated that Conditioned Latin Hypercube sampling (cLHS) (Castro Franco et al., 2015; Minasny and McBratney, 2006; Mulder et al., 2013) could help to minimize the variance of the prediction error of geostatistical interpolation, with limited sample data. Although cokriging, RF and cLHS are being successfully applied as prediction models of several soil properties, their potential to optimize the sampling scheme to map soil-depth at field scale has been underexplored due to their novelty.

The southeast Pampas plain of Argentina, one of the most important cropping regions of the world, have about four million hectares that are underlain by a petrocalcic horizon which limits the soil depth (Pazos and Mestelan, 2002). Consequently, soil depth is the key factor that limits crop yield (Sadras and Calviño, 2001). At present, expensive and laborious sampling schemes are used to map soil depth at field scale. However, most of agricultural fields in the southeast Pampas plain of Argentina have wide availability of inexpensive auxiliary information because precision agriculture technologies have been rapidly adopted in the last decades (Swinton and Lowenberg-Deboer, 2001). In this context, the potential use of this auxiliary information to optimize the sampling schemes to map soil-depth at field scale requires to be evaluated and quantified.

The objective of this study was to present, test and evaluate a method to optimize the sampling scheme to map soil depth to petrocalcic horizon at field scale, based on inexpensive auxiliary information, RF as algorithm of importance variables selection and cLHS as model-based sampling scheme. The integration of these algorithms offers a new approach to optimize the sampling scheme, to identify the most important variables of auxiliary information and to overcome the limitations of conventional methods. Also, the parameterization of cokriging, RF and cLHS is very simple and computationally slighter than other algorithms.

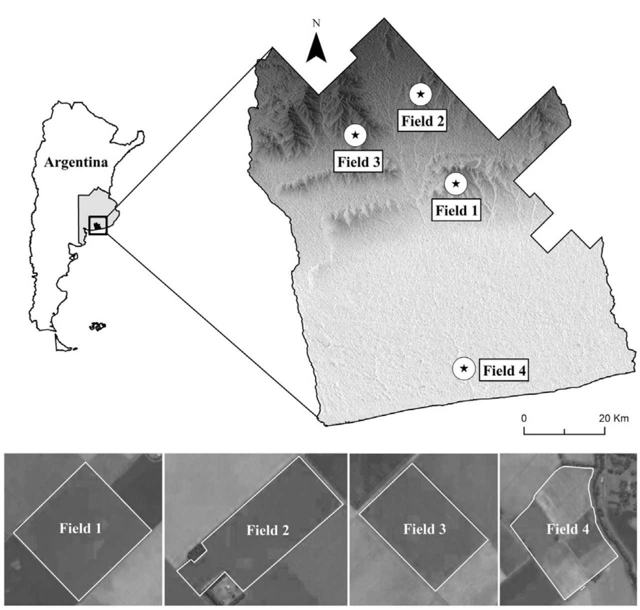

Fig. 1. Location of the study fields in the southeast Pampas of Argentina.

2. Materials and methods

2.1. Agricultural fields

The location of the fields used in this study is shown in Fig. 1. These fields were selected because they represent the variability of elevation, landscape position and spatial variability of soil depth usually found in the southeastern Pampas. The current crop rotations in all fields include corn, soybean or sunflower in summer and wheat or barley in winter (Costa et al., 2015). Specifically, the fields are located in the geological province locally termed “Sierras Septentrionales” in the southeast of Buenos Aires province of Argentina. In this zone, the loess deposits are from the Late Holocene and Pleistocene (Blanco and Stoops, 2007). The soils are classified as Subgroups Typic Argiudoll and Petrocalcic Argiudoll; Family fine, illitic, thermic (Soil Survey Staff, 2014). Table 1 shows the area and composition of soil mapping unit for each field.

The southeastern Pampas plain of Argentina has a frost-free period that extends from October to May. The mean annual temperature is 14.8 °C. It has a humid and subhumid hydric regime (Thornthwaite, 1948). The mean annual precipitation is about 756 mm. The rain regime is (i) rainy from October to March, (ii) moderately rainy in April, May and September, and (iii) scarcely rainy from June to August (Costa et al., 2015).

2.2. Auxiliary information measurement

ECa and elevation were used as auxiliary information to optimize sampling to map soil depth at field scale.

ECa measurements were collected at two different dates (July 18th, 2008 in Field 4 and June 23rd–30th, 2011 in Fields 1, 2 and 3) using a Veris® 3100 soil electrical conductivity sensor (Veris Technologies Inc., Salina, KS, USA). The accumulated rainfall reached 612 mm from January to July 2008, whereas only 211 mm were accumulated from December to June 2011. However, a rainfall of 8.5 mm occurred on June 22nd 2011. Precipitation data were provided by Agrometeorological Department of National Institute for Agricultural Technology of Argentina (INTA-CEI Barrow) from the nearest weather recording station for each field.

The coulter electrodes of the Veris® 3100 are configured as a Wenner array, an arrangement commonly used for geophysical resistivity surveys. In this sensor, the system records ECa in mS m^{-1} by electrical resistivity at a shallow depth (0–30 cm, ECa_30 cm) and a deep depth (0–90 cm, ECa_90 cm) (Moral et al., 2010). Veris® 3100 was pulled through the field by a pick-up truck. ECa measurements were made along parallel transects approximately 20 m apart on the surface of each agricultural field. An advance GPS Surveying instrument GPS Trimble® GeoXT™ handheld with submeter accuracy was used to georeferenced the ECa measurements. Latitude, longitude, ECa_30 cm and ECa_90 cm data were recorded in an ASCII text file and transferred to GIS software for further analysis. For more details of ECa measurements with Veris 3100® see Corwin and Lesch (2003), Corwin and Lesch (2005) and Allred et al. (2008).

Elevation was measured simultaneously with ECa, using an advance differential GPS Surveying instrument GPS Trimble®R3 (Trimble Navigation Limited, CA, USA), which is equipped with a GPS receiver, antenna and rugged handheld controller. Elevation data were post-processed with Trimble Business Center software V3.5 to produce a digital elevation model of spatial resolution of 10 m, in each field.

Experimental variograms were computed to describe the spatial variation of ECa and elevation following the procedure proposed by Diggle and Ribeiro (2007).

The adjusted experimental variogram was used to interpolate ECa and elevation by ordinary kriging in each field. The R package “geoR” was used to conduct the geostatistical interpolation (R Development Core Team, 2015). Finally, a 10 × 10 m grid square size was chosen for output maps.