MAIN ENVIRONMENTAL FEATURES LEADING TO RECENT LAND ABANDONMENT IN MURCIA REGION (SOUTHEAST SPAIN)

Francisco Alonso-Sarría[1*], Carlos Martínez-Hernández[2], Asunción Romero-Díaz[2], Fulgencio Cánovas-García[3], Francisco Gomariz-Castillo[1,4]

[1]*Water and Environment Institute, University of Murcia, 30100 Murcia, Spain*

[2]*Geography Department, University of Murcia, 30001 Murcia, Spain*

[3]*Departamento de Geología y Minas e Ingeniería Civil, Universidad Técnica Particular de Loja, San Cayetano Alto s/n, Loja, Ecuador*

[4]*Euro-mediterranean Water Institute, 30100 Murcia, Spain*

ABSTRACT

Land abandonment is a global phenomenon whose environmental consequences are difficult to asses. Murcia Region is one of the most arid regions in southern Europe and also one of the most prone to land abandonment. This study researches which environmental features are more relevant to explain abandonment at agricultural plot scale. Geomorphometric features were measured at different scales to investigate which scales could be more relevant. Two different models have been used: logistic regression, a statistical model that allows the interpretation of the involved features, and Random Forest, a machine learning model with a higher predictive power but lower interpretability. The combined use of both such models allows a set of predictors to be selected, which, when used with Random Forest, produce a map that is highly accurate for predicting abandonment and, when used with logistic regression, produce an interpretable model. The main conclusion is that climate is the most relevant factor to explain land abandonment.

KEY WORDS: Land Abandonment, Feature Selection, Data Analysis, Random Forest, Logistic regresion

INTRODUCTION

Land degradation is related to the human pressure on the ecosystems, either by the increase of the population or the increase of the human activities due to new technology or mismanagement. Both agricultural intensification and population increase result in soil erosion, loss of biodiversity, soil degradation, vegetation changes and human and social changes (Zhao *et al*., 2013; Mandal *et al.,* 2013; Li *et al*., 2013; Wu *et al*., 2013; Al-Awadhi *et al*., 2013). Such processes lead to land degradation and a loss in the soil services to humankind, jeopardizing societies sustainability (Keesstra *et al*., 2012; Brevik *et al*., 2015). On the other hand, land abandonment is a global phenomenon associated with a progressive reduction both in traditional agricultural practices (MacDonald *et al*., 2000) and in rural population. Several authors have stressed land abandonment as an environmental problem leading to land degradation (Moravec & Zemeckis, 2007; Lenda *et al*., 2012; Navarro & Pereira, 2012). It has been studied, among other places, in Oceania (Beilin *et al*., 2014), South America (Grau & Aide, 2008; Aide *et al*., 2012; Schneider & Geoghegan, 2006; Franco *et al.,* 2012; White *et al.,* 2013; Lugo & Helmer, 2004), North America (Walton *et al*., 2008; Ramankutty *et al*., 2010), the former Soviet Union (Prishchepov *et al.,* 2013), China (Shang *et al*., 2008; El Kateb *et al*., 2013; Deng *et al.,* 2012; Fang *et al.,* 2012), Nepal (Khanal & Watanabe, 2006), Irak (Gisbson, 2012), Iran (Raiesi, 2012) or South Africa (Kakembo & Roentree, 2003).

Land abandonment has been also widely studied in Europe (MacDonald *et al.*, 2000; García-Ruiz *et al*., 2010; García-Ruíz & Lana-Renault, 2011; Pointereau *et al.,* 2008; Parson, 2014). In the most developed European countries, abandonment reached a peak during the industrialisation in the XIX century and after World War II (Gellrich & Zimmermann, 2007). Nowadays, it is more related with European policies (Renwick *et al.,* 2013; Pointereau *et al.*, 2008) and affects thousands of squared kilometres throughout Europe (Cerdà, 1997; Suárez-Seoane *et al*., 2002; Cramer *et al*., 2008; Cammeraat *et al*., 2010). Land abandonment studies in Europe have been mostly conducted in the mediterranean basin (Kosmas *et al.,* 2002; Detsis 2010; Rey-Benayas *et al*., 2007, 2014), in countries such as France (Piegay *et al*., 2004; Cosandey *et al.,* 2005; Sluiter & De Jong, 2007; Bakker *et al*., 2008), Italy (Dunjó *et al.,* 2003; Maccherini *et al*., 2013; Ricotta *et al*., 2012; Renzi *et al*., 2002; Giupponi *et al.,* 2006; Garfi *et al*., 2007), Greece (Kosmas *et al*., 2000; Kouloui & Giourga, 2007; Zakkak *et al.,* 2014); Kizos & Koulouri, 2006; Bakker *et al*., 2008), Portugal (Nunes *et al.,* 2012; Proença *et al*., 2012) and Spain. The results of those studies show a recovery of the natural vegetation that resulted in a reduction in the runoff discharge, and a clear decrease in the sediment yield (Keesstra *et al.,* 2007) which has been demonstrated by direct measurements and modelling (Keesstra *et al.,* 2009; 2014).

In Spain, several studies have been conducted on the environmental, biological, hidrological and geomorphological consequences of land abandonment. The scope of such studies has been generally very local: some valleys in the Pyrenees (Ruíz Flaño, 1993; Lasanta *et al*., 2005; García Ruíz *et al*., 2010; García Ruiz & Lana-Renault, 2011), the Iberic System (Lasanta *et al*., 2001; Arnaez *et al*., 2011; Lasanta el al., 2014), Canary Islands (Arbelo et al., 2006), some sectors in the Ebro Valley (Ries & Hirt, 2008; Sauer & Ries, 2008), South Spain (Ruíz Sinoga & Martínez Murillo, 2009), and SouthEast Spain (Padilla, 1997; Symeonakis *et al.,* 2007; Lesschen *et al.,* 2007, 2008; Romero Díaz *et al.,* 2007, 2012; Bellin *et al*., 2009; Nadal-Romero *et al*., 2013; Calatrava *et al*., 2011; Alados *et al.,* 2011; Cerdá *et al.,* 2012; Martínez Hernández *et al.,* 2013; Robledano *et al*., 2014; among others). Land abandonment is one of the most characteristic processes in the spanish agricultural evolution since the end of the XIX century. Its highest intensity was reached during the sixties and the seventies in the XX century (García-Ruiz *et al*., 2010). More recently, the rapid urbanisation process (Jiménez-Herrero *et al*., 2005) and the European Agricultural Policies that subsidise the abandonment of less profitable crops (Errea & Lasanta, 2001) have contributed to the process.

Aridity is one of the main reason that explain the traditional underdevelopment of the Spanish agriculture, as it limited the introduction of new techniques resulting from the agricultural revolution (Cazcarro *et al*. 2014). In the first years of the 20th century, the Spanish government began to fund some hydraulic plans, but it was in the fifties and sixties when hydraulic infrastructures became one of the most important policies in Spain. In Southeast Spain, the combination of irrigation and a large insolation gave a significant advantage in the production of agricultural goods. When Spain joined the European Union, more important changes took place as a result of the Common Agricultural Policy (CAP) that encouraged certain products, as fruit and olive trees, and discouraged others, as tubers, cereals or fodder crops (Pinilla & Serrano, 2009). As a result of these historical factors, a re-allocation of agricultural production, from humid areas in the North to semi-arid areas in the South and Southeast, took place from 1930 to 2005. The contribution of Murcia to the Spanish agricultural GVA increased 2.9 percentage points during this period (Cazcarro *et al*., 2014). Nowadays, and according with the Spanish National Statistical Institute, Almeria (5.5 %) and Murcia (5.4 %) have the largest provincial contributions to the Spanish agricultural GVA (Colino *et al*., 2104). In fact, the agricultural history in Murcia is very similar to that of Almeria province (Faulkner *et al*., 2003).

However, the need of larger amounts of water and financial resources per hectare in this intensified agriculture has lead to a reduction in the cultivated land. According to the Murcia Region Department of Agriculture and Water, in the period 2004-2011, the rainfed area has reduced from 409,330 ha to 364,024 ha, while the irrigated area has declined from 200,878 ha to 194,332 ha; however, in a similar period (2007-2013), highly intensive irrigated areas have increased from 138,219 ha to 147,807 ha (Colino *et al*., 2014).

There is also a change in the farmer type. Traditional farmers with familiar exploitations have been substituted by large enterprises and cooperatives employing a large amount of inmigrant workers. On the other hand, the offsprings of traditional farmers have abandoned agriculture to other sectors (mainly building during the real estate bubble). This process is also similar to that occuring in Almeria (Rodríguez-Vaquero, 2008).

In Murcia Region, Romero Díaz *et al*. (2007) highlights that, since 1980, land abandonment has affected a large percentage of cultivated areas in Murcia region, mostly in fragile soils formed on marly Neogene-Quaternary basins. More recently, Martínez-Hernández *et al.* (2013) indicated that cultivated land decreased by 46% in the period 1991-2001, while the non cultivated area increased by 33%, especially forests and abandoned land. Causes of land use change have been divided into proximate and underlying (Geist *et al.*, 2006). Proximate causes operate at a local level and include a set of the physical actions on the land that change it. Underlying causes are forces that act more diffusely and trigger the proximate causes. Underlying causes include byophisical, economic and technological, demographic, institutional, and cultural factors. Land use changes in Murcia, as in the whole Mediterranean basin, are related with socioeconomic dynamics, especially with the decrease in rural population (Kosmas *et al.*, 2002). Other factors in western Mediterranean countries are European Community agriculture policies that try to set-aside less profitable crops (Errea & Lasanta, 2001), urban expansion (Ghosh *et al.*, 2014) or market changes (García-Ruiz 2010). However, according to the staff in *Consejería de Agricultura de la Comunidad Autónoma de la Región de Murcia* (Regional agriculture authority) just a small percentage (around 5-10%) of the abandoned land in Murcia has benefited from European Agricultural Policy subsidies (unpublished data). Although global socioeconomic factors may be responsible for establishing global trends that affect the whole territory, the traditional farmers' final decisions about abandoning individual plots are more related with environmental factors.

Given the important consequences of abandonment, both from an environmental and from a land management point of view (Quiñonero-Rubio *et al*., 2013), there is a need to evaluate the phenomenon. At the same time, it is important to identify both the areas most prone to abandonment and the most relevant environmental features leading to land abandonment in order to establish appropriate policies to manage the land and to attenuate the negative effects where they appear. Land use change has been studied using stochastic modelling (Weng, 2002), aerial photographs analysis (Lasanta-Martínez *et al.*, 2005), and multitemporal remote-sensing imagery classification (Serra *et al.*, 2008). Both statistical methods and machine learning algorithms have been used in different classification contexts, both to classify and also to identify which of the independent variables used to classify are more important in the classification. Random Forest and Logistic regression are specially adequate to obtain variable importance. Given that abandonment can be considered as a binomial classification problem, classification techniques can be used to estimate potential of abandonment and also to discover which of the features used as independent variables are more relevant to distinguish between abandoned and non abandoned plots and, consequently, are more related with the abandonment process.

Mapping is a very helpful tool to locate the land degradation processes and to foresee the expected changes. It has been also used to understand the factors and the behaviour of landscape changes (Desprats *et al*., 2013; Xu *et al*., 2014; Jafari & Bakhshandehmehr 2014). Mapping the potential of abandonment is then the second objective of this study.

The main objective of this study is to identify the environmental (biophysical) features most relevant to explain land abandonment in the time span 2001-2009 at a local, agricultural plot, scale. For the purposes of this paper, we define agricultural abandonment as the consistent cessation of farming activity in a particular area indefinitely and without any recent attempt to resume this activity or any other profitable activity. We do not include in this definition land that was abandoned in a past secular socioeconomic context, in which the socioeconomic conditions were different to those that affect the present day abandonment processes. We consider that a resolution of 25 meters is large enough to represent properly the plot scale because according to the Murcia Region agricultural census, plots smaller than 0.5 ha represent just a 0.21% of the cultivated land (INE, 2009); working with a 25 meters resolution, a surface of 0.5 ha are represented by 8 pixels. We are aware that, as has been previously mentioned, socioeconomic features are also relevant to explain land abandonment; however, at agricultural plot scale, environmenral features are more relevant; moreover, it would be very difficult to obtain socioeconomic information at such a detailed scale.

MATERIALS AND METHODS

Study Area

Murcia region is located in the southeast of the Iberian Peninsula, and its extension is about 11,317 km^2. From a geological point of view, it is part of the Betic Cordilleras. Twenty-three percent of the territory is below than 200 m.a.s.l.; 44.7% is between 200 and 600 m.a.s.l., and 32.3% is above 600 m.a.s.l. Overall, this area is quite hilly with an alternation of ranges, specially in the northwest, which often exceeds 1,000 m.a.s.l., reaching 2,000 m.a.s.l. as maximum elevation; plateaus (in the northeast area); plains (in the southeast); and numerous valleys, basins and inter-mountain corridors that compartmentalise the relief. Climate is Mediterranean with semi-arid features. The annual rainfall is less than 350 mm, with the exception of some areas in the upper northwestern lands where it exceeds 600 mm. Temporal distribution of rainfall is irregular with long dry periods combined with short and intense rainfall events that produce frequent flash floods and floodings. The average annual temperature is between 15 and 19 centigrade degrees. Winters are short and summers are long and hot. Insolation exceeds 2,800 hours per year and even 3000 in some southern areas. Evapotranspiration is between 1,200-1,300 mm. As a result of the climatic characteristics most drainage channels are dry most of the year. Just the Segura river, the main river crossing the study area, is a permanent river. Drought periods affect intensely the landscapes and aridity imposes a clear contrast between rainfed and irrigated agriculture. The population is around 1.5 million, with a density around 130 inhabitants/km^2, but with large differences among municipalities. Agriculture has long been the basis of the regional economy. In recent years, intensive agriculture, using water from the Tajo-Segura transfer, has transformed the agricultural sector. On the other hand, unprofitable rainfed lands are suffering a more intense abandonment process.

Land abandonment map

Due to the size of the study area (about 12,000 km^2), to create an abandonment map from scratch was infeasible, so we decided to use the most recent and accurate official land use map, also with the largest scale, that existed at the time. It is the SIOSE (Spanish Land Occupation Information System) map (Arozarena *et al.*, 2006). The SIOSE project was coordinated by the Spanish *Instituto Geográfico Nacional* (National Geographic Institute). Its main output is a 1:25,000 land use/land cover layer obtained by photointerpretation of several 2005 SPOT images supported by the analysis of several landsat images, orthoimagery and different basic and derived ancillary maps. The process is fully described in IGN (2005).

The land use polygons in the layer receive, in some cases, additional attributes besides the land use label. With regard to land abandonment, some of the polygons were labelled as pastures including the attribute "formerly cultivated". Such polygons might be identified as abandoned plots. However, this dataset has several limitations. It is not accurate enough to directly generate a map of abandoned agricultural plots by a simple query to its database. Besides the possible inaccuracies in the map, SIOSE uses a usual strategy in the generation of coverage maps: creating mosaics to aggregate plots whose size is under the minimum mappable area unit, which depends on the map scale. Coverage association is a similar concept; in this case, different coverages blend with each other without clear borders and are represented as a single polygon. An example is the association of forest and scrubland. As "formerly cultivated pastures" is the unique label in only 27% of the 4701 polygons that carry it, it has been necesary to to identify, at each plot occupied by an association or a mosaic, which sector really corresponds to abandoned plots. Consequently, SIOSE polygons were used as a first approximation for an abandonment map. So it was neccesary a subsequent and laborious photointerpretation process.

Moreover, such polygons are a snapshot of the situation in 2005; however, land abandonment is a complex process that has to be followed along a larger time span. In a second step, and to refine that first approximation, a photo-interpretation (not digital classification) analysis was carried out on a regional 2003 Quick-Bird image and on several digital or digitised aerial photographs corresponding to the years 2002, 2004, 2007, 2008, 2009 and 2011. The reason why we used such a large amount of images is that sometimes it is not easy to decide, looking at just one image, if a plot is abandoned; it looks abandoned due to the phenological calendar of the crop, but is actually being cultivated; or has a temporal activity cessation, but with a further intention of resume the cultivation. Having a series of several images allows us to have a clearer idea about the evolution of every crop. This large time span has allowed to verify such polygons and to detect other abandoned plots. To confirm that a SIOSE polygon labelled as abandoned was, indeed, abandoned, it had to fulfil two basic conditions:

•It must have been cultivated in 1981 and 2001. At first, we took 1981 as the base year because the beginning of that decade has been identified as the start of modern and market oriented agriculture in Murcia region (Cortina García, 1994). Besides, the first good resolution orthophotography (50 cm) was taken in 1981. On the other hand, the only available land use map from which extract the land use prior to abandonment is the *Mapa de Cultivos y Aprovechamientos* (Crops and land use map) (2000-2009) published by the Spanish *Ministerio de Agricultura, Pesca y Alimentación* (Ministry of Agriculture, Fisheries and Food) with field data collected between 2001 and 2007, at 1:50,000 scale. So, the effective time span of this research is 2001-2009.

