Geoderma 322 (2018) 101–111

Contents lists available at ScienceDirect

Geoderma

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

A pedometric technique to delimitate soil-specific zones at field scale

Mauricio Castro-Franco[a,b,*], Mariano A. Córdoba[a,c], Mónica G. Balzarini[a,c], José Luis Costa[d]

[a] Consejo Nacional de Investigaciones Científicas y Técnicas CONICET, Av. Rivadavia 1917, C1033AAJ, Buenos Aires, Argentina
[b] Instituto Nacional de Tecnología Agropecuaria INTA, CEI Barrow, Ruta 3 km 488, 7500 Tres Arroyos, Argentina
[c] Catedra de Estadística y Biometría, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ing Agr. Felix Aldo Marrone 746, Ciudad Universitaria, C.C. 509, 5000 Cordoba, Argentina
[d] Instituto Nacional de Tecnología Agropecuaria INTA, EEA Balcarce, Ruta 226 km 73.5, Balcarce, Argentina

ARTICLE INFO

Handling Editor: A.B. McBratney

Keywords:
Argentina
Digital soil mapping
Random forest
Spatial principal components
Precision agriculture

ABSTRACT

Delimitation of soil types within a farm field is key for site-specific crop management. An alternative to this, is to develop pedometric techniques that allow an efficient combination of soil survey information and high-resolution terrain attribute data. The aim of this study was to present and evaluate a pedometric technique to delimit soil-specific zones at field scale by coupled Random forest, fuzzy k-means clustering and spatial principal components algorithms (RF-KM-sPCA) and by using information from soil surveys and terrain attributes derived from a digital elevation model. The protocol involves three-steps: 1) automatic classification of small (20x20m) spatial units (SU) using the knowledge of the soil map units present in the farm landscape, 2) aggregation of SUM at farm scale and 3) validation of soil-specific zones. For the first step, we used the random forest algorithm with 10 terrain attributes. For the second step, KM-sPCA algorithms were used to cluster within field SU accounting for autocorrelation. For the third step, apparent soil electrical conductivity and yield maps was used to validate the delimitation of soil-specific zones. This technique produced more contiguous zones than other cluster methods which do not use spatiality. Six farm fields with highly differences in soils were partitioned by the proposed pedometric strategy. Apparent soil electrical conductivity and yield maps present significant differences among zones in all experimental fields. This analytic strategy, based in easy-to-obtain data, could be used to improve precision agricultural managements.

1. Introduction

Soil properties that limit crop yield within agricultural fields often vary considerably over space and time (Castro Franco et al., 2015; Gebbers and Adamchuk, 2010). Usually, this variability is intentionally ignored in soil sampling schemes, laboratory analyses and agronomic strategies for crop management. Hence, it appears that applying strategies for soil-specific conditions in the context of precision agriculture would have the potential to improve the way in which soils are currently managed. To achieve this, the method used to delimit the complexity of soils of agricultural fields in subareas according to the soil type should be simplified, so that these subareas could be individually controlled with respect to management decisions (Fraisse et al., 2001; Johnson et al., 2001).

Pedometrics techniques is the application of mathematical and statistical models to study the distribution and genesis of soils (Rossiter, 2012), which within the context of digital soil mapping (McBratney et al., 2003), could be useful to define soil-specific zones in agricultural fields. Generally, there are three pedometric approaches. The first one, known as disaggregation of soil map units (DgSMU), allows delimiting soil-specific zones at different scales by combining information obtained from conventional soil surveys with information obtained from digital soil mapping (Bui and Moran, 2001). The second one estimates the spatial distribution of soil properties by using geostatistical methods (Hempel et al., 2008). The main disadvantage of this approach, is that a huge number of soil samples have to be collected and analyzed to adequately represent the soil spatial variability (Fraisse et al., 2001). The third approach estimates soil spatial patterns through machine learning algorithms by using ancillary data such as apparent soil electrical conductivity, remote sensing, and digital elevation models (DEM) (Ahmad et al., 2010; Castro Franco et al., 2015; Nitze et al., 2012; Scull et al., 2003).

In South America, the first and third approaches have had a great potential to generate useful cartography to be implemented in soil-specific management strategies. This is because in this region several soil surveys are available, which can be disaggregated using digital soil mapping tecniques (Pennock et al., 2015; Sanchez et al., 2009); Spatially, this conventional survey is formed by polygons or Soil Map Units

* Corresponding author at: Ruta 3 Km 488 CC 50, 7500 Tres Arroyos, Argentina.
E-mail address: castrof.mauricio@inta.gob.ar (M. Castro-Franco).

https://doi.org/10.1016/j.geoderma.2018.02.034
Received 20 February 2017; Received in revised form 30 January 2018; Accepted 21 February 2018
0016-7061/

