Geoderma xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Geoderma

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

Homogenisation of a soil properties map by principal component analysis to define index agricultural insurance policies

Omar Valverde Arias [a,*], Alberto Garrido [a,b], María Villeta [c], Ana María Tarquis [a,d]

[a] CEIGRAM, ETSIAAB, Universidad Politécnica de Madrid (UPM), Practice field of School Agricultural Engineering, Senda del Rey St. 13, 28040 Madrid, Spain
[b] Department of Agricultural Economics and Social Sciences, ETSIAAB, UPM, Ciudad Universitaria s.n., 28040 Madrid, Spain
[c] Department of Statistics and Operation Research III, Universidad Complutense de Madrid. Av. Puerta del Hierro, 1, 28040 Madrid, Spain
[d] Department of Applied Mathematics, ETSIAAB, UPM, Ciudad Universitaria s.n., 28040 Madrid, Spain

A R T I C L E I N F O

Article history:
Received 17 March 2016
Received in revised form 21 December 2016
Accepted 16 January 2017
Available online xxxx

Keywords:
Soil characteristics
Homogeneous areas
Classification
Stratify sampling

A B S T R A C T

Extreme climatic events are a serious concern for agriculture and its related activities in the entire world. Some recent studies have shown that there is a change in seasonal patterns and an increasing frequency of extreme climatic events, adding higher risk on farming activities. Therefore, government institutions should implement agricultural risk management policies or enable the private sector to develop appropriate tools.

One of the main strategies for transferring agricultural risk is crop insurance. An alternative for traditional agricultural insurance is insurance based on easily measurable indicators, which use a predetermined index value as the basis for defining the indemnities. However, index insurance cannot be based only on isolated measurements. It should also be integrated into a complete monitoring system that uses many sources of information and tools (e.g., index influence areas, crop production risk maps, crop yields, and insurance claims statistics).

To establish index influence areas, it is necessary to have secondary information indicating the type of climate and soil homogeneity of the study area. Over a homogeneous area, index measurements on crops of interest will be similar, but differentiating index values for areas with different soil and climatic characteristics will reduce basis risk.

This study assesses two conventional agricultural and geographic methods (control and climatic maps) based on expert criteria. They were compared with one statistical method of multi-factorial analysis (factorial map). All these methods claim to homogenise soil and climatic characteristics.

The three resulting maps were evaluated by agricultural and spatial analysis. The factorial map showed more homogeneous classes than the climatic map but the later lost all the soil characteristics information that will influence the index value variations. On the other hand, the factorial map obtained fewer classes than the control map, although retaining the main information on soil variability in the study area. These results obtained from the statistical method (factorial map) demonstrate that this method has generalised efficiently climatic, topographic and soil characteristics of the complete dataset.

1. Introduction

The use of agricultural insurance as a means to manage and transfer agricultural risk is expanding around the world (Hoag, 2009). Climatic or vegetation indices can be effective tools to estimate crop yield reductions resulting from extreme events, such as floods, frost, or droughts. The most commonly used vegetation index is the normalised difference vegetation index (NDVI), which regularly exhibits a high correlation with crop conditions. The NDVI permits establishing a relationship between crop damage and its effects on lost production (Carlson and Ripley, 1997). Index insurance pays indemnities not based on losses verified in situ but based on one index that is highly correlated with the real losses, even when those losses do not occur immediately after the index's trigger. Index insurance policies determine a threshold and a range of index values based on which indemnities will be paid (Barnett and Mahul, 2007) and (Barraza, 2012). Many new agricultural insurance initiatives rely on this type of information (Chantarat et al., 2013).

The successful application of the NDVI for estimating the impact of weather on vegetation is compromised in non-homogeneous areas (Kogan, 1990). The problem with this approach is that the difference between vegetation index levels in these areas, in addition to weather impacts, can be associated with differences in geographic features, such as

* Corresponding author.
E-mail addresses: omar.valverde@upm.es (O.V. Arias), alberto.garrido@upm.es (A. Garrido), mvilleta@estad.ucm.es (M. Villeta), anamaria.tarquis@upm.es (A.M. Tarquis).

http://dx.doi.org/10.1016/j.geoderma.2017.01.018
0016-7061/

climate, soil, land use and topography. Thus, it is difficult to correctly interpret this index without a sound understanding of local soil conditions. It is therefore necessary to have a soil map adapted to index-based insurance (Taylor et al., 2009).

One of the most important conditions for acceptable index performance is having sufficient knowledge of the influence area in which this index works, due to the spatial variability in soil and climatic characteristics (Vedenov and Barnett, 2004). Furthermore, the surface of the homogenous areas should be representative to make the application of index-based insurance feasible (de Leeuw et al., 2014).

NDVI values are usually obtained with higher resolution images and then compared with a loss threshold curve that has been estimated based on lower resolution imagery, due to a higher annual and monthly availability of lower resolution imagery. This approach may result in the further discrimination of the spatial variation of the index in current readings, compared with the value defined by the threshold that estimates the average loss (de Leeuw et al., 2014). It can also lead to a non-negligible basis risk because the index value may not correlate properly with the potential insured losses in the entire index influence area (Chantarat et al., 2013).