•It should not have been cultivated in 2007, 2008 and 2009. The aerial othophoto corresponding to the year 2011 was later used as a validation test.

Finally, all areas besides the formerly cultivated pastures (according to SIOSE layer) were checked using the same set of images. In this way, other abandoned plots were identified. A comprehensive description of the methodology appears in Martínez-Hernández *et al.* (2013).

Features

A large amount of environmental features were considered as predictors with a mainly exploratory purpose and to avoid *a priori* assumptions about any particular feature being more important than the others. These features were:

•Land use: obtained from the aforementioned crops and land use map at 1:50,000 scale.

•Lithology: from the Geology map at scale 1:50,000 produced by the *Instituto Geológico y Minero de España* (Spanish Geology and Mining Institute) simplified into a map of 5 categories to represent the five main rock types in the area (Limestone, Detritic, Evaporitic, Metamorphic and Volcanic, and Quaternary deposits).

•Soil type: from the soil map at scale 1:100,000 produced by LUCDEME project (Alias & Ortiz, 1986–2004), in which polygons are labeled with soil associations rather than with individual soil types. The map was reclassified to meet the FAO-2010 criteria; additionally, some of the less frequent associations were also reclassified into larger groups to prevent the presence of very infrequent classes. It would have been interesting to add quantitative soil properties to this analysis; however, the uncertainities of the estimation of such properties for all the study area discouraged us of using them. Just the concentration of soil organic carbon, for which we had an accurate enough map (Blanco-Bernardeau *et al.,* 2014), was included in the analysis.

•Climate layers: Precipitation; absolute minimum, absolute maximum and average temperature; potential evapotranspiration and incident solar radiation. Every variable was estimated on a monthly scale, which means a total of 72 layers. These layers were available from previous projects (Gomariz-Castillo & Alonso-Sarría, 2013): Precipitation and temperature were obtained using GLM models (James *et al*., 2013), potential evapotranspiration using the Allen modification of the Hargreaves equation (Allen *et al*., 1994), and solar radiation using the GRASS module r.sun (Hofierka, 1997). These layers were calculated with a resolution of 400 m.

•Geomorphological features were calculated from an official DEM with a resolution of 25 meters obtained from the *Instituto Geográfico Nacional* (Spanish National Geographic Institute). The resolution of the DEM, and consequently of all the derived features, are adequate to the local, agricultural plot, scale that we are interested in. The geomorphological features were calculated using different scales (window sizes) to investigate which scales could be more relevant. Such calculations were performed using GRASS (Neteler & Mitasova, 2008) and SAGA (Olaya & Conrad, 2009) modules. These features include the most used and the most representative of the different geomorphological features groups (Hengl & Reuter, 2009):

•Elevation. The aforementioned DEM.

•Terrain derivatives obtained from the DEM: Slope, sine, and cosine of the aspect, profile and plan curvatures, all calculated with different window sizes (3, 7, 11, 15, 19, 23, 27, 31, 35 and 39 pixels) to obtain estimations at different scales (from approximately 0.1 has to 95 has). That represents a total of 50 layers. The GRASS module r.param.scale (Wood, 1996) was used to calculate them.

•Roughness using several different measurements: Melton coefficient (Melton, 1965); Terrain Ruggedness Index (Riley *et al.*, 1999); Vector Ruggedness Measure (Sappington *et al.*, 2007); Iwahashi and Kamiya coefficient cited in Iwahashi & Pike (2007). These four coefficient were calculated using the same window sizes used to calculate the terrain derivatives, representing a total of 40 layers.

•Topographic position: Topographic Position Index (Weiss, 2001) was calculated with SAGA using the same scales than in the previous indices. In this way, 10 layers were created.

•Derived topographic indices: Topographic Wetness Index (Quinn *et al.*, 1991; Beven *et al.*, 1995); Multiresolution Valley Bottom Flatness Index (Gallant & Dowling, 2003) and USLE LS factor using the formulation proposed by Moore & Burch (1986).

•Finally, other features: Distance to a major road and distance to natural areas, both calculated with GRASS.

Feature selection

The list of features previously presented make a total of 182 different layers (179 quantitative and 3 qualitative). Besides the computational complexity, other problems when dealing with such a large number of predictors include collinearity and the risk of overfitting the model.

Both statistical and machine learning based models are sensitive to collinearity (Dormann *et al.*, 2013). In the case of logistic (and also linear) regression, parameter estimates may be unstable as standard errors become inflated. So, it is difficult to asses the importance of the features. When using stepwise regression, small variations in a collinear dataset might result in one or the other collinear predictor being dropped from the model, leading, through a different trajectory, to a complete different model.

Machine learning based methods share the problem when using collinear predictors; the obtained model is sensitive to slight changes in the data set, being difficult to interpret the final model or to separate the effects of collinear variables. In particular, one of the most interesting characteristics of Random Forest is that it trains several decission trees using different, randomly selected, feature subsets. This allows a reduction in correlation among trees that increases the power of the final voting system. Collinearity among features increase the correlation among trees and, consequently, decrease the power of the method.

One of the most usual approaches for dealing with collinearity is Principal Component Analysis (PCA); however, we preferred not to used PCA because the new components are difficult to interpret and because the objective of this study is to identify which features are more relevant to explain land abandonment. In contrast, Variance Inflation Factors (VIF) (Zuur *et al.*, 2007) can be used to produce a non collinear subset of the original feature set, so that the resulting models are quite easier to interpret.

The R program (R Core Team, 2014) was used for the VIF analysis and the rest of statistical analysis in this research. The algorithm presented in Zuur *et al.* (2009) was adapted for this research to recursively calculate the VIF for each feature, select the feature with the highest VIF and omit it from the data set, recalculate VIF for the rest of the features, and continue with the next iteration. This procedure continues until no feature has a VIF higher than a given threshold. Zuur *et al.* (2007) recommend a threshold smaller than 10, but, following a more restrictive recomendation from other authors (Kutner *et al.*, 2004; O'brien, 2007), a threshold of 5 was used in this study.

Classification methods

a) Logistic regression

Logistic regression is a particular case of the Generalized Linear Model (McCullagh & Nelder, 1989; Zuur *et al.*, 2009) that can be used with response variables which range from 0 to 1. Thus, it is especially useful for modeling percentages or probabilities. It is also used as a classification method with binomial responses (although it can also be adapted to the multinomial case).

The main advantage of logistic regression, over more sophisticated machine learning methods such as Random Forest, is that the resulting model is easy to interpret and the environmental value can be assessed by examining the coefficients of the predictors entering the model.

This methodology has been used in several research fields as species distribution (Guisan *et al.*, 2002), or snow depth (López Moreno & Nogues Bravo, 2006). It has also been used to analyze land abandonment (Verburg *et al.*, 2002; Dirnböck *et al.,* 2003; Bakker *et al.*, 2005; Serra *et al.*, 2008; Corbelle-Rico *et al.*, 2012).

In this case, we used a stepwise regression method. This begins with a saturated model with all variables; variables are then recursively eliminated if their significance to the model is low. Once the stepwise process has finished, predictors with p-value less tan 0.01 are rejected one at a time.

b) Random Forest

Decision trees (Breiman *et al.*, 1984) are among the best known supervised classification methodologies. They conform a non-parametric, robust, and non-sensitive to missing or noisy data prediction technique (Schmidt *et al.*, 2008) whose decision rules are easy to interpret. However, the main problem of classifying with just one tree is its high sensitivity to the input data, because small modifications in the dataset can lead to very different models. Ensemble learning techniques have received much attention as a way to overcome this limitation.

Random Forest is an ensemble method proposed by Breiman (2001). It consists of using several decision trees (500 to 2,000) that are trained using a random subset of cases (obtained by bootstrapping) and a randomized subset of the features. Trees are grown to maximum size without pruning, and then each new case is classified by a voting system among all the trees. The randomness added to the process decreases the correlation between trees, and the large number of trees reduces generalization error (Breiman, 2001; Pal, 2005; Prasad *et al.*, 2006), providing better results than other classification methods (Breiman, 2001; Liaw & Wiener, 2002).

Random Forest algorithm uses two parameters: Number of trees and number of variables used to train each tree; however, it is not very sensitive to the particular values used (Liaw & Wiener, 2002; Hastie *et al.*, 2008). So, in this study, the default values were used for the parameters: 500 trees and a number of variables in each tree equal to the integer part of the square root of the number of features.

The main drawback of Random Forest, comparing with the single classification tree approach is that it becomes a “black box” approach (Prasad *et al.*, 2006). However, Random Forest provides several metrics that allows the model to be interpreted. Variable importance is evaluated based on how the prediction would change if the data for that predictor were randomly permuted. Several statistics can be used as estimator of variable importance. In this study, the mean decrease in accuracy for each predictor was used. Thus, Random Forest is much more interpretable than other machine learning methods, and it has been called a “gray box” approach (Prasad *et al.*, 2006).

The output of a Random Forest classification model might be a hard classification of abandonment and non abandonment or a fuzzy approach that uses the percentage of trees with a positive output as a measurement of the potential of abandonment.

Random Forest has been used in Remote Sensing (Guhimre *et al.*, 2010), Genetics (Cutler & Stevens, 2006), Ecology (Cutler *et al*., 2007), Soil Science (Schmidt *et al.*, 2008) or groundwater characterization (Baudron *et al.,* 2013).

The R package randomForest (Liaw & Wiener, 2002) was used in this study.

c) Integration of Random Forest and Logistic Regression

In this study, we used the importance of variables, one of the outputs of the Random Forest model, to support feature selection both for Random Forest and Logistic Regression in order to obtain a set of features that may give an interpretable model with the latter and a model with higher predictive power with Random Forest.

Sampling

The whole analysis was carried out with a cell size of 25x25 meters, meaning more than 18 million cells inside Murcia region. Because of the huge amount of pixels in the raster layers, a random sample of 2,000,000 pixels was obtained to do the variable selection. To guarantee that the feature selection results are relevant to the objectives, these pixels were obtained from areas that appears as cultivated in the MCA map, but are now abandoned.

Classification methods were calibrated using a random subsample of 100,000 pixels taken from the pixels used for the VIF analysis. Because the samples were random, they are not balanced; that is, there are many fewer abandoned cases than non abandoned cases (3.46 % of abandoned pixels in the training dataset and 3.27 % in the validation dataset). Finally, a different subset of 100,000 pixels was randomly sampled to build ROC curves (James *et al*., 2013) in the validation stage.

Validation

Both classification methods give a probability of abandonment value as output. However, these values should not be considered as a direct probability predictor, taking 0.5 as a threshold to perform a hard classification, because classes are not balanced. In this cases it is better to use a threshold that maximises prediction accuracy (Kuhn & Johnson, 2014). Such threshold also depends on the cost associated with each type of error (false positives and false negatives). However, in this study, the objective was not to obtain a hard classification but to obtain a map of potential of abandonment that would provide its users with the flexibility to set the threshold depending on their needs. In such cases, accuracy statistics as the percentage of correctly classified pixels or the kappa index (Congalton & Green, 2008) are not adequate.

Receiver Operating Characteristic (ROC) curves (James *et al.,* 2013) can be used to quantify the accuracy of the model in such circumstances. These curves (Figure 2) are obtained by plotting, for each threshold value, the sensitivity (true positive rate) and the specificity (the complementary of the false positive rate) of the model. The result is a convex curve from the bottom-left corner of the graph to the upper-right corner (point 1,1). If the curve resembles a straight line the classification is not different from a random one. If the convexity of the curve increases, approaching the point where both sensitivity and specificity are equal to 1, the classification is better. The usual method to quantify this shape is the Area Under the Curve (AUC) which can be interpreted as the probability of correctly classify a pair of random cases, one with Y=1 and the other with Y=0. AUC values range between 0.5 (or even lower), indicating that the classification is not different from a random one (or even worst); and 1, indicating a perfect classification for the optimum threshold. The R package ROCR (Sing *et al*., 2005) was used to calculate ROC curves.

RESULTS

Land abandonment map

After applying the revision process to the SIOSE map, only 17.12 % of the area labelled as "pastures formerly cultivated" was actually abandoned land. On the other hand, 1.1 % of the cultivated area according to the SIOSE map was actually abandoned. Figure 1 shows the obtained land abandonment map; 73.3 % of the identified abandoned plots were included as pastures formerly cultivated in the SIOSE map and the remaining 26.7 % were identified analysing the area that was labelled as cultivated in the map. According to Figure 1, 4.28 % of the cultivated surface in 2001 had been abandoned in 2009, representing a 3.27 % of the whole region.

Variance Inflation Factors

Table 1 shows the 42 quantitative features selected by the VIF-based process of feature selection. These features summarise the environmental variability in the study area, preventing collinearity, and maintaining the interpretability of both the variables and the models. The aforementioned 3 qualitative variables (land use, soil type and lithology) were added to the data set, making a total of 45 predictors.

Classification models

Figure 2 shows the results of both classification models calibrated with the 45 features resulting from the variable selection process. It is clear that Random Forest model has the highest prediction capacity. Figure 3 shows the 30 more important features according with the Random Forest algorithm. It is clear that the climatic features are more important than the geomorphometric features.

Although the VIF test allowed a substantial reduction in the number of features, a model with 45 variables is still too large because of the risk of overfitting the training data and the difficulty to interpret the model from an environmental point of view. To discover how many features are really needed to fit a model with high enough predictive power, we decided to calibrate Random Forest with an increasing number of variables, beginning with the most important one, until a model with the 20 most important variables was reached. The results of this calibration process appear in Figure 4. Although only the 20 most important features were analysed, a sill is rapidly reached using just 13 or even 6 features. This result demonstrates that an accuracy similar, or even better, to the obtained with the complete set of features can be reached selecting just the most important features.

The problem with variable importance is that it does not provide information on the effect of the features on the model. In order to understand what role the different features play, we repeated the sequential calibration process with a logistic regression. Features, once more, were entered into the model according to their importance measured by Random Forest; the difference is that features were maintained in the model or omitted according to two criteria: The p-value of the effect had to be smaller than 0.05, and the effect of the feature on the model should not be counter to accepted knowledge. Although this decision may not be considered sensible, we think it is better to be safe not including a dubious predictor and to avoid reaching the wrong conclusions. Features in black Figure 4 were rejected in this sequential process; while features in red were maintained in the model.

Having a subset of features that both are important for prediction and have an environmentally sound effect on the logistic model, we tried to asses the predictive capacity of such model and repeat the sequential calibration process, but this time for both the Random Forest and the logistic model. Figure 5 shows the importance of variables in this reduced model. Figure 6 shows the increase in the area under the ROC curve with the addition of new features in order of importance. Once more, it is possible to reduce the number of features used in the model to a final model with just 8 features without loss in accuracy (AUC=0.917).

During the sequential calibration process, the sign of the coefficients and p-values of the predictors in the logistic regression model were quite stable when each new variable were introduced. It is noteworthy that, for the logistic model, although the AUC value is quite lower (AUC=0.748), there is not much difference between calibration and validation data and the curves are monotone increasing.

Table 2 and Figure 7 show the effect of the six quantitative variables finally selected. All of them are highly significant, with really small p-values. The estimated coefficients show the direction and magnitude of the effects in the linear part of the logistic regression. Figure 7 shows the effects on potential of abandonment and the confidence intervals. Table 3 and figures 8 and 9 show the effects and confidence intervals of the different land use classes and soil types. Finally, Figure 10 shows the final map using the reduced Random Forest model with an AUC of 0.917.

DISCUSSION

The feature set resulting from the VIF analysis should be interpreted as a subset of the original dataset that summarizes the original information reducing, at the same time, collinearity. This means that every feature in the subset does not only represent itself, but also all the features correlated with it that are not included in the subset. For example, precipitation in December can be considered a proxy for winter precipitation.

Regarding Random Forest importance, climatic features were seen to be more important than the geomorphometric features. One reason that might explain this is that, in Murcia Region, climate is, indeed, more relevant than geomorphology; however, it is also possible that farmers perceive climatic factors more easily than other factors. Murcia Region is a semi-arid area that suffers frequent drought periods. In the time span we are studying, the year 2005 was characterized by a significantly shorter amount of rainfall. Although the climatic layers were build with long series, not just with the data for the 2001-2009 period, this fact can partly explain the strong importance reached by climatic variables.

Among the climatic features, October and December precipitation have a clearly negative correlation with abandonment. It should be noted that Murcia region receives precipitation both from winter frontal systems coming from the Atlantic Ocean and from convective cells generated during the autumn months in the Mediterranean Sea. Both mechanisms are quite uncorrelated. Thus, precipitation in October and December act as proxies for these separate mechanisms.

Being a semi-arid region, water availability is a very important issue. The spatial differences in precipitation at regional scale may represent the difference among profitability or not profitability. This is especially important to explain land abandonment in dryland areas. Although irrigated plots have other water sources, they also depend partly on precipitation. So, in both cases, areas receiving less rainfall, both in winter and in autumn, are more prone to abandonment.

