



## RESEARCH ARTICLE

## OPEN ACCESS

# A GIS-based multivariate clustering for characterization and ecoregion mapping from a viticultural perspective

Francisco J. Moral<sup>1</sup>, Francisco J. Rebollo<sup>2</sup>, Luis L. Paniagua<sup>3</sup>, and Abelardo García-Martín<sup>3</sup>

<sup>1</sup> Universidad de Extremadura, Escuela de Ingenierías Industriales, Departamento de Expresión Gráfica. Avda. de Elvas s/n, 06006 Badajoz, Spain. <sup>2</sup> Universidad de Extremadura, Escuela de Ingenierías Agrarias, Departamento de Expresión Gráfica. Avda. Adolfo Suárez s/n, 06007 Badajoz, Spain. <sup>3</sup> Universidad de Extremadura, Escuela de Ingenierías Agrarias, Departamento de Ingeniería del Medio Agronómico y Forestal. Avda. Adolfo Suárez s/n, 06007 Badajoz, Spain

## Abstract

In wine-growing regions, zoning studies define areas according to their potential to produce specific wines and also identify the key drivers behind their variability and optimize vineyard management for sustainable viticulture. However, delineation of homogeneous zones is difficult because of the complex combination of factors which could affect zone classifications. One possibility to capture potential variability is the use of natural environmental properties as they are related to success in grape growing. With the aim of characterizing the spatial variability of the main vine-related environmental variables and determining different zones, climate and topographical data were obtained for Extremadura (southwestern Spain), an important wine region. Firstly, accurate maps of all climate indices were generated by using regression-kriging as the most suitable algorithm in which exhaustive secondary information on elevation was incorporated, and maps of topography-derived variables were obtained using GIS (Geographical Information System) tools. Secondly, principal component analysis and multivariate geographic classification were used to define homogeneous classes, resulting in three zones. Each zone was further characterized by overlaying the zonation map with a geology map and all environmental layers. It was obtained that although a wide part of the Extremaduran territory has warm climate characteristics, the zones have different viticultural potential and a high proportion of the region lays on suitable substrate. This zonation in Extremadura is the basis for further zoning studies at more detailed field scale and the modeling of vineyard response to climate change.

**Additional key words:** regression-kriging; zonation; viticulture; Extremadura.

**Abbreviations used:** AP (Annual Precipitation); CI (Cool Night Index); DEM (Digital Elevation Model); DI (Dryness Index); DO (Denomination of Origin); GIS (Geographic Information Systems); GSP (Growing Season Precipitation); HI (Huglin Index); ISR (Incoming Solar Radiation); MGC (Multivariate Geographic Clustering); PCA (Principal Component Analysis); SP (Summer Precipitation).

**Authors' contributions:** Acquisition, analysis, and interpretation of data: FJM, FJR, LLP, and AGM. Contributed analysis tools: FJM and FJR. Statistical analysis: LLP and AGM. Coordinating the research project and wrote the paper: FJM. Revision of the manuscript: FJR, LLP, and AGM.

**Citation:** Moral, F. J.; Rebollo, F. J.; Paniagua, L. L.; García-Martín, A. (2016). A GIS-based multivariate clustering for characterization and ecoregion mapping from a viticultural perspective. Spanish Journal of Agricultural Research, Volume 14, Issue 3, e0206. <http://dx.doi.org/10.5424/sjar/2016143-9323>.

**Received:** 18 Jan 2016. **Accepted:** 05 Jul 2016

**Copyright © 2016 INIA.** This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0 Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Funding:** Cofinanced by Junta de Extremadura / European FEDER funds through grants to the research groups Alcántara (TIC008) and Ingeniería Aplicada en Hortofruticultura y Jardinería (TPR009).

**Competing interests:** The authors have declared that no competing interests exist.

**Correspondence** should be addressed to Francisco J. Moral: [fjmoral@unex.es](mailto:fjmoral@unex.es)

## Introduction

Identification of zones with similar characteristics within a region or any other area of interest is necessary to describe land resources limitations and potentials. The basis at zonation is to delimit geographic units relevant to crop present and future uses, and to provide a basis for strategic planning production (e.g., Liu & Samal, 2002; Munier *et al.*, 2004).

