Boletín de la Asociación de Geógrafos Españoles N.º 68 - 2015, págs. 523-529
I.S.S.N.: 0212-9426

ESTIMATION OF SOIL MOISTURE LEVELS AT FIELD CAPACITY AND WILTING POINT BY PREDICTIVE MODELS ON A REGIONAL SCALE

Pedro Pérez Cutillas
Departamento de Geografía. Universidad de Murcia. Centro de Edafología y Biología Aplicada del Segura, Consejo Superior de Investigaciones Científicas (CEBAS-CSIC)
pedrope@um.es, perezcutillas@cebas.csic.es
Gonzalo G. Barberá
Centro de Edafología y Biología Aplicada del Segura, Consejo Superior de Investigaciones Científicas (CEBAS-CSIC)
gbarbera@cebas.csic.es
Carmelo Conesa García
Departamento de Geografía. Universidad de Murcia
cconesa@um.es

I. INTRODUCTION

The soil moisture content is an essential factor in the dynamics of water flows in semiarid regions (Dunkerley, 2002). So, knowledge of the relationship between the matric potential (ψ) and the water content in the soil (θ) is critical in order to describe the hydrodynamic properties of the soil. The water content is usually regulated by the water retention capacity of each soil at a given negative pressure, such that the water availability is determined by the relationship between these processes, together with the previous moisture content of the hydrological systems under study (Malik et al., 1987). The water retention curve expresses the relationship between the water content (θ; $m^3\ m^{-3}$) and soil matric potential (ψ, Pa) and provides information on important features such as the maximum storage capacity for water in the soil, field capacity, and wilting point.

The measurement of hydric properties, in both the field and laboratory, is complex with regard to time and resources as well as requiring a large number of samples due to the spatial variability of soil (Klute, 1986). Numerous studies have recognized the usefulness of predictive regression models for the determination of soil hydraulic functions from such properties over large areas (Wösten and Van Genuchten, 1988; Schuh and Cline, 1990; Twarakavi et al., 2010). However, there are very few reports dealing with the analysis and application of such techniques that, based on the use of statistical models, try to estimate soil hydraulic properties at the regional level by means of environmental variables.

Such techniques are based primarily on the spatial dependence of the studied variables, are suggested for areas that are not very large, and require quite intensive and regular sampling (Chen et al., 2000; Fox and Sabbagh, 2002). As the density of measurements is not usually high on a regional scale, the use of other methods is required to predict the values of the soil variables in a simple and affordable manner (Liess et al., 2012). A particularly effective alternative is to model the relationship between the variable of interest in the soil and environmental variables for which spatially distributed data are available (Gessler et al., 1995; Brus and De Gruijter, 1997; McKenzie and Ryan, 1999; Thompson et al., 2001; McBratney et al, 2003).

From a different perspective, remote sensing allows an overview of large areas, providing massive amounts of quantitative, georeferenced data and spectral values that show good correlations with the areas occupied by different land uses, besides having a very dynamic capacity for temporal updates (Ben-Dor, 2002). The assessment, characterization, and determination of soil properties, using data from remote sensors, have been widely applied in recent years (Scull et al., 2002; Rawls et al, 2004; Uno et al, 2005; Vrieling, 2006; Lagacherie et al, 2012; Poggio et al, 2013). The capacity of spectrometry, under laboratory conditions, to predict important soil properties has been shown (Schulten and Schnitzer, 1997; ViscarraRossel et al., 2006; Ben-Dor et al., 2009).

The results of the cited work served as the basis - in this study - for the modeling processes, using regression analysis with a series of environmental variables, needed to calculate the values of field capacity and wilting point from the soil moisture retention curve.