On the other hand, the effect of precipitation in September is positive. This effect can be interpreted as the effect of late summer convective storms in the interior of the study area. This kind of precipitation events can produce substantial damages and economic losses in cultivated areas. Another positive effect is that of August absolute maximum temperature. Since Murcia is one of the hottest and driest areas in Europe, this positive effect is related with the amount of heat and potential evapotranspiration in summer, being a factor that leads to land abandonment.

Among the geomorphometric features, only the slope calculated at a maximum scale (39 pixels, that is, around 95 ha) was introduced in the final model. In this case the steeper the slope the more likely the abandonment. We are aware that the selected features are relevant to the local, agricultural plot, scale we are working in and that, at more detailed scales, geomorphometric variables might be more relevant. The positive correlation between slope and abandonment has been already stated by several researchers (e.g. Koulori *et al*., 2007).

Distance to natural vegetation areas is the final quantitative feature included in the model; it has also a negative effect; however, its uncertainty is larger than in the other cases. Although the p-value is quite small, the confidence interval of the effect seems to be very large. The interpretation is that the probability of abandonment of cultivated plots very near natural land cover (less than 500 m) is slightly higher than in plots farther from non cultivated areas. Similar effects have been found by Bieling (2013).

Land use has also been considered by different authors (Dunjó *et al*., 2003; Lesschen *et al.,* 2007; García Ruíz, 2010; Nadal Romero *et al*., 2013) as an important feature in relation with land abandonment. Figure 8 shows the effects of the different uses on the model. In general, rainfed crops are more prone to abandonment than the irrigated crops. Murcia is an arid region and the supply of water in rainfed areas is far from guaranteed. In a first group, irrigated fruit trees and vineyards are the most profitable crops and thus the less prone to abandonment. In a second group, carob trees (a very marginal crop) and rainfed grass crops are the most prone to abandonment; however the confidence intervals are quite large in these cases. Finally, the "Other crops" group includes a small number of plant nurseries and palm trees that are not abandoned but, because of their small number, the effect shows large uncertainty.

Figure 9 and Table 3 show the effects of different soil types; Table 3 includes the keys to interpret the codes in figure 9. Most of the soil types whose effect on abandonment is positive correspond to soils with different limitations for agriculture. Arenosols are poorly developed soils with a low nutrient content. Solonchaks are highly saline soils, and in Murcia Region they are associated with highly saline marls and clays deposited during the Keuper period. The dispersive nature of soils developed on marls and some land use practices, such as levelling and terracing, increase the risk of piping erosion processes (Romero-Díaz *et al*., 2007). Vertisols are soils with a high proportion of swelling clays, and, as a result, deep wide cracks can appear when drying. Regosols are poorly developed soils formed from unconsolidated materials. They appear mostly in loamy depressions where water erosion processes occur (Romero-Díaz *et al*., 2007). Lithosols are very shallow soils (less than 10 cm). By contrast, it is noteworthy that calcium Fluvisols, typically alluvial and fertile soils, were seen to be prone to abandonment. The cause is probably not the soil characteristics itself, but its occurrence in areas of recurrent flooding, one of the most important environmental hazards in Murcia region (Romero-Díaz & Maurandi Guirado, 2000; Castejón Porcel & Romero Díaz, 2014).

Figure 4 shows that the accuracy in the final model with 8 features is almost perfect for the Random Forest model and calibration data; with validation data its AUC is 0.917, slightly better than when using the 45 variables (Figure 2). In relation with the logistic model, the increase in the area under the ROC curve for validation and calibration data are both monotone increasing and almost identical. This surprising result can be explained because each new predictor is actually adding valuable information to the model without overfitting it.

Figure 10 shows the prediction obtained by Random Forest using the 8 selected features. This map reflects quite accurately the land abandonment map obtained by photointerpretation (Figure 1). The lower potential of abandonment appears in irrigated areas in the south of the study area where water availability combines with fertile soils; and in the mountain areas of the northwest where precipitation is higher and temperatures milder. The highest abandonment potential appears in the northeast, where climate is more continental, arid and extreme.

Land abandonment is a very complex problem with several underlying factors affecting differently at different spatio-temporal scales. Due to the specific location, temporal span, scale, and type of features analyzed, we do not think that the results obtained and the conclusions that we draw from them are universally valid. Socioeconomic pressures are probably more important at a broader (regional) scales. Besides, geomophologic features might have been more relevat in the past, or even nowadays at a more detailed (larger) scales.

The results show that our approach may be quite useful to disect the multidimensional problem of land use change analysing a subspace of it. The methodology has been proven useful, and can be easily implemented in any other area and can also be adapted to the amount of information available.

CONCLUSIONS

The Variance Inflation Factor methodology allowed a substantial reduction in the number of predictors from 182 to 45. These predictors summarize all the information in the initial dataset using a subset of variables that are not correlated with each other.

Combining feature importance obtained by Random Forest with logistic regression information enabled us to reduce this quantity to a very tractable subset that can produce a model with high predictive power and, at the same time guarantee that the features included into the model are environmentally interpretable.

In Murcia, climate is the main factor to explain land abandonment.

ACKNOWLEDGEMENTS

This study was carried out within the framework of the research project 15233/PI/10, funded by Fundación Séneca. We also thank the two anonymous reviewers who have contributed to substantially improve this manuscript.

REFERENCES

Aide MT, Clark ML, Grau HR, López-Carr D, Levy MA, Redo D, Bonilla-Moheno M, Riner G, Andrade-Núñez MJ, Muñiz M. 2012. Deforestation and reforestation of Latin America and the Caribbean (2001-2010), *Biotropica* **45**: 262-271. DOI: 10.1111/j.1744-7429.2012.00908.x

Alados CL, Puigdefábregas J, Martínez-Fernández J. 2011. Ecological and socio-economical thresholds of land and plant-community degradation in semi-arid Mediterranean areas of southeastern Spain. *Journal of Arid Environments* **75** (12): 1368-1376. DOI: 10.1016/j.jaridenv.2010.12.004

Al-Awadhi J M. 2013. A case assessment of the mechanisms involved in human-induced land degradation in northeastern kuwait. *Land Degradation and Development* **24**: 2-11. DOI 10.1002/ldr.1090

Alias L, Ortiz R. 1986–2004. *Memorias y mapas de suelos de las hojas del MTN a escala 1:50.000*. Proyecto LUCDEME. Ministerio de Medio Ambiente.

Allen R., Smith M., Perrie A., Pereira L. 1994. An update for the calculation of reference evapotranspiration. *ICID Bulletin* **43(2)**: 35-92.

Arbelo CD, Rodriguez-Rodriguez A, Guerra JA, Mora JL, Notario JS, Fuentes F. 2006. Soil degradation processes and plant colonization in abandoned terraced fields overlying pumice tuffs. *Land Degradation and Development* **17**: 571-588. DOI: 10.1002/ldr.735.

Arnaez J, Lasanta T, Errea, MP, Ortigosa L. 2011. Land abandonment, landscape evolution, and soil erosion in a spanish mediterranean mountain region: The case of Camero Viejo. *Land Degradation and Development* **22**: 537-550. DOI: 10.1002/ldr.1032.

Arozarena A, Villa G, Valcárcel N, Peces JJ, Domenech E, Porcuna A. 2006. New concept on Land Cover/Land Use information system in Spain. Design and Production. In *Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover*.

Bakker, M.M., Govers, G., Kosmas, C., Vanacker, V., van Oost, K., Rounsevell, M. (2005): Soil erosion as a driver of land-use change, *Agriculture, Ecosystems & Environment*, **105(3):** 467-481. doi:10.1016/j.agee.2004.07.009

Bakker MM, Govers G, Van Doorn A, Quetier F, Chouvardas D, Rounsevell M. 2008. The response of soil erosion and sediment export to land-use change in four areas of Europe: the importance of landscape pattern. *Geomorphology* **98:** 213–226. DOI: 10.1016/j.geomorph.2006.12.027

Baudron P, Alonso-Sarría F, García-Aróstegui J, Cánovas-García F, Martínez-Vicente D, Moreno-Brotóns J. 2013. Identifying the origin of groundwater samples in a multi-layer aquifer system with random forest classification, *Journal of Hydrology* **499:** 303–315.

Beilin R, Lindborg R, Stenseke M, Pereira HM, Llausàs A, Slätmo E, Cerqueira Y, Navarro L, Rodrigues P, Reichelt N, Munro N, Queiroz C. 2014. Analysing how drivers of agricultural land abandonment affect biodiversity and cultural landscapes using case studies from Scandinavia, Iberia and Oceania. *Land Use Policy* **36**: 60-72. DOI: http://dx.doi.org/10.1016/j.landusepol.2013.07.003

Bellin N, Van Wesemael B, Meerkerk A, Vanacker V, Barberá GG. 2009. Abandonment of soil and water conservation structures in Mediterranean ecosystems. A case study from southeast Spain. *Catena* **76**: 114–121. DOI: 10.1016/j.catena.2008.10.002

Beven K, Lamb R, Quinn P, Romanowicz R, Freer J. 1995. TOPMODEL. In V P Singh (ed.), *Computer Models of Watershed Hydrology*. Water Resour. Publ., pages 627–668.

Bieling C. 2013. Perceiving and Responding to Gradual Landscape Change at the Community Level: Insights from a Case Study on Agricultural Abandonment in the Black Forest, Germany. Ecology and Society, **18(2)**:36

Blanco-Bernardeau A., Alonso-Sarría F., Gomariz-Castillo F. 2014. Elaboración de un mapa de carbono orgánico del suelo en la Región de Murcia, *XVI Congreso Nacional de Tecnologías de la Información Geográfica*, Universidad de Alicante. https://www.researchgate.net/publication/271201680_Elaboracin_de_un_mapa_de_carbono_orgnico_del_suelo_en_la_Regin_de_Murcia

Breiman L, Friedman JH, Olshen RA, Stone CJ. 1984. *Classification and Regression Trees*. Chapman and Hall/CRC, 368 pages. ISBN-13: 978-0412048418.

Breiman L. 2001. Random Forests. *Machine Learning* **45**: 5-32. DOI: 10.1023/A:1010933404324

Brevik EC, Cerdà A, Mataix-Solera, J, Pereg, L, Quinton, JN, Six J and Van Oost K.. 2015. The interdisciplinary nature of SOIL, *SOIL*, **1**: 117-129. DOI:10.5194/soil-1-117-2015

Cammeraat E, Cerdà, A and Imeson AC. 2010. Ecohydrological adaptation of soils following land abandonment in a semiarid environment. *Ecohydrology* **3:** 421-430. DOI:10.1002/eco.161

Calatrava J, Barberá GG, Castillo VM. 2011. Farming practices and policy measures for agricultural soil conservation in semi-arid mediterranean areas: the case of the Guadalentín Basin in Southeast Spain. *Land Degradation & Development* **22**: 58-69. DOI: 10.1002/ldr.1013

Castejón Porcel G, Romero Díaz A. 2014. Inundaciones en la Región de Murcia en los inicios del Siglo XXI. Biblio 3W. Revista Bibliográfica de Geografía y Ciencias Sociales. Vol. XIX, nº 1102 http://www.ub.es/geocrit/b3w-1102.htm

Cazcarro I, Duarte R, Martín-Retortillo M, Pinilla V, Serrano A. 2014. Water scarcity and agricultural growth in Spain: From curse to bleesing? Working paper of the Spanish Society of Economic History. http://www.aehe.net/2014/10/dt-aehe-1419.pdf

Cerdà A. 1997. Soil erosion after land abandonment in a semiarid environment of southeastern Spain *Arid Soil Research and Rehabilitation* **11(2):** 163 - 176 http://www.tandfonline.com/doi/abs/10.1080/15324989709381469

Cerdà A, Giménez A, Burguet M, Arcenegui V, González FA, García-Orenes F, Pereira P. 2012. El impacto del cultivo, el abandono y la intensificación de la agricultura en la pérdida de agua y suelo. El ejemplo de la vertiente norte de la Serra Grossa en el este peninsular. *Cuadernos de Investigación Geográfica* **38**: 75-94.

Colino J, Martínez-Carrasco, F and Martínez-Paz, J.M. 2014. El impacto de la PAC renovada sobre el sector agrario de la Región de Murcia. CES 207 pages. https://www.researchgate.net/publication/269747214_El_impacto_de_la_PAC_renovada_sobre_el_sector_agrario_de_la_Regin_de_Murcia

Congalton RG, Green K. 2008. *Assessing the Accuracy of Remotely Sensed Data Principles and Practices*. CRC Press, 2 edn., 183 pages.

Corbelle-Rico E, Crecente-Maseda R, Santé-Riveira I. 2012. Multi-scale assessment and spatial modelling of agricultural land abandonment in a European peripheral region: Galicia (Spain), 1956-2004, *Land Use Policy*, **29**: 493-501. doi:10.1016/j.landusepol.2011.08.008

Cortina García J. 1994. La agricultura murciana antes y después del mercado común 1975-1992. *Murcia: Serie Estudios*, 12, Consejería de Agricultura, Ganadería y Pesca de la CARM, 169 pp.

Cosandey C, Andréassian V, Martin C, Didon-Lescot JF, Lavabre J, Folton N, Mathys N, Richard D. 2005. The hydrological impact of the Mediterranean forest: a review of French research. *Journal of Hydrology* **301**: 235–249.

Cramer VA, Hobbs RJ and Standish RJ. 2008. What's new about old fields? Land abandonment and ecosystem assembly. *Trends in Ecology & Evolution* **23:** 104-112

Cutler D, Edwards T, Beard K, Cutler A, Hess K, Gibson J, Lawler J. 2007. Random forest for classification in ecology, Ecology **88:** 2783–2792.

Deng L, Shangguan ZP, Sweeny S. 2012. Changes in soil carbon and nitrogen following land abandonment of farmland on the Loess Plateau, China. *PLOS ONE* 8(8). DOI:10.1371/journal.pone.0071923

Desprats JF, Raclot D, Rousseau M, Cerdan O, Garcin M, Le Bissonnais Y, Ben Slimane A, Fouche J, Monfort-Climent D. 2013. Mapping linear erosion features using high and very high resolution satellite imagery. Land Degradation and Development **24**: 22- 32. DOI 10.1002/ldr.1094

Detsis V. 2010. Placing land degradation and biological diversity decline in a unified framework: methodological and conceptual issues in the case of the North Mediterranean region. *Land Degradation and Development* **21**: 413–422. DOI: 10.1002/ldr.980

Dirnböck T, Dullinger S, Grabherr G. 2003. A regional impact assessment of climate and land-use change on alpine vegetation *Journal of Biogeography* **30(3):** 401-417. 10.1046/j.1365-2699.2003.00839.x

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D and Lautenbach S. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography,* **36**: 27‑46.

Dunjó G, Pardini G, GispertM. 2003. Land use change effects on abandoned terraced soils in a Mediterranean catchment, NE Spain. *Catena* **52**: 23–37. DOI: 10.1016/S0341-8162(02)00148-0

El Kateb H, Zhang H, Zhang P, Mosand R. 2013. Soil erosion and surface runoff on different vegetation covers and slope gradients: A field experiment in Southern Shaanxi Province, China. *Catena* **105**: 1-10. DOI: 10.1016/j.catena.2012.12.012

Errea MP, Lasanta T. 2001. Les posibilités de development de l'élevage dans les champs abandonnés à la suite du gel des terres par la PAC en milieu semi-aride : L'exemple du « Campo de Saragosse » (Espagne). *Sud-Ouest Européen* **9**: 75-84.

Fang X, Xue Z, Li B, An S. 2012. Soil organic carbon distribution in relation to land use and its storage in a small watershed of the Loess Plateau, China. *Catena* **88**: 6-13. DOI: 10.1016/j.catena.2011.07.012

Faulkner H, Ruiz J, Zukowskyj P and Downward S. 2003. Erosion risk associated with rapid and extensive clearances on dispersive materials near Sorbas, S.E. Spain. *Environmental Science and Policy* **6:** 115-127

Franco A, Cruz A, Ramírez B. 2012. Cambio tecnológico y tecnología comunitaria en El Valle Morelia-Queréndaro, Michoacán, México. *Revista Mexicana de Ciencias Agrícolas* **3(7):** 1305-1320.

Gallant JC, Dowling TI. 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research* **39**: 14. DOI: 10.4225/08/512EF27AC3888.

García-Ruiz JM. 2010. The effects of land uses on soil erosion in Spain: a review. *Catena* **81**: 1–11. DOI: 10.1016/j.catena.2010.01.001

García-Ruiz JM, Lana-Renault N, Beguería S, Lasanta T, Regüés D, Nadal-Romero E, Serrano-Muela P, López-Moreno JI, Alvera B, Martí-Bono C, Alatorre LC. 2010. From plot to regional scales: Interactions of slope and catchment hydrological and geomorphic processes in the Spanish Pyrenees. *Geomorphology* **120**: 248-257. DOI: 10.1016/j.geomorph.2010.03.038.