In viticulture, the unit of land in which topography, climate, soil properties, and other characteristics (e.g., variety, rootstock, historical experience, and winemaking) are homogeneous is denominated terroir (OIV, 2010). Delineation of terroir unit, which are likely not only to produce grapes and wines with similar characteristics but also to enable crop operational decisions, is the objective of viticultural zoning. However it is not easy to accurately define homogeneous manage-

ment areas due to the complex interactions of all factors. Initial zoning approaches were performed considering only a single terroir factor (*e.g.*, Huglin, 1978; Mackenzie & Christy, 2005). However, nowadays it is clear that an improved zoning process has to be developed from a multivariate perspective.

The development of geomatics, combining spatial information and geodesy, and the growth in computing power have opened up interesting prospects for using multivariate models and geographic information systems (GISs) as essential tools to be applied in zoning approaches (*e.g.*, Zhou *et al.*, 2003; Herrera-Nuñez *et al.*, 2011). Georeferenced information can be managed in a GIS, also enabling different databases to be associated with maps and different layers can be cross referenced in order to assess the agricultural potential of a geographical area. The main limitation of using a GIS is in the reliability and quality of the information contained in the layers.

Use of multivariate models in combination with GIS can potentially constitute an effective way to characterize the environmental variability. According to the information generated, this variability can be further processed and managed with the aim of optimizing winegrape production and establishing the most appropriate crop methods in any particular location (*e.g.*, Williams *et al.*, 2008), which is important for a sustainable viticulture.

The objective of this study was to present an approach for the delineation of relatively homogeneous zones using a multivariate clustering approach in a GIS. These zones represent areas of similarity of combined climatic, topographic, and edaphic conditions. Results are used to analyze a traditional winemaking area.

## Material and methods

### The study area

This work is centred in the Autonomous Community of Extremadura (latitude between 37° 57' and 40° 29' N, longitude between 4° 39' and 7° 33' W). This region is located in the southwest of Spain on the Portuguese border, and it is one of the largest regions in Europe (approximately 41,600 km<sup>2</sup>, similar to Belgium). Elevation ranges from 2091 m a.s.l., at Gredos mountains, to 116 m a.s.l., in the Guadiana valley, near the border between Spain and Portugal, being the mean elevation about 425 m a.s.l. Figure 1 shows the digital elevation model (DEM), with a spatial resolution of 1 km, used in this research.

Extremadura is a region with a great contrast, with wide agricultural and forest areas; it is considered to be one of the most important ecological enclave in

Europe. Extremadura's climate is Mediterranean, modified by the interior location of the region and by oceanic influences that penetrate the peninsula due to its proximity to the Atlantic. It is a semiarid region, where the water balance is negative. Mean annual precipitation is less than 600 mm in the majority of the areas of the region, even less than 400 mm in the center of the Guadiana valley, but it may exceed 1000 mm in the northern (Gredos) and eastern (Guadalupe) mountainous areas. One of the most important characteristics of the precipitation is its intraannual variability. There is a dry season, from June to September, and a wet season, from October to May, in which 80% of the precipitation occurs. Periodic droughts with duration of 2 or more years, with an occurrence of once in the 8-9 years, are frequent (Almarza, 1984).

The Extremaduran wines gained their identity and quality recognition in 1999, when on April, 16, the Order of the Ministry of Agriculture, Fisheries and Food ratified the regulation of the regulatory board (BOE, 1999). Thus, the Ribera del Guadiana Denomination of Origin (DO) was established with six Extremaduran wine-growing subzones (Fig. 1). This recognition was not based on potential productivity but rather on adequate product quality. Traditionally, a great amount of the wine produced in Extremadura was exported without bottling, as a bulk-wine. Since the DO recognition, this is less common and different white, red and rose bottled wines are produced and their quality is increasingly recognized.

Vineyard surface area currently covers ~ 81,963 ha in Extremadura, ~8.5 % of vineyards in Spain, ranking second in terms of vineyard surface; regarding the wine production, Extremadura ranks also second in Spain with 2,920,965 hL (MAGRAMA, 2015). Wine exports from Extremadura, about 2,600,000 hL in 2011, increased more than 49 % respect to the previous year (OEMV, 2012).

### Environmental data set

The first step of this study was to generate an environmental data set for the region. Although information about the main environmental variables of interest exists in Extremadura, usually its density is too low to represent the regional variability, so it is necessary to contact different national and regional agencies to obtain the basic data.