II. STUDY AREA AND ENVIRONMENTAL CONDITIONS

The study area selected was the region of Murcia (11313 km^2, INE, 2013), within the Segura River basin (18208 km^2, CHS, 2013) in Southeast Spain. Overall, this is a rather hilly landscape due to the presence of numerous mountain ranges belonging to the Betic Cordilleras, aligned in an ENE-WSW direction, often with altitudes above 1,000 m. Interspersed among them is a series of valleys, basins, plains, and plateaus which together configure a contrasting topography and a unique area of great variety in terms of landscape. Most of the study area is subjected to a semi-arid Mediterranean climate (Capel-Molina, 2000): rainfall in the basin is around 375 mm/year, with average annual values of 472 mm in the top part of the basin and 317 mm near the mouth, but with high seasonal and inter-annual irregularity of rainfall, prolonged droughts and, often, torrential rain.

This whole region has a strong annual insolation, with an average of 2800 hours of sunshine per year, with a maximum in July (340 h) and minimum in December (160 h); the hilly nature of the regional relief means that the local variations in the received radiation are notable. Meanwhile, desertification and loss of soil by erosion are the most important causes of land degradation in this area, as they imply the loss of its main physical, chemical, and biological components (Boix-Fayos et al., 2005). In this basin there is a wide diversity of soils, caused by the action of various processes associated with the influence of multiple environmental factors. These include the weather conditions, lithology, and relief, which largely condition the erosion and soil processes (Alvarez Rogel et al., 2001).

III. MATERIAL AND METHODS

III.1. Analytical data of the LUCDEME Project

All the empirical information on the suction potential (pF) in the soil was obtained from data of the project *Lucha contra la Desertificación en el Mediterráneo* (LUCDEME) (ICONA, 1986). The information used in this part of the work derives from the complete data of a selection of soil profiles, sampled in the field and analyzed in the laboratory. The different taxonomic units of the FAO-UNESCO system (1974) - extracted from the LUCDEME project for the study area - are characterized by 307 profiles, sampled *ad hoc*, distributed across the territory in places where these units are better represented. The information on the profiles appears classified by the depths of the different horizons observed in the field, depending on their physicochemical properties listed in 1880 records.

III.2. Environmental variables used in modeling

III.2.1. Topographical and hydrological variables

These variables were obtained from the Digital Terrain Model (DTM) of the NASA mission TERRA, performed with the Japanese ASTER sensor (NASA - METI, 2013). From this DTM (**1**) were obtained: Slopes (**2**), using the method of maximum change in the altitude value with respect to the surrounding pixels for each pixel, with values expressed in degrees of inclination (Burrough and McDonnell, 1998). The curvature was calculated analogously to the map of slopes, but using the second derivative, and was expressed as the slope of the slope (Moore et al., 1991). The curvature can be broken down (Zeverbergen and Thorne, 1987) into the profile curvature (**3**), which is calculated from the direction of maximum slope, and the perpendicular curvature (**4**). The flow accumulation (**5**) represents the number of pixels in the basin draining to a particular pixel (Tarboton et al., 1991). It should be noted that this variable has a distribution that is highly skewed to the right, so its logarithm was used for the modeling.

III.2.2. Types of soils and lithology

The pedological and lithological information was obtained from different sources. In the soil mapping of the LUCDEME project (Scale 1: 100,000) (**6**), with the spatial distribution of taxonomic units classified according to the FAO System (1974), most of the soil polygons do not identify a taxonomic unit but rather an association of two or three units. Also employed were the map of Saline Phase soils (**7**), identified as areas whose electrical conductivity exceeded 2 mmhos/cm, and the lithology map (**8**), made using the data of the National Geological Map (1:50,000) (MAGNA) of the Institute of Geology and Mining of Spain (IGME); the geological mapping has been reclassified to obtain the lithological values through interpretation of the legend of each geological formation that appears in the report that accompanies each page of the map.

III.2.3. Climatic Variables

The climatic variables used are the layer of precipitation (mm) (**9**) and the layer of temperature (°C) (**10**) for the study area, developed by the Laboratory of Biogeographical Informatics (LBI, 2013) of the National Museum of Natural Sciences of the CSIC, and various variables related to solar radiation (Rich et al., 1994). Calculation of the radiation has been performed for a full year, yielding three layers of information: the direct solar radiation (**11**) (Fu and Rich, 2000) (the radiation that the surface receives directly from the sun), the diffuse solar radiation (**12**) (Fu and Rich, 2000), corresponding to the radiation that reaches a point not directly from the sun but is the proportion that has been scattered by the atmosphere, and the duration of the solar radiation (**13**), indicated in hours of direct sunlight for each of the pixels.