García-Ruiz JM, Lana-Renault N. 2011. Hydrological and erosive consequences of farmland abandonment in Europe, with special reference to the Mediterranean region – A review. *Agriculture, Ecosystems and Environment* **140**: 317-338. DOI: 10.1016/j.agee.2011.01.003

Garfi G, Bruno DE, Calcaterra D, Parise M. 2007. Fan morphodynamics and slope instability in the Mucana River basin (Sila Massif, southern Italy): significance of weathering and role of land use changes. *Catena* **69**: 181-196. DOI: 10.1016/j.catena.2006.06.003

Gellrich M, Zimmermann N. 2007. Investigating the regional-scale pattern of agricultural land abandonment in the Swiss mountains: a spatial statistical modelling approach. *Landscape Urban Plan* **79**: 65-76. DOI: 10.1016/j.landurbplan.2006.03.004.

Ghosh A, Sharma R, Joshi PK. 2014. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. Applied Geography **48**: 31-41. DOI: http://dx.doi.org/10.1016/j.apgeog.2014.01.003

Gibson GR. 2012. *War and agriculture: three decades of agricultural land use and land cover change in Iraq*. Blacksburg, VA, 137 pages.

Giupponi C, Ramanzin M, Sturato E, Fuser S. 2006. Climate and land uses changes, biodiversity and agri-environmental measures in the Belluno province, Italy. *Environmental Sciences and Policy* **9**: 163–173.

Gomariz-Castillo F., Alonso-Sarría F. 2013. An R script to model monthly climatic variables with GLM to be used in hydrological modelling. *9th International R User Conference.* Universidad de Albacete, Spain. https://www.researchgate.net/publication/280932279_An_R_script_to_model_monthly_climatic_variables_with_GLM_to_be_used_in_hydrological_modeling_in_River_Segura_basin

Grau HR, Aide MT. 2008. Globalization and land-use transitions in Latin America. *Ecology and Society* 13 (2): 16. DOI: http://www.ecologyandsociety.org/vol13/iss2/art16/

Guhimre B, Rogan J, Miller J. 2010. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the getis statistic. *Remote Sensing Letters* **1**: 45-54.

Guisan A, Edwards, T, Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene, *Ecological Modelling* **157**: 89-100.

Hastie T, Tibshirani R, Friedman J. 2008. *The Elements of Statistical Learning*. Springer, 745 pages.

Hengl T, Reuter HI. 2009. *Geomorphometry: Concepts, Software, Applications.*, Elsevier, 772 pages.

Hofierka J. 1997. Direct solar radiation modelling within an open GIS environment, *Proceedings of JEC-GI'97 conference* in Vienna, Austria, IOS Press Amsterdam, 575-584.

IGN. 2005. *Sistema de Información de ocupación del suelo en España. Documento Técnico SIOSE 2005* Versión 2, D.G. Instituto Geográfico Nacional, Servicio de Ocupación del Suelo, Madrid, Spain.

Iwahashi J, Pike RJ. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* **86**: 409–440. DOI: 10.1016/j.geomorph.2006.09.012.

Jafari R, Bakhshandehmehr L. 2014. Quantitative mapping and assessment of environmentally sensitive areas to desertification in central Iran. *Land Degradation and Development*. DOI: 10.1002/ldr.2227

James G, Witten D, Hastie T, Tibshirani R. 2013. *An introduction to Statistical Learning with Applications in R.* Springer. 426 pages.

Jiménez-Herrero L, Prieto del Campo F, Riechmann-Fernández J, Gómez-Sal A. 2005. *Informe de la Sostenibilidad en España 2005*. Universidad de Alcalá, Observatorio para la Sostenibilidad en España.

Kakembo V, Rowntree KM. 2003. The relationship between land use and soil erosion in the communal lands near Peddie Town, Eastern Cape, South Africa. Land Degradation and Development **14**: 39-49. DOI: 10.1002/ldr.509

Keesstra SD, Bruijnzeel LA, van Huissteden J. 2009. Meso-scale catchment sediment budgets: combining field surveys and modeling in the Dragonja catchment, southwest Slovenia *Earth Surface Processes and Landforms* **34**: 1547-1561. DOI: 10.1002/esp.1846

Keesstra SD, Geissen V, van Schaik L, Mosse K, Piiranen S. 2012. Soil as a filter for groundwater quality. *Current Opinions in Environmental Sustainability* **4**: 507-516. DOI:10.1016/j.cosust.2012.10.007

Keesstra SD, Temme AJAM, Schoorl JM, Visser SM. 2014. A new, simple model for temporal and spatial sediment fluxes in meso-scale catchments: LAPSUS-D: the hydrological functioning. *Geomorphology* **212**: 97-107. DOI:10.1016/j.geomorph.2013.04.021

Khanal NRK, Watanabe T. 2006. Abandonment of Agricultural Land and Its Consequences. A Case Study in the Sikles Area, Gandaki Basin, Nepal Himalayaatanabe. *Mountain Research and Development* **26:**32-40. DOI: http://dx.doi.org/10.1659/0276-4741(2006)026[0032:AOALAI]2.0.CO;2

Kizos T, Koulouri M. 2006. Agricultural landscape dynamics in the Mediterranean: Lesvos (Greece) case study using evidence from the last three centuries. *Environmental Science & Policy* **9**: 330-342.

Kosmas C, Gerontidis S, Marathianoy M. 2000. The effect of land use change on soils and vegetation over various lithological formations on Lesvos (Greece). *Catena* **40**: 51–68.

Kosmas C, Danalatos NG, López-Bermúdez F, Romero-Díaz A. 2002. *The effect of Land Use and Soil Erosion and Land Degradation under Mediterranean Conditions*. In N. A. Geeson, C. J. Brandt y J. B. Thornes (Eds.): Mediterranean Desertification: a mosaic of processes and responses: John Wiley & Sons, 57-70.

Koulouri M, Giourga C. 2007. Land abandonment and slope gradient as key factors of soil erosion in Mediterranean terraced lands. *Catena* **69**: 274–281. DOI: 10.1016/j.catena.2006.07.001

Kuhn M, Johnson K. 2014. *Applied predictive modeling*. Springer.

Kutner MH, Nachtsheim C, Neter J. 2004. *Applied linear regression models*. McGraw Hill. ISBN 9780072386912.

Lasanta T, Arnáez J, Oserín M, Ortígosa L. 2001. Marginal lands and erosion in terraced fields in the Mediterranean mountains: A case in the Camero Viejo (Nothwestern Iberian System, Spain). *Mountain Research and Development* **21**: 69-76.

Lasanta T, Nadal-Romero E, Errea P, Arnaez J. 2014. The effect of landscape conservation measures in changing landscape patterns: a case study in Mediterranean mountains. *Land Degradation and Development*. DOI: 10.10021 ldr.2359

Lasanta-Martínez T, Vicente-Serrano SM, Cuadrat JM. 2005. Mountain Mediterranean landscape evolution caused by the abandonment of traditional primary activities: a study of the Spanish Central Pyrenees. *Applied Geography* **25**: 47-65. DOI: 10.1016/j.apgeog.2004.11.001.

Lasanta-Martínez T, Vicente-Serrano SM, Cuadrat JM. 2005. Mountain Mediterranean landscape evolution caused by the abandonment of traditional primary activities: a study of the Spanish Central Pyrenees. *Applied Geography* **25**: 47-65. DOI: 10.1016/j.apgeog.2004.11.001.

Lenda M, Skórka P, Knops JMH, Moron D, Tworek S, Woyciechowski M. 2012. Plant establishment and invasions: an increase in a seed disperser combined with land abandonment causes an invasion of the non-native walnut in Europe. *Proc. R. Soc. B.* **279**: 1491-1497. DOI: 10.1098/rspb.2011.2153

Lesschen JP, Kok K, Verburg PH, Cammeraat LH. 2007. Identification of vulnerable areas for gully erosion under different scenarios of land abandonment in Southeast Spain. *Catena* **71**: 110-121. DOI: 10.1016/j.catena.2006.05.014

Lesschen JP, Cammeraat LH, Nieman T. 2008. Erosion and terraces failure due to agricultural land abandonment in a semi-arid environment. *Earth Surface Processes and Landforms* **33**: 1574-1584. DOI: 10.1002/esp.1676

Li XL, Gao J, Brierley G, Qiao YM, Zhang J, Yang YW. 2013. Rangeland degradation on the Qinghai-Tibet plateau: Implications for Rehabilitation. *Land Degradation and Development* **24**: 72- 80. DOI 10.1002/ldr.1108

Liaw A, Wiener M. 2002. Classication and Regression by randomForest. *R News* **2(3)**: 18–22.

López-Moreno JI, Nogués-Bravo D. 2006. Interpolating local snow depth data: an evaluation of methods. *Hydrological Processes* **20**: 2217–2232.

Lugo AE, Helmer E. 2004. Emerging forests on abandoned land: Puerto Rico's new forests. *Forest Ecology and Management* **190**: 145-161. DOI: 10.1016/j.foreco.2003.09.012

Maccherini S, Santi E, Bonini L, Amici V, Pruscini S, Palazzo D, Cortés F. 2013. The impact of land abandonment on the plant diversity of olive groves. *Biodiversity Conservetion* **22**: 3067-3083. DOI: 10.1007/s10531-013-0571-8

MacDonald D, Crabtree J R, Wiesinger G, Dax T, Stamou N, Fleury P, Gutiérrez J, Gibon A. 2000. Agricultural abandonment in mountain areas of Europe: Environmental consequences and policy response. *Journal of Environmental Management* **59**: 47-69. DOI: 10.1006/jema.1999.0335.

Mandal D, Sharda VN. 2013. Appraisal of soil erosion risk in the Eastern Himalayan region of India for soil conservation planning. *Land Degradation and Development* **24**: 430-437. DOI 10.1002/ldr.1139

Martínez-Hernández C, Cánovas-García F, Alonso-Sarría F, Romero-Díaz A, Belmonte-Serrato F. 2013. Cartografía de áreas agrícolas abandonadas mediante técnicas de SIG y fotointerpretación Comarcas de la huerta y Campo de Murcia y Alto Guadalentín. In *Espacios insulares y de frontera, una visión geográfica*. Universitat de les Illes Balears, pages 393–403. https://www.researchgate.net/publication/273144061_Cartografa_de_reas_agrcolas_abandonadas_mediante_tcnicas_de_SIG_y_fotointerpretacin._Comarcas_de_la_Huerta_y_Campo_de_Murcia_y_Alto_Guadalentn

McCullagh P, Nelder JA. 1989. *Generalized Linear Models*. Chapman and Hall/CRC.

Melton MA. 1965. The geomorphic and paleoclimatic significance of alluvial deposits in southern Arizona. *Journal of Geology* **73**: 1-38. http://www.jstor.org/stable/30066379.

Moore ID, Burch GJ. 1986. Modelling erosion and deposition: Topographic effects. *Transactions of the American Society of Agricultural and Biological Engineers* **29**: 1624-1630. DOI: 10.13031/2013.30363.

Moravec J, Zemeckis R. 2007. Cross Compliance and Land Abandonment, Deliverable D17 of the CC Network Project, SSPE-CT-2005-022727: 6-16.

Nadal-Romero E, LasantaT, García-Ruíz JM. 2013. Runoff and sediment yields from land under various uses in a Mediterranean mountain area: long-term results from an experimental station. *Earth Surface Processes and Landforms* **38**: 346-355. DOI: 10.1002/esp.3281

Navarro L, Pereira H. 2012. Rewilding Abandoned Landscapes in Europe. *Ecosystems* **15**: 900-912. DOI: 10.1007/s10021-012-9558-7

Neteler M, Mitasova H. 2008. *Open Source GIS: A GRASS GIS Approach*. Springer, New York. 406 pages.

Nunes A, Figueiredo A, Almeida A. 2012. The effects of farmland abandonment and plant succession on soil properties and erosion processes: a study case in centre of Portugal. *Revista de Geografia e Ordenamento do Território* **2**: 165-190.

O'Brien RM. 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity* **41**: 673-690. ISSN 0033-5177.

Olaya V, Conrad O. 2009. Geomorphometry in SAGA. In Tomislav Hengl Hannes I Reuter (eds.), *Geomorphometry*, vol. 33 of *Developments in Soil Science*. Elsevier, pages 141–169.

Padilla A. 1997. *Colonización vegetal en campos de cultivo abandonados en la provincia de Alicante.* Publicaciones de la Universidad de Alicante. Alicante, 365 pp.

Pal M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* **26**: 217-222. DOI: 10.1080/01431160412331269698.

Parsons AJ. 2014. Abandonment of agricultural land, agricultural policy and land degradation in Mediterranean Europe. In E.N. Mueller, J. Wainwright, A. J. Parsons, L. Turnbull (Eds.). *Patterns of Land Degradation in Drylands: Understanding Self-Organised Ecogeomorphic Systems*. Springer: 357-366. DOI: 10.1007/978-94-007-5727-1_14

Piégay H, Walling DE, Landon N, He Q, Liébault F, Petiot R. 2004. Contemporary changes in sediment yield in an alpine mountain basin due to afforestation (the upper Drôme in France). *Catena* **55**: 183–212.

Pinilla V and Serrano R. 2009. Agricultural and Food Trade in the European Community since 1963. Patel K.K. (ed) *Fertile Ground for Europe? The History of European Integration and the Common Agricultural Policy since 1945*. Baden-Baden: Nomos, 273-300.

Pointereau P, Coulon F, Girard P, Lambotte M, Stuczynski T, Sánchez Ortega V, del Río A. 2008. *Analysis of farmland abandonment and the extent and location of agricultural areas that are actually abandoned or are in risk to be abandoned.* Institute for Environment and Sustainability, European Commission, 115 pp.

Prasad A M, Iverson LR, Liaw A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* **9**: 181-199. DOI: 10.1007/s10021-005-0054-1.

Prishchepov AV, Müller D, Dubinin M, Baumann M, Radeloff VC. 2013. Determinants of agricultural land abandonment in post-Soviet European Russia. *Land Use Policy* **30**: 873-884. DOI: http://dx.doi.org/10.1016/j.landusepol.2012.06.011

Proença V, Honrado J, Miguel H. 2012. From Abandoned Farmland to Self-Sustaining Forests: Challenges and Solutions. *Ecosystems* **15**: 881–882. DOI: 10.1007/s10021-012-9557-8

Quinn P, Beven K, Chevallier P, Planchon O. 1991. The prediction of hillslope paths for distributed hydrological modeling using digital terrain models. *Hydrological Processes* **5**: 59-79. DOI: 10.1002/hyp.3360050106.

Quiñonero-Rubio JM, Boix-Fayos C, de Vente J. 2013. Desarrollo y aplicación de un índice multifactorial de conectividad de sedimentos a escala de cuenca. Cuadernos de Investigación Geográfica 39 (2), 203-223. https://publicaciones.unirioja.es/ojs/index.php/cig/issue/view/167

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raiesi F. 2012. Land abandonment effect on N mineralization and microbial biomass N in a semi-arid calcareous soil from Iran. *Journal of Arid Environments* **76**: 80-87. DOI: 10.1016/j.jaridenv.2011.08.008

Ramankutty N, Heller E, Rhemtulla J. 2010. Prevailing myths about agricultural abandonment and forest regrowth in the United States. *Annals of the Association of American Geographers* **100**: 502-512. DOI: 10.1080/00045601003788876

Renwick A, Jansson T, Verburg PH, Revoredo-Giha C, Britz W, Gocht A, McCracken D. 2013. Policy reform and agricultural land abandonment in the EU. *Land Use Policy* **30**: 446-457. DOI: http://dx.doi.org/10.1016/j.landusepol.2012.04.005

Renzi R, Bochicchio M, Bacchi B. 2002. Effects on floods of recent afforestation and urbanisation in the Mella River (Italian Alps). *Hydrology and Earth System Sciences* **6**: 239-253.

Rey Benayas JM, Martins A, Nicolau JM, Schulz J. 2007. Abandonment of agricultural land: an overview of drivers and consequences, *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, **2** (57): 1-14.

Rey Benayas JM, Galván I, Carrascal LM. 2010. Differential effects of vegetation restoration in Mediterranean abandoned cropland by secondary succession and pine plantations on bird assemblages. *Forest Ecol Manag* **260**: 87-95.

Ricotta C, Guglietta D, Migliozzi A. 2012. No evidence of increased fire risk due to agricultural land abandonment in Sardinia (Italy). *Nat. Hazards Earth Syst. Sci.* **12**: 1333-1336. DOI: 10.5194/nhess-12-1333-2012

Ries JB, Hirt U. 2008. Permanence of soil surface crusts on abandoned farmland in the Central Ebro Basin/Spain. *Catena* **72**: 282–296. DOI: 10.1016/j.catena.2007.06.001

Riley SJ, De Gloria SD, Elliot R. 1999. A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity. *Intermountain Journal of Sciences* **5**: 23–27.

Robledano F, Romero-Díaz A, Belmonte-Serrato F, Zapata VM, Martínez-Hernández C, Martínez-López V. 2014. Ecogeomorphological consequences of land abandonment in semiarid Mediterranean areas: Integrated assessment of physical evolution and biodiversity. *Agriculture, Ecosystems & Environment* **197**: 222-242. DOI: 10.1016/j.agee.2014.08.006

Rodríguez-Vaquero JE. 2008. Gestión territorial e inmigración. El caso de Almería. *Nimbus*. **21-22:** 177-187. http://repositorio.ual.es:8080/jspui/bitstream/10835/1526/1/Rodrigu%C3%A9z%20Gesti%C3%B3n.pdf

Romero Díaz A, Maurandi Guirado A. 2000. Las inundaciones en la Cuenca del Segura en las dos últimas décadas del siglo XX. Actuaciones de prevención. Serie Geográfica 9: 93-120.