Climate, topographical, and geological data constitute the environmental information that was obtained digitally and subsequently incorporated, managed, and displayed into a GIS architecture, in ArcGIS v.10 (Environ. Syst. Res. Inst., Redlands, CA, USA).

### Climate data

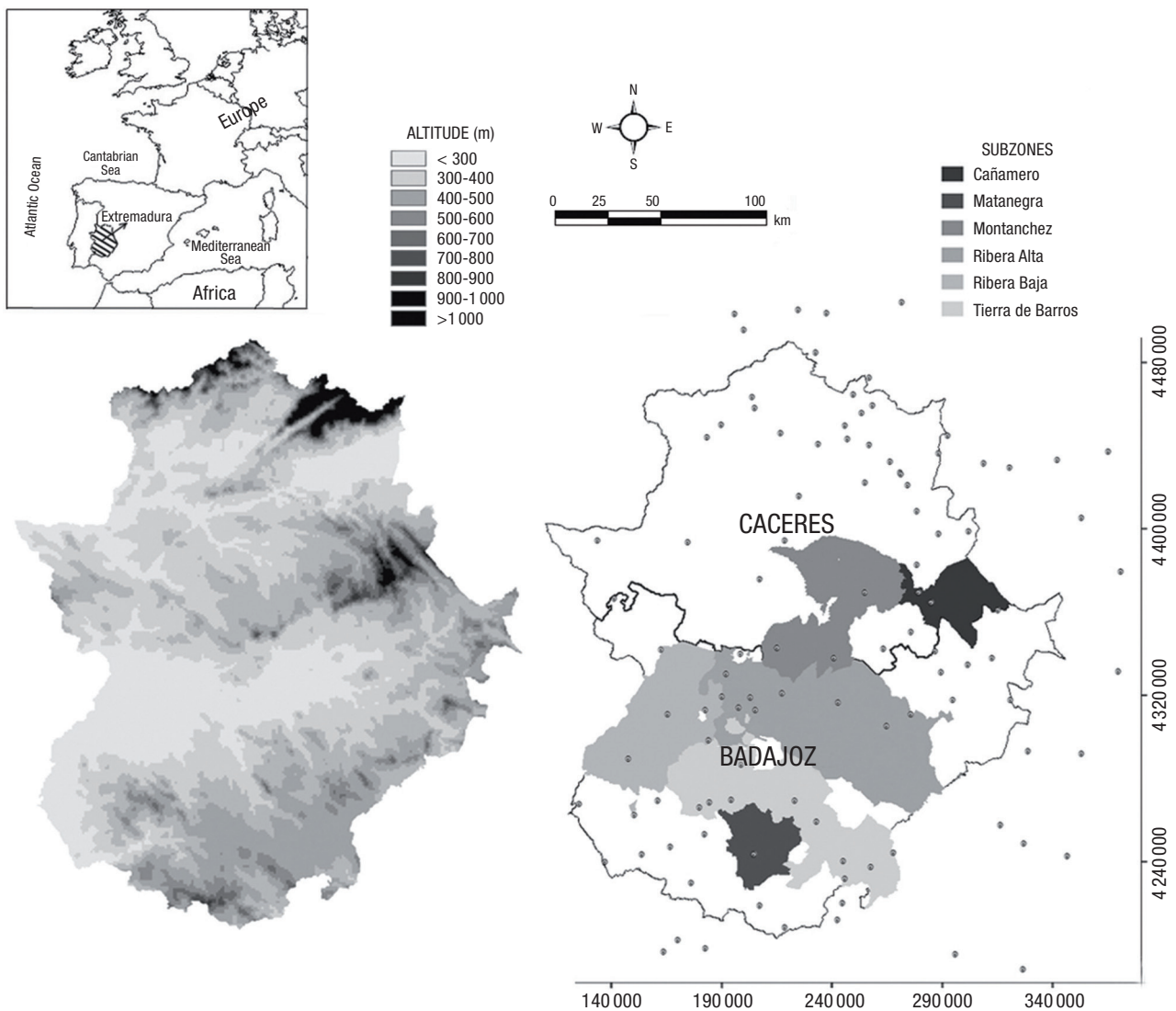
Daily weather observations were obtained for the period 1980-2011, from 117 georeferenced meteorological stations belonging to the Meteorology Agency of the Spanish Government, 90 located within Extremadura and 27 along boundaries (Fig. 1). Six climate indices were computed for each meteorological station. Two of them are temperature-derived indices, the Huglin index (HI) and the cool night index (CI), and three are precipitation-based indices, the annual (AP) and summer (SP, considering June, July, and August) precipitation, and the growing season precipitation (GSP, considering all days between 30 April and 1 October). In addition, the dryness index (DI), which takes into account the potential water balance of the soil over the growing cycle, was also computed. All these indices

are simple indicators of the long-term suitability of climate for the successful production of winegrapes.