III.2.4. Additional Variables obtained by Remote Sensing

Landsat 5 satellite images from two dates were used to obtain additional information. These images correspond to two different seasons of the year, summer (24/07/2009) and winter (14/02/2009). In this way the possible seasonal bias is minimized. For each period the Normalized Difference Vegetation Index (NDVI) (**14** and **15**, data for winter and summer, respectively) was estimated. This index allows determination of the quantity, quality, and development of the vegetation according to the measurement of the intensity of radiation of wavelengths in the Red (R) and Near Infrared (NIR) parts of the electromagnetic spectrum that the vegetation emits or reflects (Townshend et al., 1985).

Also, a series of algorithms have been applied to the selected images, to yield mineralogical indicators (Sabins, 1981; Crosta et al., 2003) by combining the following standardized functions: «Clay Minerals Index» (clay minerals, CMI) (**16**, **19**), «Ferrous Minerals Index» (iron minerals, FMI) (**17** and **20**), and «Iron Oxide Index» (iron oxides, IOI) (**18** and **21**). For each of the functions the calculations were performed for the summer (July) and winter (February) seasons.

III.3. Predictive models for estimating soil moisture content at field capacity and wilting point

From the coordinates of the samples of the database, the values of 21environmental variables were obtained at a pixel resolution of 400 m. Due to the discrete nature of the variables of the block SL, in the models proposed there is a coefficient for the effect of each class level (i.e., for each type of soil or lithology).

The *pF* values available for each point were related, in a table, to the corresponding values of the selected environmental variables - with which regression models were built according to the method described below. A regression model in which the predictors are GIS layers can easily be represented by map algebra since it is a simple linear combination of these layers.

The pixel resolution for the models was 400 m; this spatial resolution provides certain detail and is consistent with the modeling of these values at the regional level. In addition, Pérez-Cutillas (2013) showed that the use of lower pixel resolutions (< 400 m) does not substantially improve the predictive ability of the models. To obtain the model of *pF* as a function of the environmental variables, a forward stepwise linear regression was applied, using the Akaike Information Criterion (AIC) (1974).

IV. RESULTS AND DISCUSSION

IV.1. Modeling pF values as a function of environmental variables on a regional scale

The best predictive model for the soil water content at field capacity (*pF* 2.5), generated by the method described and in accordance with the AIC values obtained for each combination of variables, is described in the following formula:

$$\textbf{\textit{Modelo pF}}\ \mathbf{2{,}5} = [\boldsymbol{DTM + FAc^2 + RaDi + RaDr + RaDu^2 + FMI_{feb} + NDVI_{jul} + FSLu}]$$

The best model obtained for the values at wilting point (*pF* 4.2) is shown below:

$$\textbf{\textit{Modelo pF}}\ \mathbf{4{,}2} = [\boldsymbol{DTM + DTM^2 + CuP + CuP^2 + FAc^2 + Tmp^2 + Pr^2 + RaDi + RaDr + NDVI_{jul} + LM + FSLu}]$$

The models obtained show an uneven coefficient of determination: $R^2 = 0.15$ for *pF* 2.5, and $R^2 = 0.31$ for $pF = 4.2$.

IV.2. Influence and effect of environmental variables on the models of pF

To determine the effect exerted by different variables in the models used to estimate the values of *pF*, one can get an overview by observing the model coefficients (Tables 2 and 3). The effects analyzed are based on the impacts of the variables on the different models, obtained through consideration of the coefficients obtained in the modeling processes.

Before analyzing the effect of the variables, a review of the inequalities found between the two models is worthwhile. In fact, significant differences were found in the number of variables of which the models were composed, in particular model at *pF* 4.2 - more complex and with greater predictive power than the model at *pF* 2.5. The greater uncertainty of the model describing the Field Capacity may be due to the lower precision of the measurement processes at low *pF* (Richards, 1947).