Romero-Díaz A, Marín-Sanleandro P, Sánchez-Soriano M, Belmonte-Serrato F, Faulkner H. 2007. The causes of piping in a set of abandoned agricultural terraces in Mediterranean Spain. *Catena* **69**: 282–293. DOI: 10.1016/j.catena.2006.07.008551.4(460)(063).

Romero Díaz A, Martínez Hernández C, Belmonte Serrato F. 2012. Cambios de usos del suelo en la Región de Murcia. El almendro como cultivo de referencia y su relación con los procesos de erosión. *Nimbus*. **29-30**: 607-626.

Ruiz Sinoga JD, Martínez Murillo JF. 2009. Hydrological response of metamorphic soil surface components on hillslopes affected by land abandonment along a climatic gradient (Littoral Betic

Range, southern Spain), *Earth Surface Processes and Landforms* **34**: 2047-2056. DOI: 10.1002/esp.1890

Ruiz-Flaño P. 1993. *Procesos de erosión en campos abandonados del Pirineo. El ejemplo del valle de Aisa*. Geoforma Ediciones. logroño. 191 pp.

Sappington JM, Longshore KM, Thompson DB. 2007. Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management* **71**: 1419–1426.

Sauer T, Ries JB. 2008. Vegetation cover and geomorphodynamics on abandoned fields in the Central Ebro Basin (Spain). *Geomorphology* **102**: 267-277. DOI: 10.1016/j.geomorph.2008.05.006

Schmidt K, Behrens T, Scholten T. 2008. Instance selection and classification tree analysis for large spatial datasets in digital soil mapping. *Geoderma* **146**: 138-146. DOI: 10.1016/j.geoderma.2008.05.010.

Schneider L, Geoghegan J. 2006. Land Abandonment in an Agricultural Frontier After a Plant Invasion: The Case of Bracken Fern in Southern Yucatán, Mexico. *Agricultural and Resource Economics Review* **35**: 167–177.

Serra P, Pons X, Saura D. 2008. Land-cover and land-use change in a Mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. *Applied Geography* **28**: 189–209. DOI: 10.1016/j.apgeog.2008.02.001.

Shang ZH, Ma YS, Long RG, Ding JM. 2008. Effect of fencing, artificial seeding and abandonment on vegetation composition and dynamics of 'Black soil land' in the headwaters of the Yangtze and the Yellow Rivers of the Qinghai-Tibetan Plateau. *Land Degradation and Development* **19**: 554-563. DOI: 10.1002/ldr.861

Sing T, Sander O, Beerenwinkel N, Lengauer, T. 2005. ROCR: visualizing classifier performance in R. *Bioinformatics* **21**: 3940-3941. DOI: 10.1093/bioinformatics/bti623.

Sluiter R, De Jong SM 2007. Spatial patterns of Mediterranean land abandonment and related land cover transition. *Landscape Ecology* **22**: 559-576. DOI: 10.1007/s10980-006-9049-3

Suárez-Seoane S, Osborne PE and Baudry J. 2002. Responses of birds of different biogeographic origins and habitat requirements to agricultural land abandonment in northern Spain. *Biological Conservation*, **105**: 333-344.

Symeonakis E, Calvo-Cases A, Arnau-Rosalen E. 2007. Land Use Change and Land Degradation in Southeastern Mediterranean Spain. *Environ Manage* **40**: 80-94. DOI: 10.1007/s00267-004-0059-0

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SSA. 2002. Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model *Environmental Management* **30(3):** 391-405. DOI: 10.1007/s00267-002-2630-x

Walton JC, Lambert DM, Roberts RK, Larson JA, English BC, Larkin SL, Martin SW, Marra MC, Paxton KW, Reeves JM. 2008. Adoption and abandonment of precision soil sampling in cotton production. *Journal of Agricultural and Resource Economics* **33**: 428-448.

Weiss A. 2001. Topographic position and landforms analysis. In *ESRI User Conference*. ESRI, San Diego, CA, USA. Weng Q. 2002. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management* **64**: 273-284. DOI: 10.1006/jema.2001.0509.

White J, Shao, Y, Kennedy LM, Campbell JB. 2013. Landscape dynamics on the Island of La Gonave, Haiti, 1990-2010. *Land* **2**: 493-507. DOI: 10.3390/land2030493

Wood J. 1996. *The Geomorphological characterisation of Digital Elevation Models.* Diss., Department of Geography, University of Leicester, U.K. http://www.soi.city.ac.uk/~jwo/phd/7

Wu JP, Liu ZF, Sun Y X, Zhou LX, Lin YB, Fu SF. 2013. Introduced Eucalyptus urophylla plantations change the composition of the microbial community in subtropical China. *Land Degradation and Development* **24**: 400-406. DOI 10.1002/ldr.2161

Xu EQ, Zhang HQ, Li MX. 2014. Object based mapping of karst rocky desertification using a support vector machine. *Land Degradation and Development*. DOI: 10.1002/ldr.2193

Zakkak S, Kakalis E, Radovic A, Halley JM, Kati V. 2014. The impact of forest encroachment after agricultural land abandonment on passerine bird communities: the case of Greece. *Journal for Nature Conservation* **22**: 157-165. DOI: http://dx.doi.org/10.1016/j.jnc.2013.11.001

Zhao G, Mu X, Wen Z, Wang F, Gao P. 2013. Soil erosion, conservation, and Eco-environment changes in the Loess Plateau of China. *Land Degradation and Development* **24**: 499- 510. 2013. DOI 10.1002/ldr.2246

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, 549 pages.

Zuur AF, Leno EN, Smith GM. 2007. *Analyzing Ecological Data*. Springer, 700 pages.

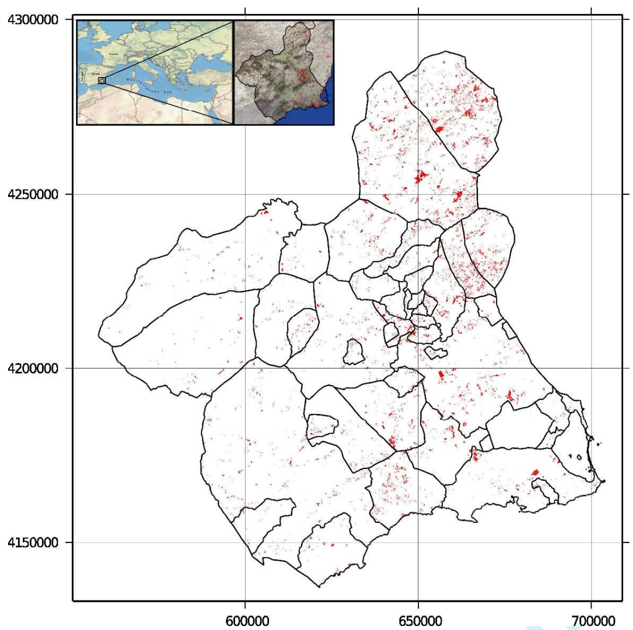


Figure 1: Abandonment map in Murcia Region. The black boundaries correspond to the limits of municipalities.

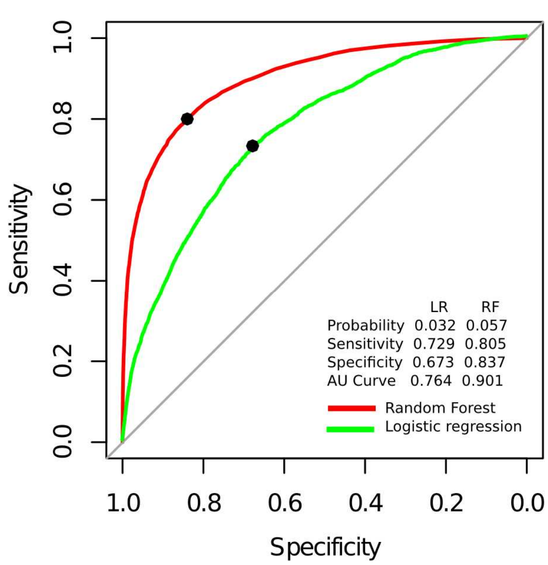


Figure 2: ROC curves for the two classification methods. The 44 features selected by the VIF-based selection process and the 3 qualitative variables were used to fit the models. The optimum points are marked; probability, sensitivity, specificity, and AUC values are included.

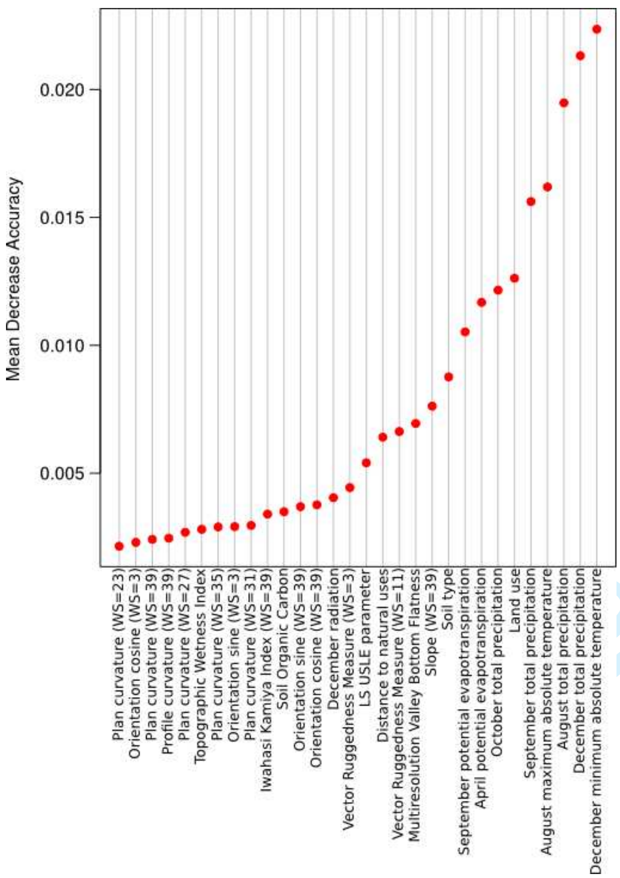


Figure 3: Feature importance in the first Random Forest model. WS stands for window size.

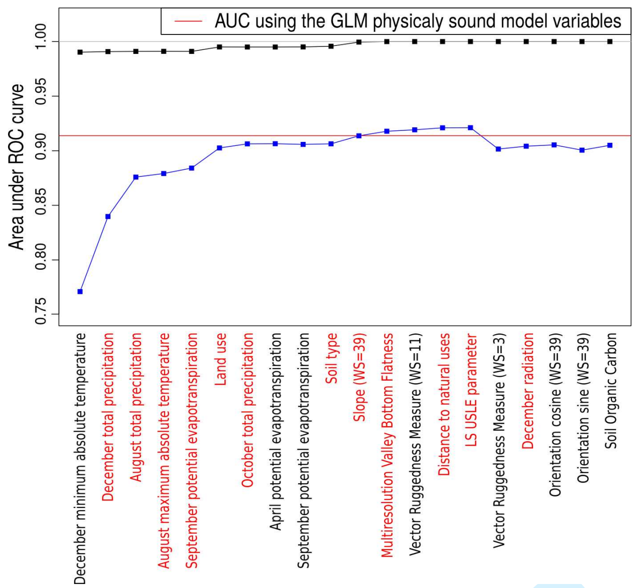


Figure 4: Increase in AUC when adding features to the Random Forest model in order of importance (only the 20 most important). The black line is the accuracy measured using the calibration set and the blue line the accuracy values obtained with the validation set. Features in black were rejected because the p-value was not significant enough or the sign of the effect was counter to accepted knowledge. Features in red were maintained in the model. The red horizontal line shows the accuracy reached using all the features maintained in the model (AUC=0.917). WS stands for window size.

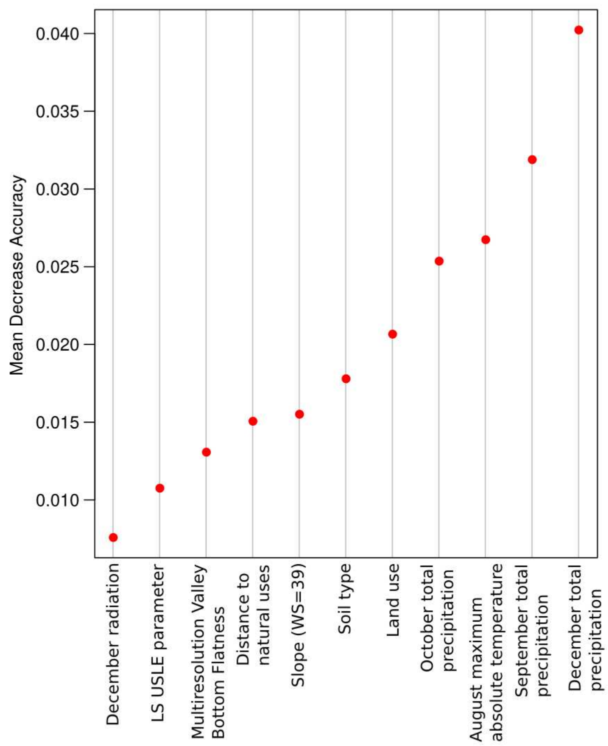


Figure 5: Feature importance in the second Random Forest model. WS stands for window size.

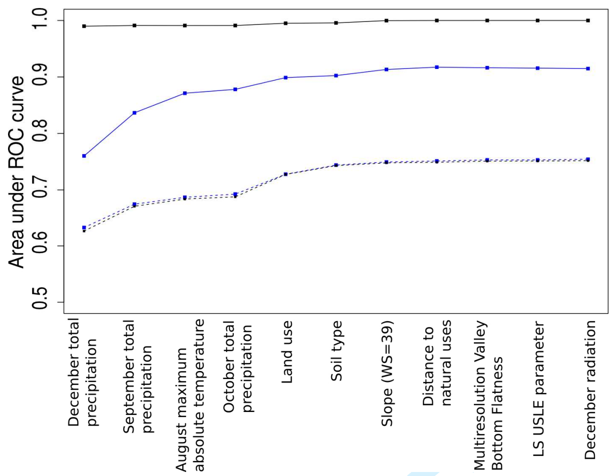


Figure 6: Increase in AUC adding variables to the Random Forest model in order of importance. Dotted lines represent the logistic model and solid lines the Random Forest model. Black lines represent the results for the calibration data, and blue lines the results for the validation data. WS stands for window size.

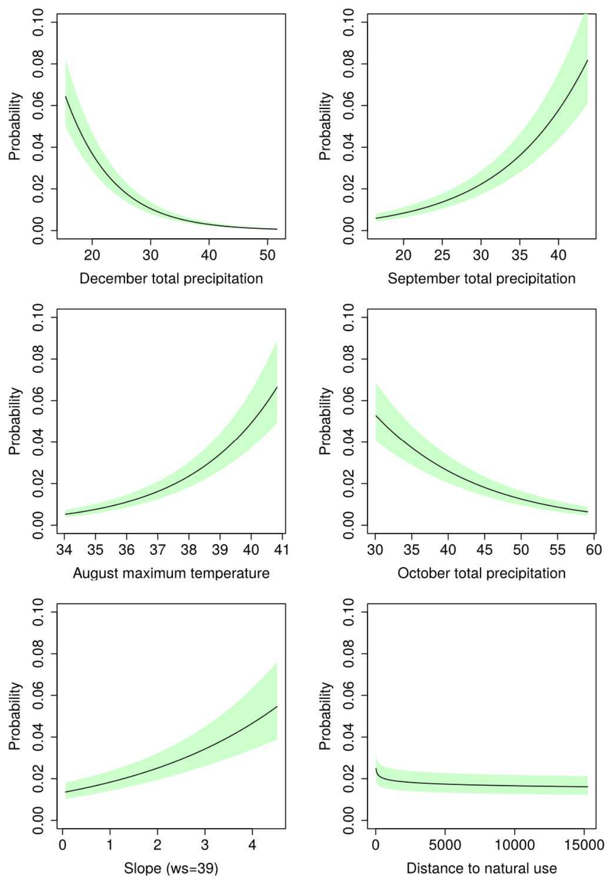


Figure 7: Effects of the quantitative variables. Precipitation is measured in millimeters, temperature in degrees Celsius, slope in degrees and distance to natural land in meters. WS stands for window size.

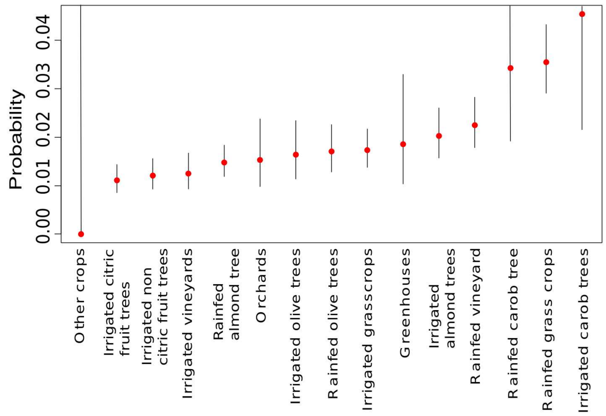


Figure 8: Land use effects.