The HI (Huglin, 1978) is defined as:

$$HI = \sum_{1\text{April}}^{30\text{Sept}} \left[ \frac{(T-10) + (T_{\max}-10)}{2} \right] d \quad [1]$$

where T is the mean air temperature (°C), T<sub>max</sub> is the maximum air temperature (°C), and d is the length of day coefficient. This index provides information regarding the level of heliothermal potential, being calculated on a period biologically acceptable, and also provides a good idea of the sugar potential according to varieties (Tonietto & Carbonneau, 2004). According to Hall & Jones (2010) the HI is the most suitable index and most representative of a region's viticultural suitability.



**Figure 1.** Location map of the autonomous community of Extremadura. Digital elevation model (left) and Ribera del Guadiana Denomination of Origin subzones (right) in both provinces, Badajoz and Cáceres, of Extremadura. Meteorological stations used in this study are also indicated as points in the right map. Coordinate system is ETRS89/UTM zone 30N.

The CI represents the average of the daily minimum air temperature (°C) in the month of September. CI is a night coolness variable that takes into account the mean minimum night temperatures during the month when ripening usually occurs. Its purpose is to improve the assessment of the qualitative potentials of wine-growing regions in relation to the production of secondary metabolites, such as polyphenols and aromas, that are important in regards to grape wine colour and aromas (Tonietto & Carbonneau, 2004).

Both HI and CI permit a good discrimination of the regional climate, as regards global heliothermal conditions during vegetative cycle of the grape and cool night conditions during ripening period. Tonietto & Carbonneau (2004) added the DI to describe the climate of viticultural regions worldwide. The DI can be computed using:

$$DI = Wo + P - Tv - Es \quad [2]$$

where  $Wo$  is the initial useful soil water reserve (mm), which can be accessed by the roots, usually adopting the value of 200 mm,  $P$  is the precipitation (mm),  $Tv$  is the potential transpiration of the vineyard (mm), and  $Es$  is the direct evaporation from the soil (mm). In the northern hemisphere, DI is calculated from 1 April to 30 September. DI can be negative, when there is a potential water deficit, and should not be greater than  $Wo$ .

After computing the indices for the locations of all meteorological stations, it was necessary to estimate their values across the region. A multivariate geostatistical algorithm, concretely regression-kriging (Odeh *et al.*, 1994, 1995; Goovaerts, 1999), was used to estimate values of all indices, using elevation as covariable (secondary variable). This interpolation algorithm was chosen because it generates lower errors when applied to estimate climate variables in Extremadura, where data are very sparse (Moral, 2010). Thus, from sampled point data (meteorological stations) estimates can be generated for any other unsampled location, which is then used to produce a continuous surface for the region. Values for each index were then available for each smallest resolution unit (pixel) within the region. Six digital models, one for each climate index, in raster format at  $1000 \times 1000$  m resolution were finally generated. Accuracy of all surfaces were determined with the cross-validation process. Some statistics showed that the models are reasonable because low values were obtained in all cases (Table 1). The root mean square errors computed for all indices are lower than 6.92% of their mean values, being even lower than 4% for AP and GSP. In consequence, all maps represent very closely the real spatial distribution of the considered climatic indices, in accordance with Bishop & McBrat-

**Table 1.** Cross-validation statistics (RMSE, root mean square error; PERC, ratio (%) between RMSE and the average value of each index).

|      | HI     | CI   | DI    | AP    | SP   | GSP   |
|------|--------|------|-------|-------|------|-------|
| RMSE | 145.02 | 0.59 | 11.16 | 20.87 | 1.55 | 11.55 |
| PERC | 5.11   | 4.01 | 6.92  | 3.11  | 4.72 | 3.91  |

HI, Huglin Index; CI, Cool Night Index; DI, Dryness Index; AP, Annual Precipitation; SP, Summer Precipitation; GSP, Growing Season Precipitation.

ney (2001) who established that the accuracy of a model is higher when kriging of the residuals is incorporated in the interpolation algorithm. All operations were performed in ArcGIS, using the Geostatistical Analyst and Spatial Analyst extensions.