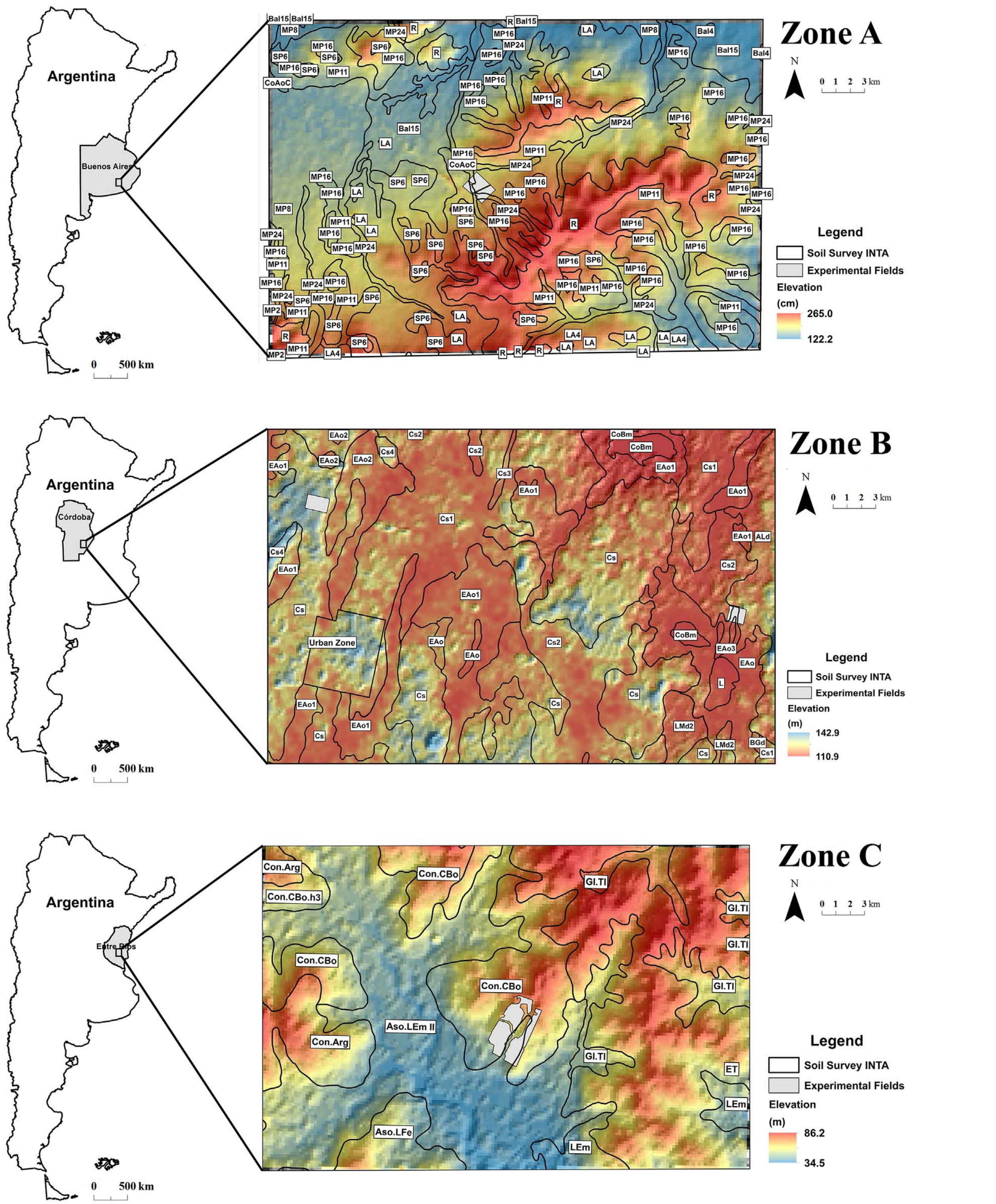

Fig. 1. Mapping of Soil Map-units (SMUs) according to INTA soil survey at scale 1:50,000 for each agricultural zone, Argentina. Spatial distribution of elevation from digital elevation model (MDE-Ar) (back)

(SMU) according to their soil-landscape relationships (Jenny, 1941; McBratney et al., 2003). Each SMU represents the "*aggregation*" of a number of soil series which are identified by their spatial correspondence; hence, each SMU is considered as a spatial generalization that can be disaggregated (Nauman and Thompson, 2014). Also, multiple ancillary information is available, which can be used to classify SMUs from machine learning algorithms (Brungard et al., 2015; Heung et al., 2016; Massawe et al., n.d.). Generally, these algorithms are used to determine the spatial correlation among SMUs and ancillary information of environmental data derived from DEM, remote sensing, and soil sensing, in order to develop a training dataset (McBratney et al., 2003). The learning relationships between SMU and environmental data are adjusted in a model which is then applied in the validation procedure. Within machine learning, Random Forest (RF) is an outstanding algorithm (Gambill et al., 2016). The RF technique is an ensemble learning technique which generates many classification trees that are aggregated