The effect of the variables on *pF* is enormously interesting due to the influence exerted by the ecological parameters according to their distribution and spatial variability. The TG group variables that appear in the model at *pF* 2.5 are [DTM] and [FAc]. In the first case, increasing altitude is linked with an increase in the value of soil moisture at field capacity. The [FAc] is presented as the square values, with a progressive increase in estimated soil moisture at *pF* 2.5 in areas with high values of flow accumulation. For the model at *pF* 4.2, the effects of these variables are similar, with the difference that [DTM] occurs in a quadratic model - with the result that, at altitudes above 1300 m, the calculated tendency of pF to increase is inverted and the values decrease. The model at *pF* 4.2 includes the curvature of the profile of maximum slope [CuP] - giving it a non-linear form, with increased moisture content in the areas of convex curvature and decreased values in the concave areas. This influence of the topographic values on the models is consistent with the results obtained in other studies (Poggio et al., 2010), according to which the parameters derived from the DTM proved useful as auxiliary information in the prediction of the available water capacity in multiple spatial areas.

In group C, only the variables [Pr] and [Tmp] are involved in the model at *pF* 4.2. The effects are similar and operate in the same way, increasing the moisture content at higher temperatures and precipitation. With regard to radiation, [RaDi] and [RaDr] show similar behavior in the two models of *pF*: increasing values of diffuse radiation produce a decrease, while for the direct radiation this effect is amplified. The variable [RaDu] only appears in the *pF* 2.5 model, acting in a non-linear way to produce an increase in moisture content at *pF* 2.5 as the number of sunshine hours increases.

Finally, the RS group appears only with [IOI_{feb}] for the model at *pF* 2.5 and with [$NDVI_{jul}$] for both *pF* models. Regarding the mineralogical index, the effects cause a decrease in the *pF* 2.5 values as the levels of mineralogical alteration increase for the elements detected in these parts of the electromagnetic spectrum. For [$NDVI_{jul}$], the effects are similar for both models, the estimated values of *pF* increasing with increasing values of the vegetation index. This information is quite consistent, generating little bias in the type of ecosystem associated with the vegetation activity, which is what this variable really represents, and showing a connection between vegetative development and the available water capacity in the soil (Todoroff et al., 2010).

The Saline Phase of the LUCDEME data [SPLu] appears as a qualitative variable that contains boolean information, in which the assignment of a value of 1 or 0 to the defined portion of each polygon expresses, respectively, the presence or absence of salinity. Meanwhile, following the methodological criteria specified above, the effects of the categories of the variable Lithology [LM] have been evaluated with reference to the 'Alluvial' class. The areas in the Saline Phase [SPLu] produce increased values for both *pF* models, being the only qualitative variable that affects the *pF* 2.5. In the model *pF* 4.2 the variable [LM] affects almost all of the categories, with a negative effect on the *pF* with respect to the reference level 'Alluvial', while the classes 'Colluvium' and 'Metamorphic - Siliceous' reach more extreme values. Only the category 'Dolomites' shows a positive effect, albeit a very weak one.

V. CONCLUSIONS

The models of soil water content at different *pF* diverge structurally, suggesting that the environmental processes that determine the water holding capacity of the soil at different *pF* values are, at least partially, distinct. These results show the model *pF* 2.5 to be less complex than model *pF* 4.2.

For the *pF* 4.2 model, the variables that are best represented are in the [TG] and [C] groups. The effects of DTM, CuP, and FAc on the Wilting Point are associated with an increase in the *pF* in zones of high altitude (up to 1500 m), in concave surfaces, and in larger drainage areas.

Also significant is the effect of $NDVI_{jul}$, which participates in the two *pF* models, giving an increase in the *pF* values in areas with more vegetation, due to the relationship between vegetation and organic matter and the capacity for retention of soil moisture.

Finally, for our results, it can be stated that the information provided by remote sensing has allowed the improvement of the models. So, the use of remote sensing data, in combination with other environmental variables, can be recommended.