Researchers have developed new methods for remote sensing technologies that minimise the bias found in NDVI temporal series to eliminate the effects that have no relationship with vegetation changes within the process of established index values (Barraza, 2012).

To establish loss thresholds from satellite imagery, we have to sample representative pixels where the crop of interest is located, in critical stages that represent NDVI variations in the study area both temporally and spatially. If we have land usage information, we can focus sampling only on the crop area of interest. If we also have soil characteristics, better results can be obtained by stratifying the sample into different main homogeneous areas with similar soil and climatic conditions, anticipating soil variation and their effects on index values (Milich and Weiss, 2000). Therefore, we can reduce basis risk because index values will have a higher correlation with losses that depend on the soil conditions of each homogeneous class (Tan, 2005).

The physical and chemical characteristics of the soil as well as the topographical and climatic parameters directly influence the availability of soil water content for crops, the crop conditions and yield (Lascano et al., 2007). Therefore, we need to analyse which of these variables are the most important and identify those that have little influence on the spatial variability of the data. This approach will facilitate identifying homogeneous areas with similar characteristics that will be reflected on a more robust index.

The study area of this paper is in the coastal region of Ecuador. In this country, there is subsidised conventional agricultural insurance (Ecuador, 2015) in which the main insured crops are rice and hard maize, both in terms of insured value and subsidies. Ecuador is a small (276,841 km^2) but very heterogeneous country, featuring many soil types, landforms, climates and microclimates. Although the coastal region is more regular than the Andean region, it is still necessary to determine different homogeneous areas that have the same soil and climatic conditions.

Traditional approaches to homogenise a soil characteristics map in the context of insurance are based on a combination of experience and intuitive surveyor criteria, which is supported by extensive expertise on soils, crops and climate (Bourennane et al., 2014). However, depending on different expert judgements, the results could be very different (Taylor et al., 2009).

The aim of this paper is to generate a map of homogeneous classes for areas cultivated with rice and maize in several provinces of Ecuador, to create a basis for the implementation and management of index-based crop insurance. To this end, we have used several statistical methods to delineate homogeneous areas based on soil characteristics, temperature and precipitation. The resulting map was compared to two other maps based on an expert classification and a simple combination of precipitation and temperature without considering some chemical soil characteristics. A detailed discussion of the three maps is presented, taking into account the accuracy and the number of classes that best fit the needs of index-based crop insurance.

2. Materials and methods

2.1. Location of study area and data acquisition

The study area is located (Fig. 1A) in the coastal region of Ecuador, where rice and hard maize are grown, in the provinces of Manabí, Santa Elena, Los Rios, Guayas, El Oro and Loja, with a total area of 6,290,452 ha.

This study was performed using a 1:250,000 scale agro-ecological map of Ecuador (MAGAP, 2015), which includes topographical, physical and chemical soil characteristics as well as climatic variables (isohyets and isotherms). The values of these variables are coded in a geographic database (*Shapefile* downloadable at http://geoportal.magap.gob.ec).

2.2. Description of variables

The database of map variables was generated by characterising soil units through field observation measurements, laboratory soil sample analysis and digital photo interpretation adapted from FAO (2006). We had one topographic variable, slope, which was expressed as a percentage and was obtained from a digital elevation model. Soil variables were divided into physical and chemical characteristics (see Table 1).

The soil physical characteristics are defined as follows (see units in Table 1):

a) Texture (T) is the percentage of sand, loam and clay particles present in the soil.
b) Effective depth (ED) is the usable soil depth for crops.
c) Stoniness (ST) is the percentage of rocks in the soil surface.
d) Soil drainage (D) is the amount time that water is retained in the soil profile.
e) Ground water (GW) is the depth to the ground water table.
f) The chemical soil characteristics are as follows:
g) pH is a measure of the concentration of hydrogen ions in the soil solution.
h) Organic matter (SOM) is the organic matter content in the soil.
i) Salinity (SS) is a measure of the electrical conductivity of the soil solution.
j) Toxicity (Tx) is determined by the amount of $CaCO_3$ and Aluminium in the soil.
k) Fertility (F) is a subjective classification that qualifies the natural fertility level of the soil based on pH, SOM, base saturation, cationic exchangeable capacity and exchangeable bases.

The climatic variables include the following: a) isotherms, which are the annual average temperature ranges, with values from 4 to 19 °C, and b) isohyets, which show precipitation ranges with values from 0 to 1200 mm year^{-1}. These values correspond with those found in the study area. Both isotherms and isohyets were generated based on meteorological station data.

In the present study, rice (266,235 ha) and hard maize (310,691 ha) cultivated areas in 2014 were also used (*Shapefile* given by General Coordination of National Information System -CGSIN in Spanish- of Ministry of Agriculture, Livestock, Aquaculture and Fisheries of Ecuador -MAGAP in Spanish-), concentrating this analysis only on those areas (Fig. 1B).

2.3. Map generation

The study area was classified into homogeneous zones according to their agricultural and technical purposes. The variability of the study area is represented by the combination of topographic, soil and climatic

Please cite this article as: Arias, O.V., et al., Homogenisation of a soil properties map by principal component analysis to define index agricultural insurance policies, Geoderma (2017), http://dx.doi.org/10.1016/j.geoderma.2017.01.018