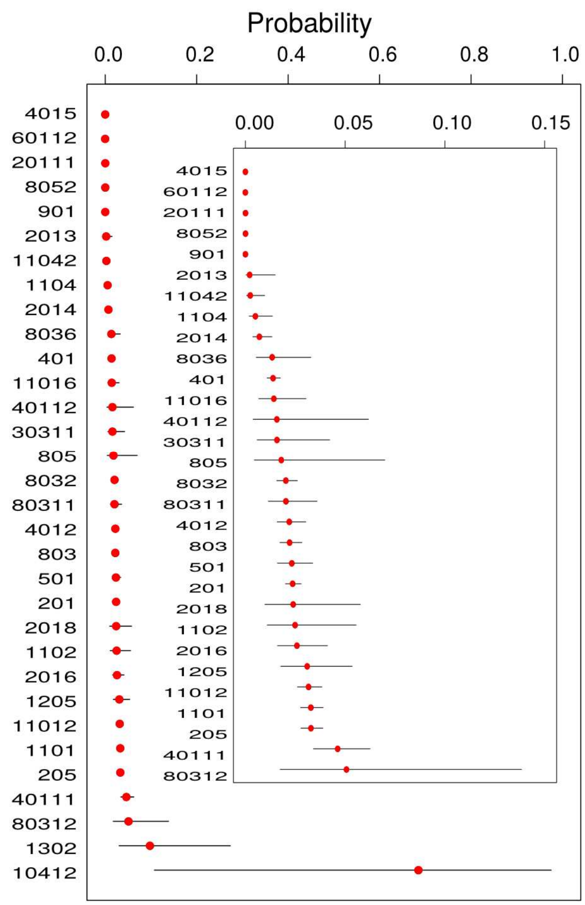


Figure 9. Soil type effects. Due to the large confidence intervals of two of the soil types, a second figure appears inserted in the main one, showing the effects of all soil types except the two with largest effect. This way, the probability range can be reduced to 0-0.15 and the differences in the effects of the different soil types are easier to see. Table 3 includes the key to interpret the soil type numerical codes.

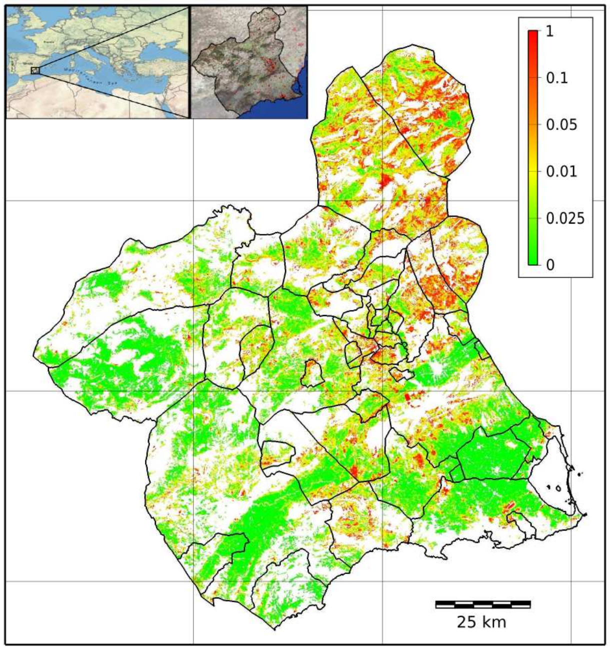


Figure 10: Potential of abandonment map, from the Random Forest model.

Climate	August total precipitation, September total precipitation October total precipitation, December total precipitation, August maximum absolute temperature, December minimum absolute temperature, December radiation, April potential evapotranspiration, September potential evapotranspiration.
Terrain Derivatives	Slope (WS=39), Aspect sine (WS=3), Aspect sine (WS=39), Aspect cosine (WS=3), Aspect cosine (WS=39), Profile curvature (WS=3), Profile curvature (WS=7), Profile curvature (WS=11), Profile curvature (WS=23), Profile curvature (WS=39), Plan curvature (WS=3), Plan curvature (WS=7), Plan curvature (WS=11), Plan curvature (WS=15), Plan curvature (WS=19), Plan curvature (WS=23), Plan curvature (WS=27), Plan curvature (WS=31), Plan curvature (WS=35), Plan curvature (WS=39)
Roughness	Vector Ruggedeness Measure (WS=3), Vector Ruggedeness Measure (WS=11), Iwahashi and Kamiya (WS=3), Iwahashi and Kamiya (WS=7), Iwahashi and Kamiya (WS=11), Iwahashi and Kamiya (WS=39),
Topographic Position	Topographic Position Index (WS=3)
Derived Topo. indices	Multiresolution Valley Bottom Flattness, USLE LS, Topographic Wetness Index
Other	Soil Organic Carbon, Distance to Natural Uses, Distance to roads

Table 1: Variables with VIF less than 5 after the variable selection process. WS stands for window size.

Predictor	Estimate	Std. Error	z value	P-value
(Intercept)	-15.16	1.186	-12.780	<0.000001
December total precipitation	-0.1281	0.005348	-23.959	<0.000001
September total precipitation	0.09919	0.006961	14.249	<0.000001
August maximun abs. temperature	0.3821	0.03	12.737	<0.000001
October total precipitation	-0.07419	0.005183	-14.314	<0.000001
Slope (WS=39)	0.3211	0.04258	7.540	<0.000001
Distance to natural uses	-0.06908	0.01443	-4.789	0.00000168

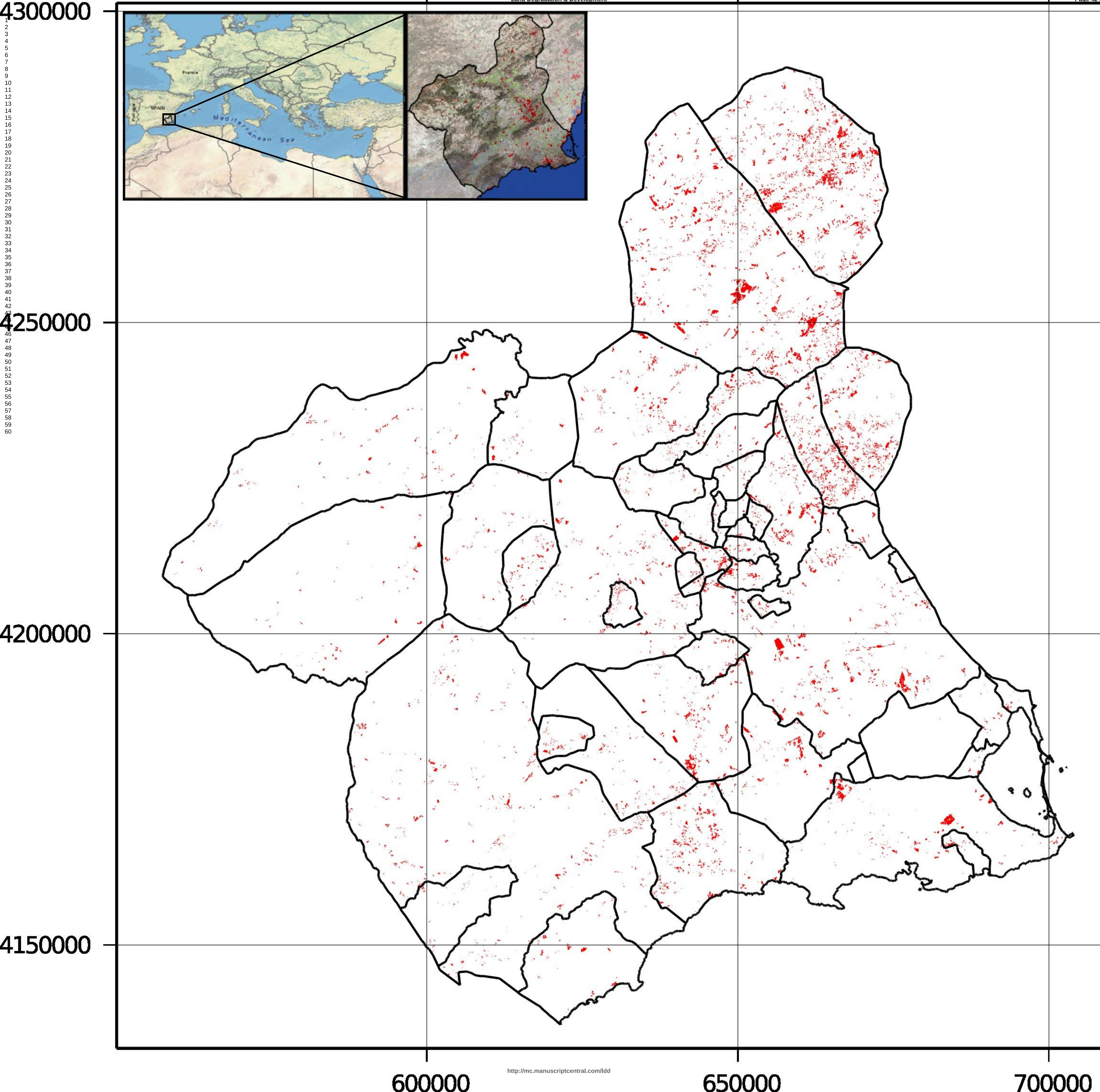
Table 2: Results for the quantitative predictors in the logistic model resulting from the iterative inclusion of predictor. WS stands for window size.

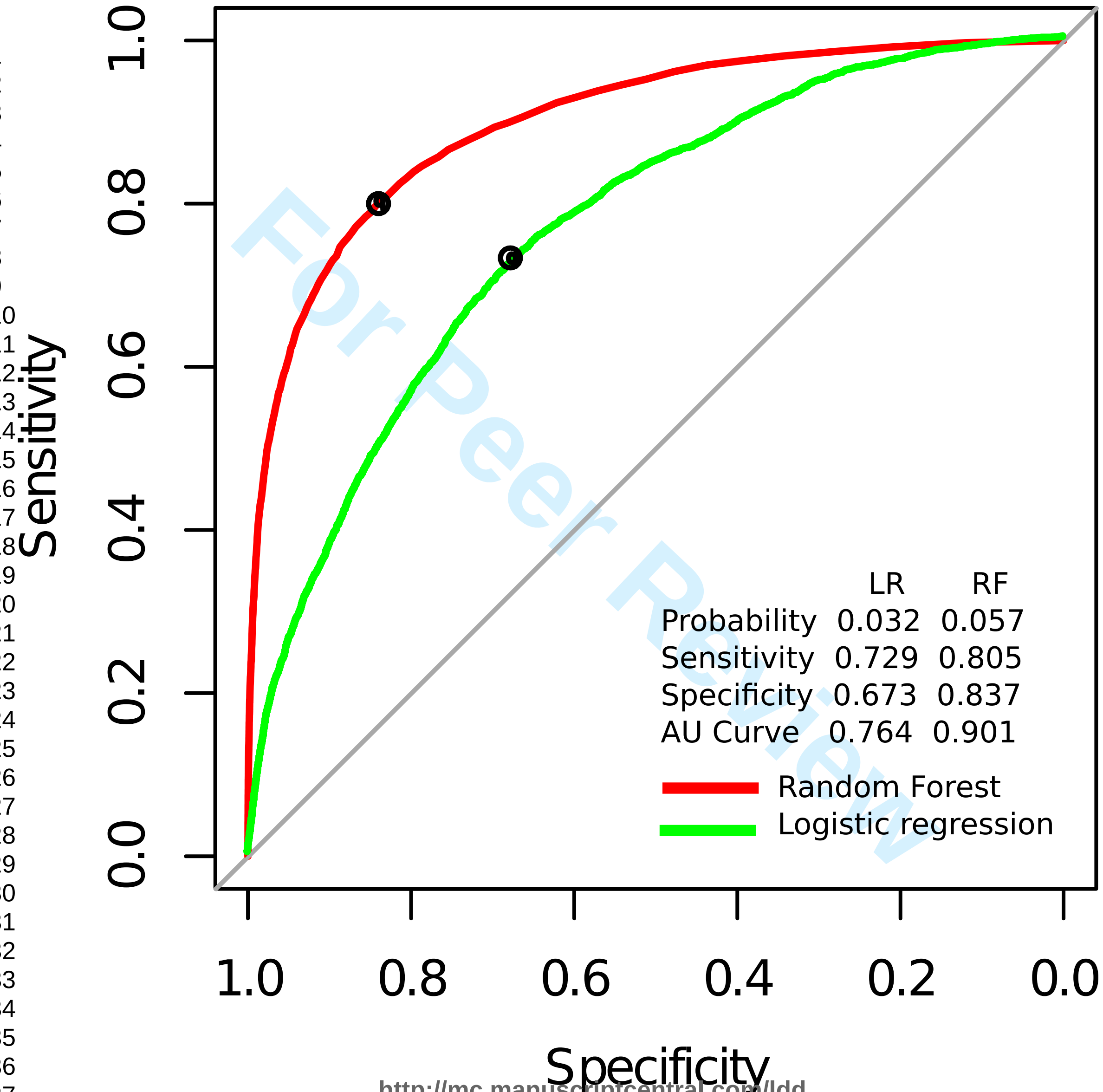
Land use class	Estimate	Std. Error	p-value
Irrigated citric fruit trees			
Irrigated non citric fruit trees	0.0827	0.1152	0.472981
Rain feed olive trees	0.4339	0.1343	0.001233
Rainfed carob trees	1.150	0.2998	0.000125
Rain feed almond trees	0.2896	0.09722	0.002898
Rainfed vineyard	0.7155	0.1060	<0.000001
Rainfed grass crops	1.185	0.08957	<0.000001
Irrigated grass crops	0.4493	0.1018	0.0000103
Orchards	0.3249	0.2199	0.139603
Greenhouses	0.5205	0.2973	0.079962
Other crops	-12.29	360.2	0.972770
Irrigated olive trees	0.3919	0.1751	0.025210
Irrigated carob trees	1.441	0.3879	0.000203
Irrigated almond trees	0.6089	0.1152	<0.000001
Irrigated vineyards	0.1201	0.1368	0.379874

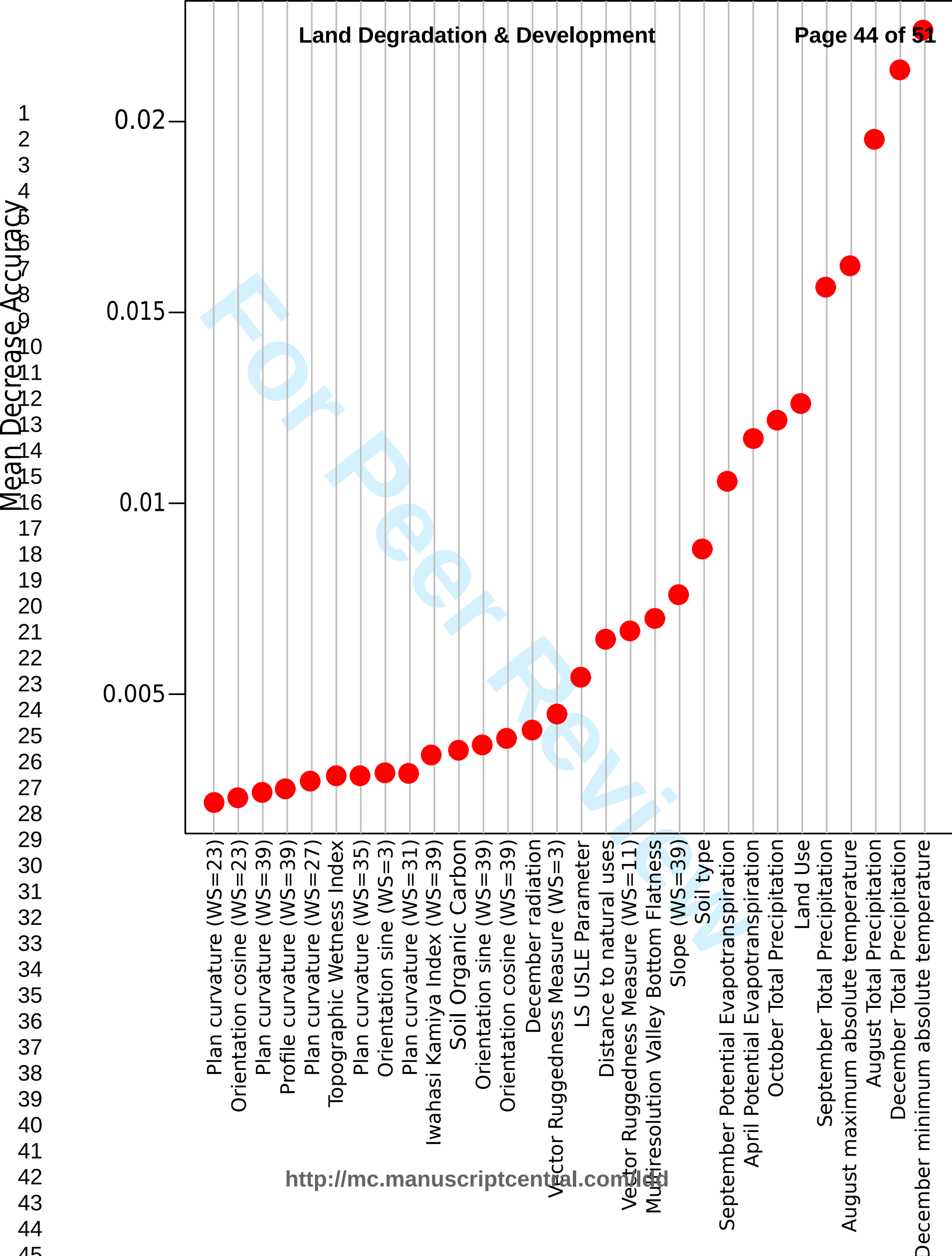
Table 3: Results for land use predictor in the logistic model resulting from the iterative inclusion of predictors. Irrigated citric fruit trees is the baseline level in the model.