#### Topographical data

Three indices included in this category were considered: elevation, slope, and incoming solar radiation (ISR). A DEM for Extremadura, in raster format at  $1000 \times 1000$  m resolution was acquired. From this raster map, *i.e.* the DEM, by using the tools included in ArcGIS, and specifically the Surface and Solar Radiation in Spatial Analyst Tools, a raster map of slopes and a raster map of incoming solar radiation from April to September were generated at the same resolution of the DEM.

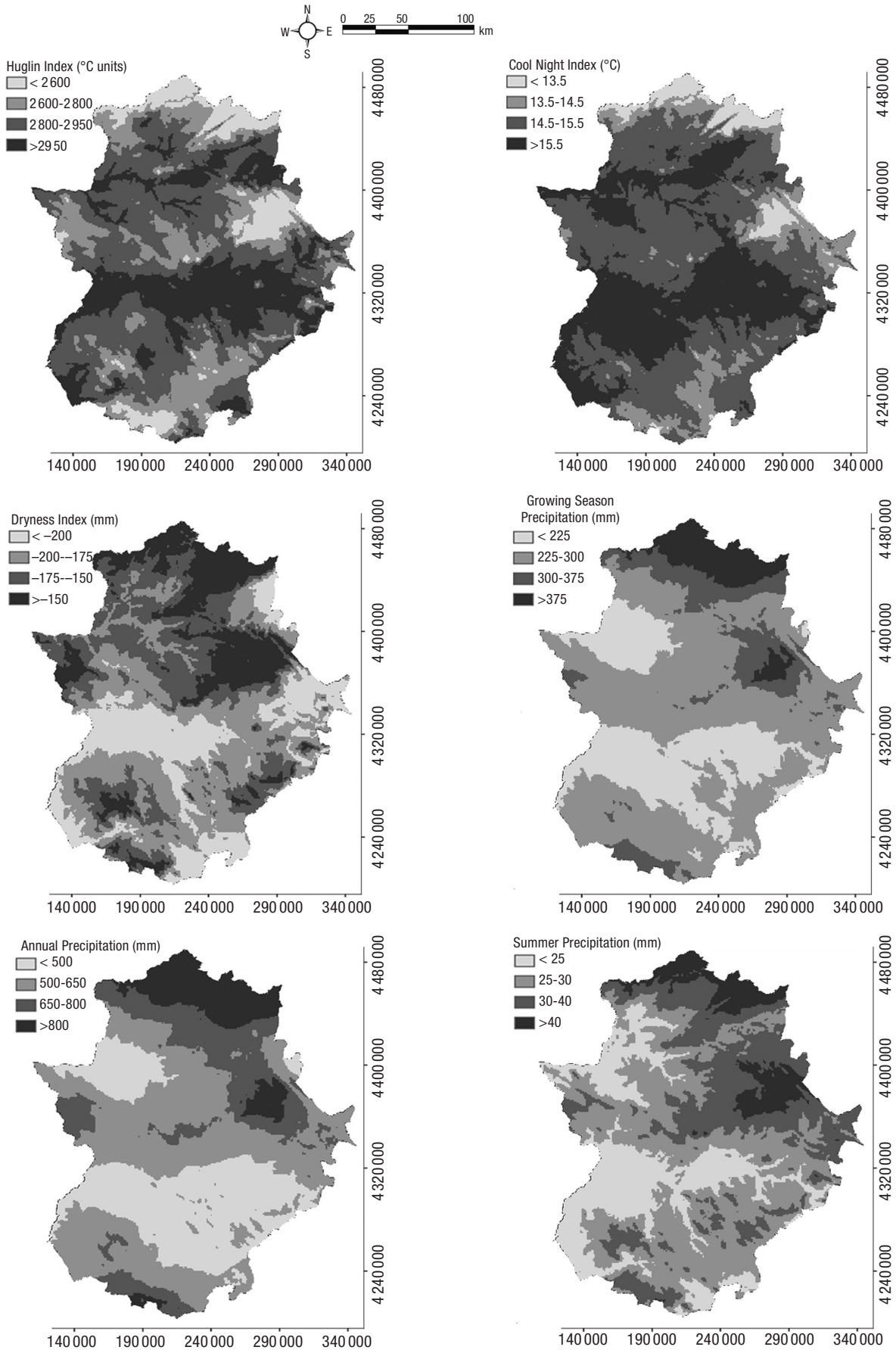
Nine raster layers (HI, CI, DI, GSP, AP, SP, elevation, slope, and ISR), at 1 km resolution, were finally generated (Fig. 2).