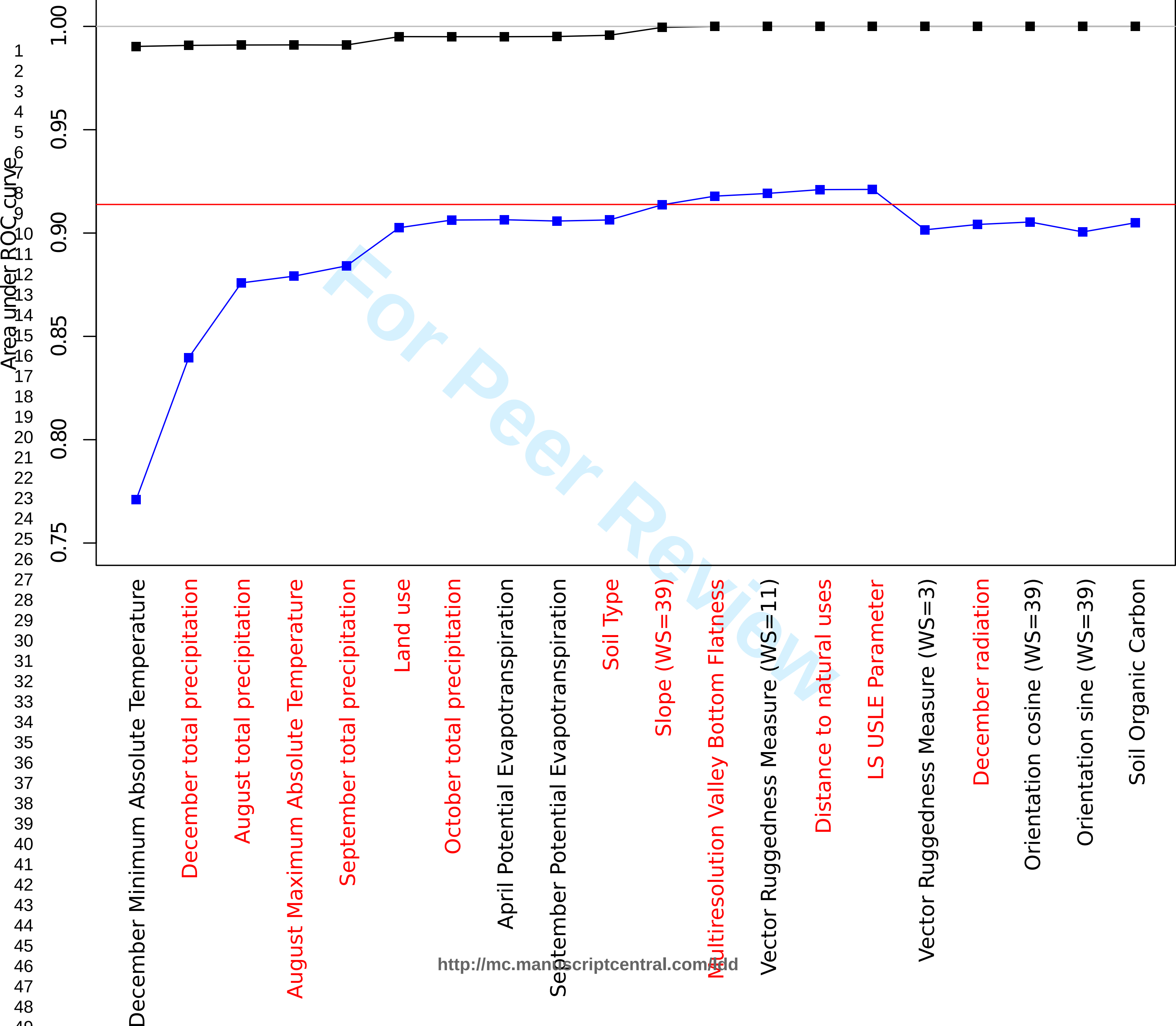
Key	Soil type	Estimate	Std. error	p-value
201	Cambisols and gipsic Xerosols			
205	gypsic Xerosols	0.3364	0.0495	<0.000001
401	calcaric Fluvisols	-0.5499	0.1014	<0.000001
501	calcic Kastanosems	-0.01799	0.1814	0.921005
803	Litosols	-0.06998	0.1052	0.505778
805	Rendzinas	-0.2813	0.7167	0.694684
901	luvic Xerosols	-11.53	38.81	0.976305
1101	calcaric Regosols	0.3361	0.0611	<0.000001
1102	eutric Regosols	0.04914	0.4199	0.906834
1104	litosolic and litorrodic Regosols	-1.573	0.5061	0.001883
1205	gleic and orthic Solonchaks	0.2753	0.2803	0.326009
1302	chromic Vertisols	1.496	0.6307	0.017712
2013	calcic and petrocalcic Cambisols	-2.441	1.002	0.014904
2014	Cambisols, calcic Xerosols and calcaric Fluvisols	-1.233	0.3215	0.000125
2016	calcic and gypsic Xerosols	0.08.697	0.2349	0.711188
2018	calcic Cambisols and orthic Rendzinas	0.01165	0.4603	0.979817
4012	calcaric Fluvisols and calcic and petrocalcic Xerosols	-0.07702	0.1517	0.611611
4015	calcaric Fluvisols and gypsic Xerosols	-13.72	31.56	0.965329
8032	Litosols, Cambisols, Xerosols and Rendzinas	-0.1606	0.1088	0.139978
8036	Litosols and gypsic Xerosols	-0.5804	0.4606	0.207628
8052	aridic Rendzinas and calcic Xerosols	-12.08	413.3	0.976690
10412	albic Arenosols and gleic Solonchaks	4.494	1.474	0.002296
11012	calcaric Regosols, calcic and petrocalcic Xerosols and calcic Cambisols	0.298	0.0721	0.0000358
11016	calcaric Regosols and Xerosoles gpsicos	-0.5159	0.3875	0.183075
11042	litosolic Regosols and calcic Xerosols	-2.312	0.7114	0.001153
20111	calcic Cambisols and calcaric Regosols	-12.29	113.1	0.913486
30311	Cambisols and eutric Regosols	-0.4093	0.5084	0.420689
40111	calcaric Fluvisols and calcaric Regosols	0.6919	0.1471	0.00000258
40112	calcaric Fluvisols and orthic Solonchaks	-0.4163	0.7174	0.561726
60112	calcaric Gleysols and gleic Solonchaks	-13.14	1697	0.993819
80311	Litosols and calcaric and litosolic Regosols	-0.1556	0.2886	0.589825
80312	Litosols and litosolic and gleic Solonchaks	0.7876	0.5569	0.157330

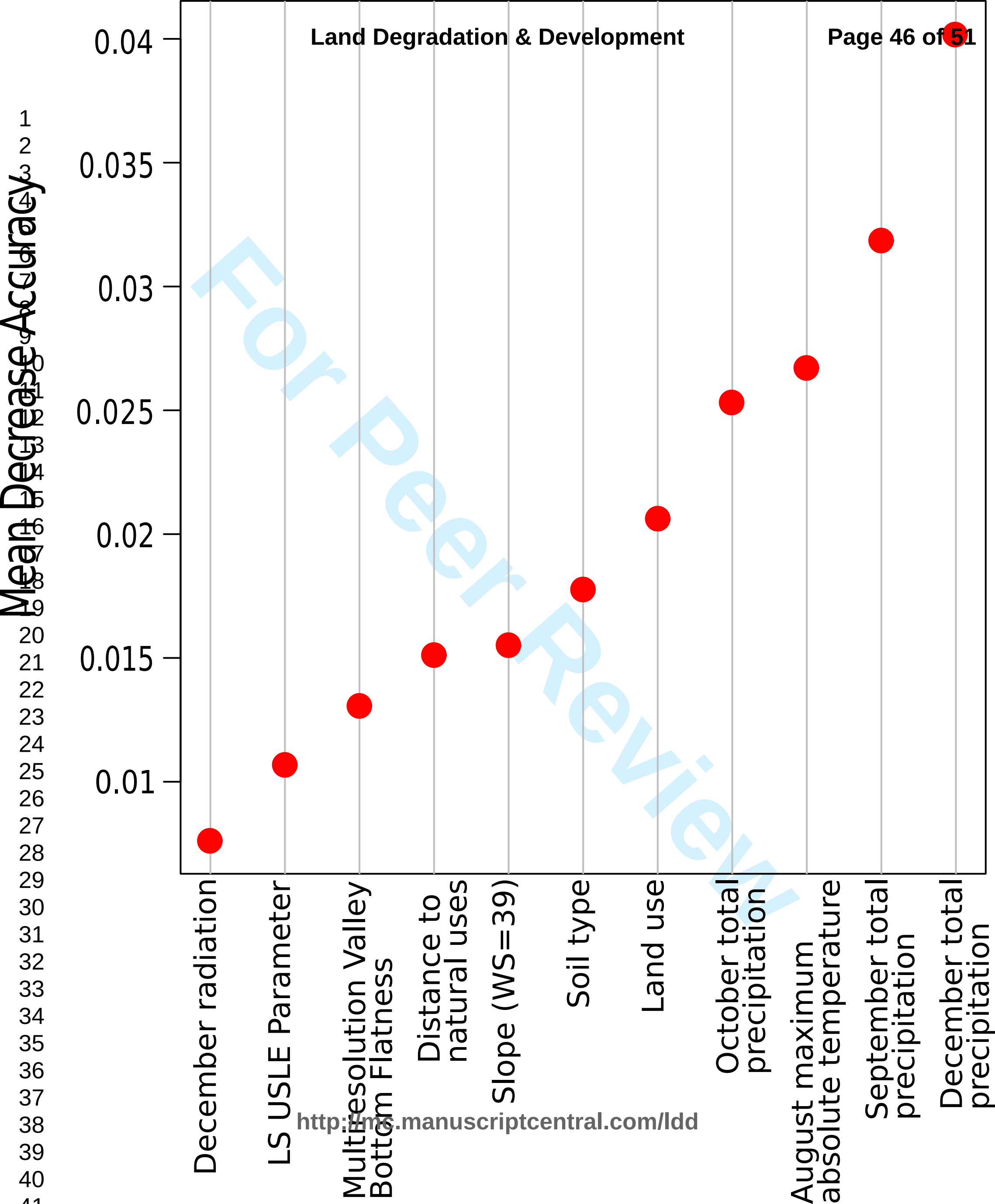
Table 4: Results for soil type predictor in the logistic model resulting from the iterative inclusion of predictors. Cambisols and gipsic Xerosols is the baseline level in the model.

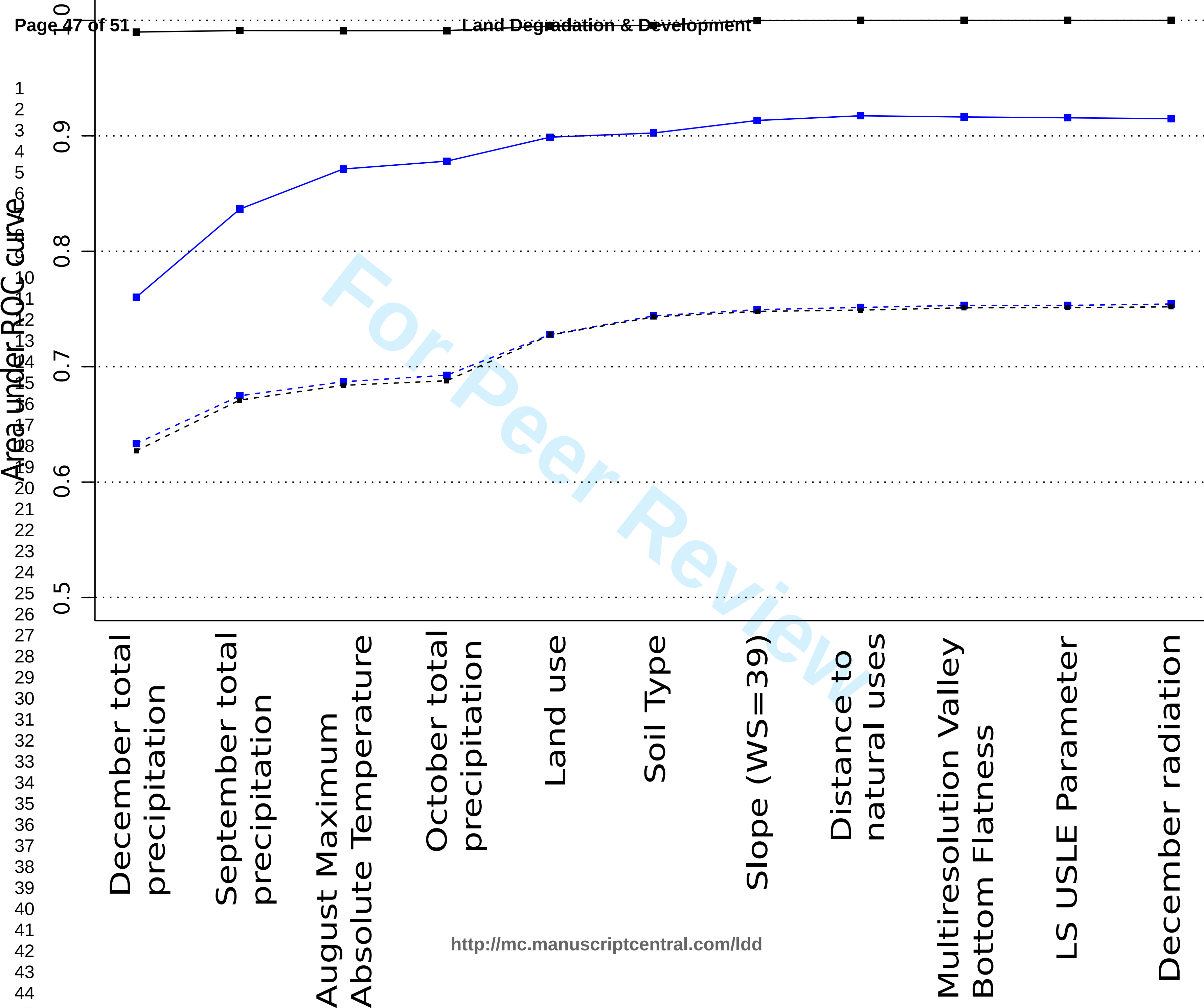

4300000
4250000
4200000
4150000
600000
650000
700000
France
SPAIN
Mediterranean Sea

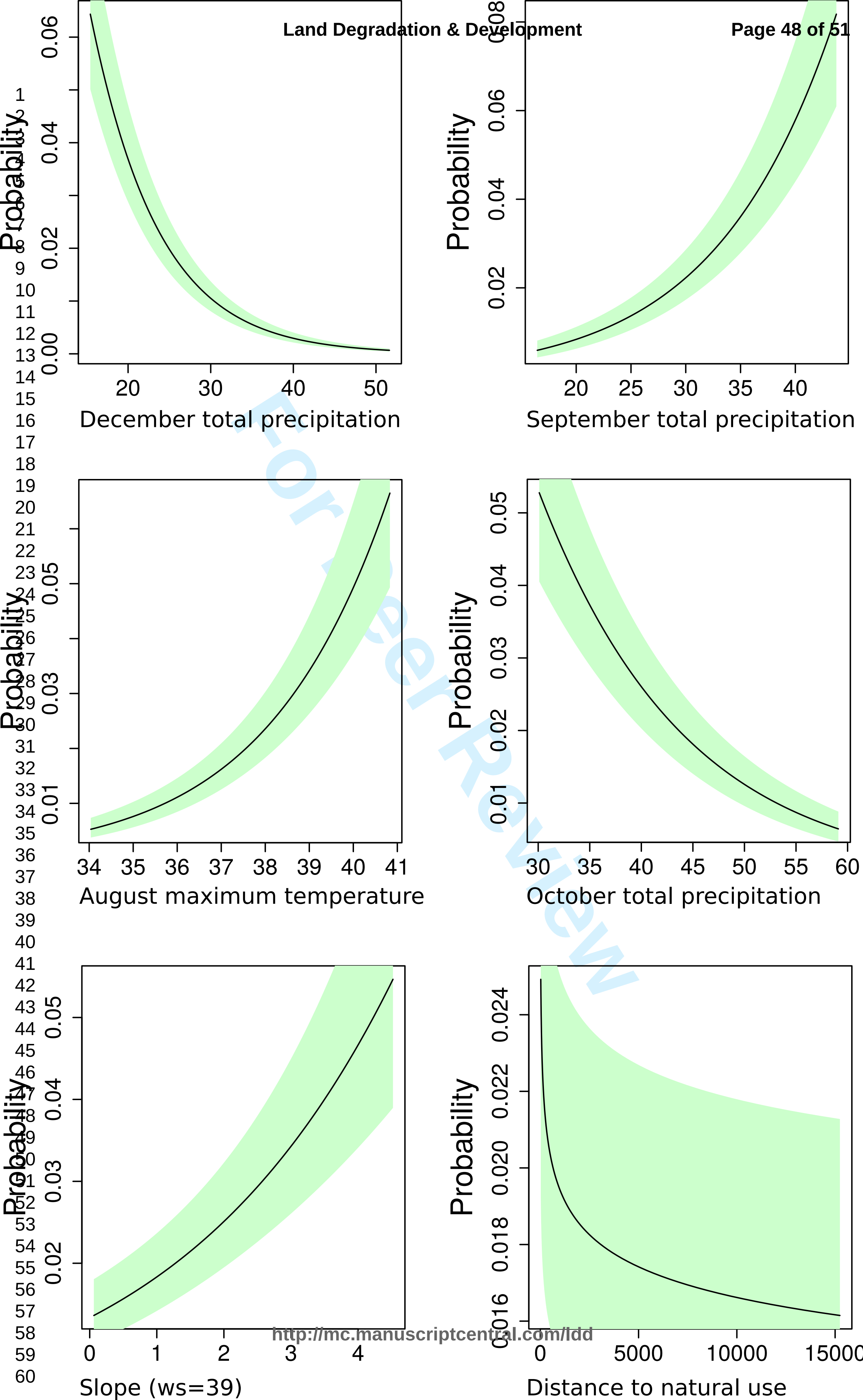

1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
Specificity
LR RF
Probability 0.032 0.057
Sensitivity 0.729 0.805
Specificity 0.673 0.837
AU Curve 0.764 0.901
Random Forest
Logistic regression

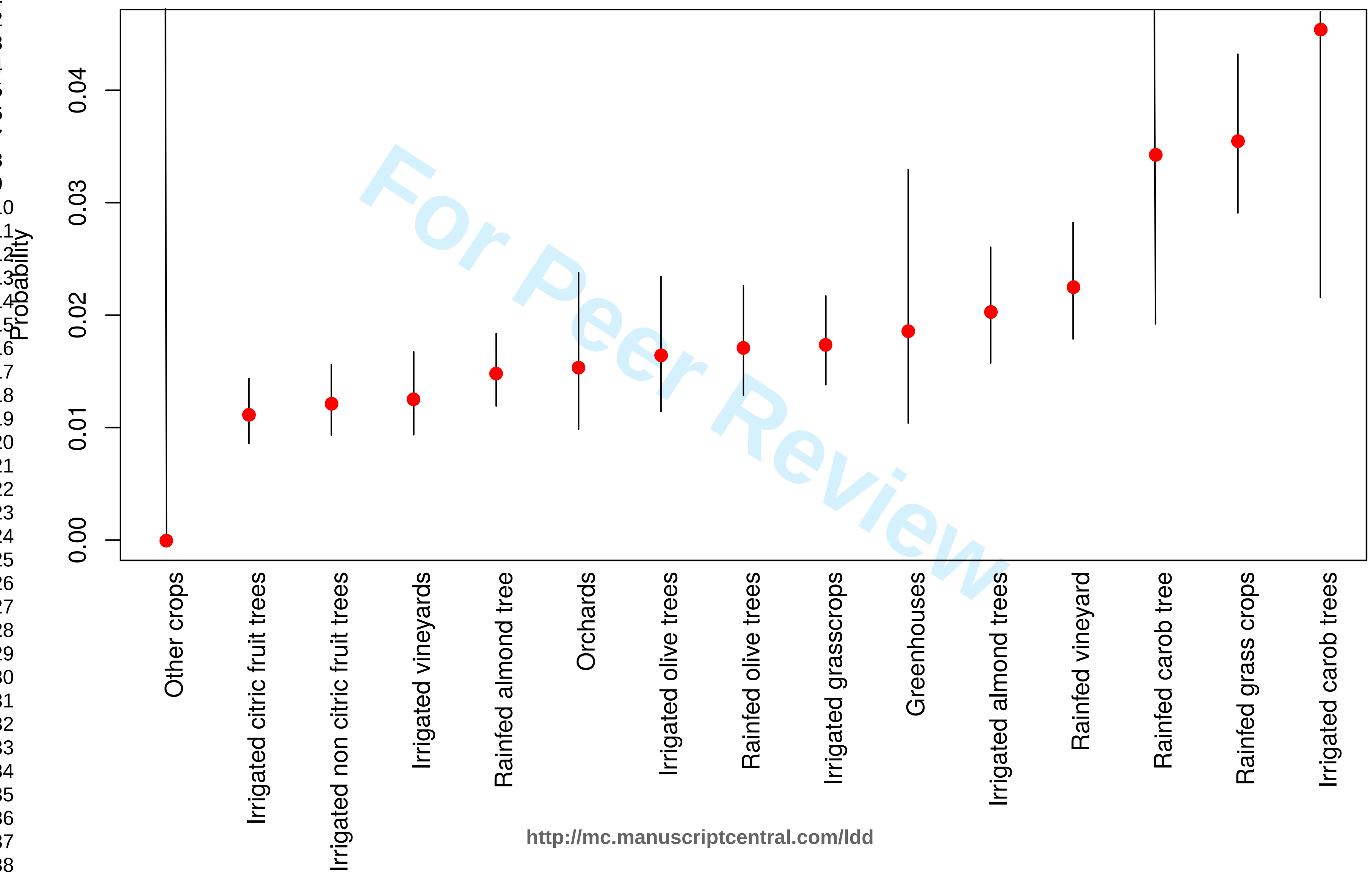

Mean Decrease Accuracy
0.02
0.015
0.01
0.005
Plan curvature (WS=23)
Orientation cosine (WS=23)
Plan curvature (WS=39)
Profile curvature (WS=39)
Plan curvature (WS=27)
Topographic Wetness Index
Plan curvature (WS=35)
Orientation sine (WS=3)
Plan curvature (WS=31)
Iwahasi Kamiya Index (WS=39)
Soil Organic Carbon
Orientation sine (WS=39)
Orientation cosine (WS=39)
December radiation
Vector Ruggedness Measure (WS=3)
LS USLE Parameter
Distance to natural uses
Vector Ruggedness Measure (WS=11)
Multiresolution Valley Bottom Flatness
Slope (WS=39)
Soil type
September Potential Evapotranspiration
April Potential Evapotranspiration
October Total Precipitation
Land Use
September Total Precipitation
August maximum absolute temperature
August Total Precipitation
December Total Precipitation
December minimum absolute temperature

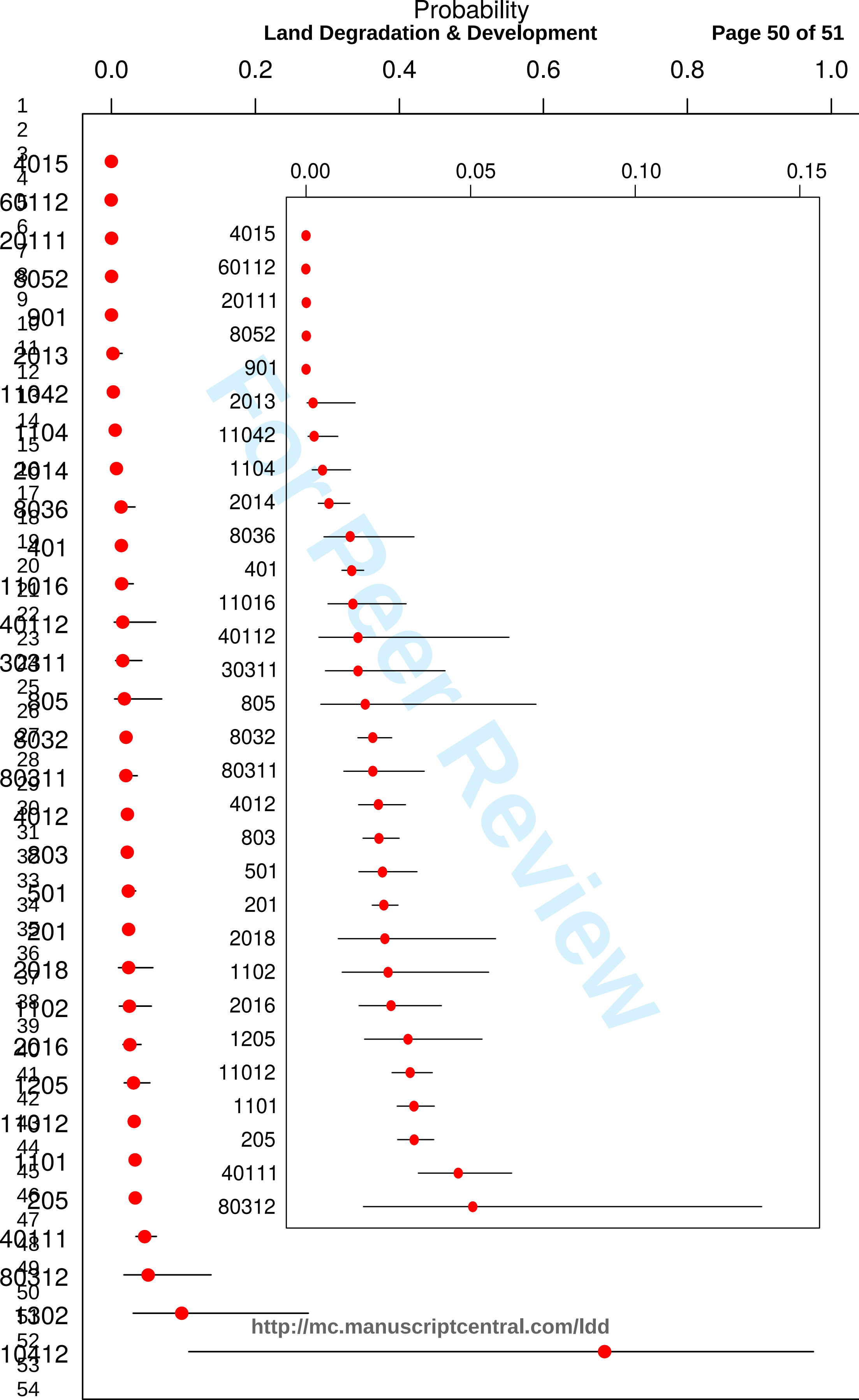

Area under ROC curve
0.75
0.80
0.85
0.90
0.95
1.00
December Minimum Absolute Temperature
December total precipitation
August total precipitation
August Maximum Absolute Temperature
September total precipitation
Land use
October total precipitation
April Potential Evapotranspiration
Setember Potential Evapotranspiration
Soil Type
Slope (WS=39)
Multiresolution Valley Bottom Flatness
Vector Ruggedness Measure (WS=11)
Distance to natural uses
LS USLE Parameter
Vector Ruggedness Measure (WS=3)
December radiation
Orientation cosine (WS=39)
Orientation sine (WS=39)
Soil Organic Carbon

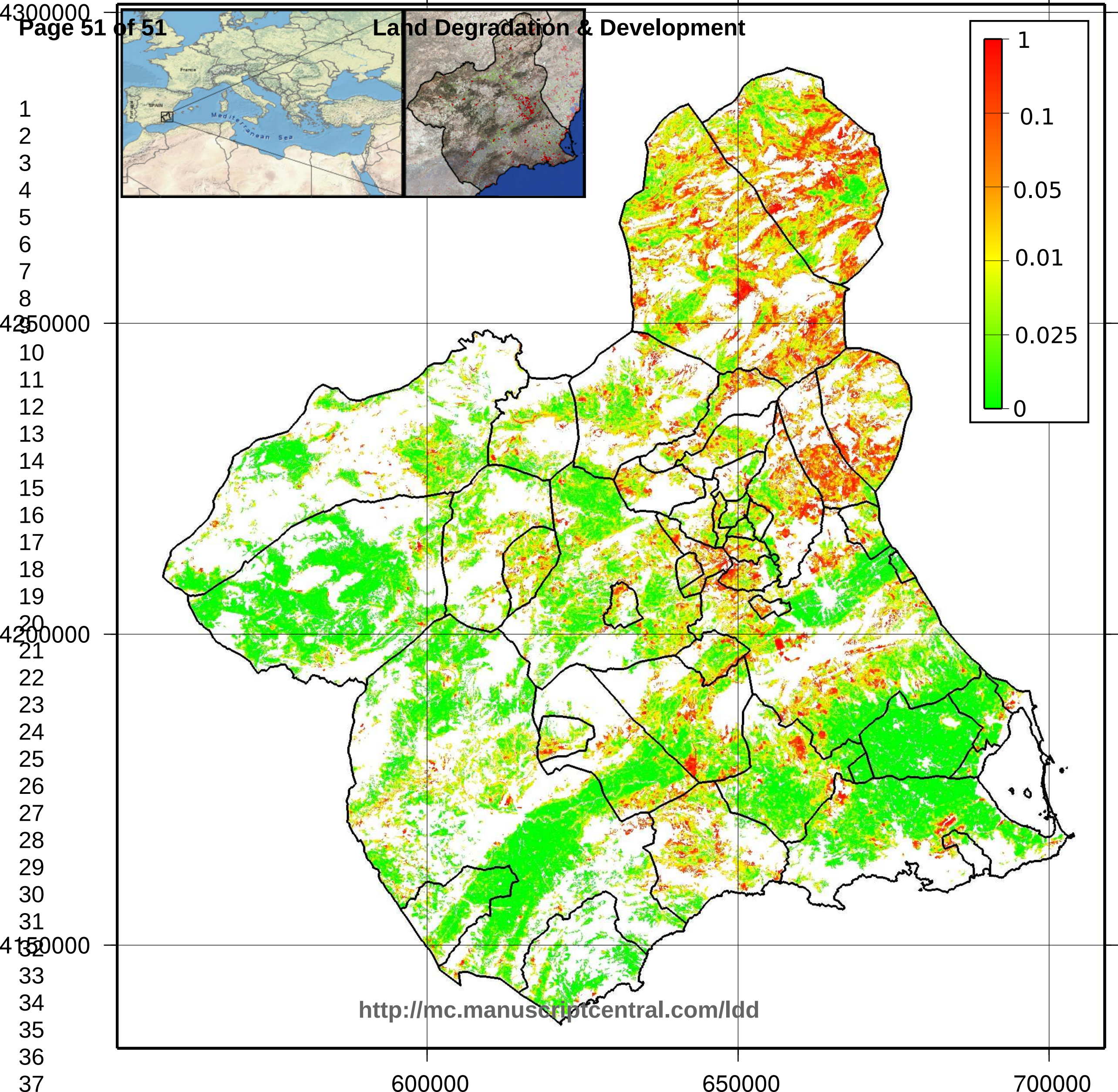

Mean Decrease Accuracy
0.04
0.035
0.03
0.025
0.02
0.015
0.01
December radiation
LS USLE Parameter
Multiresolution Valley Bottom Flatness
Distance to natural uses
Slope (WS=39)
Soil type
Land use
October total precipitation
August maximum absolute temperature
September total precipitation
December total precipitation


Area under ROC curve
0.5
0.6
0.7
0.8
0.9
1.0
December total precipitation
September total precipitation
August Maximum Absolute Temperature
October total precipitation
Land use
Soil Type
Slope (WS=39)
Distance to natural uses
Multiresolution Valley Bottom Flatness
LS USLE Parameter
December radiation


Probability
December total precipitation
Probability
September total precipitation
Probability
August maximum temperature
Probability
October total precipitation
Probability
Slope (ws=39)
Probability
Distance to natural use
For Peer Review


Probability
0.00
0.01
0.02
0.03
0.04
Other crops
Irrigated citric fruit trees
Irrigated non citric fruit trees
Irrigated vineyards
Rainfed almond tree
Orchards
Irrigated olive trees
Rainfed olive trees
Irrigated grasscrops
Greenhouses
Irrigated almond trees
Rainfed vineyard
Rainfed carob tree
Rainfed grass crops
Irrigated carob trees


Probability
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
4015
60112
20111
8052
901
2013
11042
1104
2014
8036
401
11016
40112
30311
805
8032
80311
4012
803
501
201
2018
1102
2016
1205
11012
1101
205
40111
80312
1302
10412


4300000
4250000
4200000
4150000
600000
650000
700000
1
0.1
0.05
0.01
0.025
0
France
SPAIN
Mediterranean Sea