#### Geological data

Although information about the main soil properties is important, these data are unavailable for the Extremadura region. However, as parent rock is related to soil distribution, a 1:250 000 scale geology map of Extremadura was used. It can be freely downloaded through the web of the Geological and Mining Information System of Extremadura (<http://sigeo.gobex.es/portalsigeo/web/guest>).

### Principal component analysis and geographic clustering

Delineation and characterization of homogeneous zones in Extremadura was carried out using a principal component analysis (PCA) followed by a multivariate geographic clustering (MGC) algorithm, considering



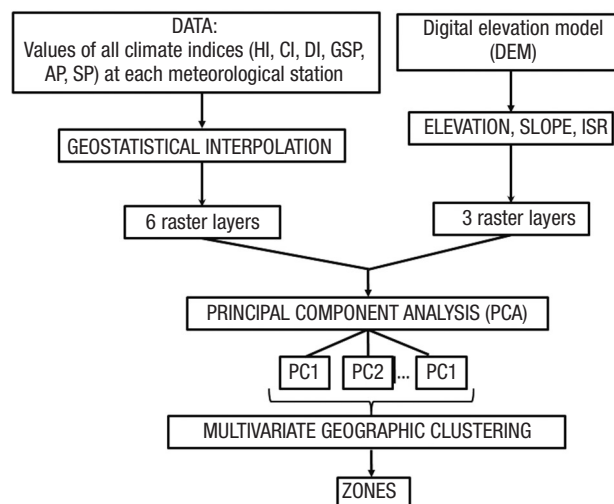
**Figure 2.** Kriged maps of all climate indices used in the study. Coordinate system is ETRS89/UTM zone 30N.

the nine aforementioned layers. Previously, variables values of each raster cell were normalized to a standard normal distribution. The PCA and MGC were conducted using Multivariate in Spatial Analyst Tools, included in ArcGIS.

The PCA is used to transform the data in the input layers from the input multivariate attribute space to a new multivariate attribute space whose axes are rotated with respect to the original space, obtaining uncorrelated axes in the new space. The main reason to transform the data in a PCA is to compress them by eliminating redundancy. The result of the PCA is a multiband raster with the same number of bands as the specified number of components (one band per axis or component in the new multivariate space). In this case study, the nine components (layers) formed 9 orthogonal axes in the data space into which each cell was plotted. Similarity of cells within the 9-dimensional data space was then coded as Euclidean separation distance. The PCA is a linear transformation that reorganize the variance of a multiband image into a new set of image bands. Each band in the output receives some contribution from all of the inputs bands. The first principal component has the greatest percent of variation explained, the second shows the second most important percent of variation not described by the first, and so forth. Usually, the first three or four components of the resulting multiband raster describes more than 90% of the variance, thereby the remaining individual raster bands can be dropped.

After performing the PCA, a MGC was conducted with the resultant principal components explaining a cumulative variance higher than 90%. An isodata clustering unsupervised classification algorithm was used. It constitutes an iterative self-organizing way of performing clustering, which, during each iteration, all samples are assigned to existing cluster centers and new means are recalculated for every class. Firstly, this algorithm assigns an arbitrary mean, one for each initial cluster. The second step classifies each pixel to the closest cluster. Later, the new cluster means are calculated based on all the pixels in one cluster, considering a multidimensional attribute space. These second and third stages are repeated until the change between the iteration is small. The number of iterations should be large enough to ensure that, after running the specified number of iterations, the migration of cells from one cluster to another is minimal; therefore, all the clusters become stable. As Williams *et al.* (2008) indicate, the advantages of using MGC include production of self-seeded, self-describing zones, and similarity of environmental heterogeneity among zones.

The contribution of the variables considered in this study to obtain homogeneous zones was performed through the stages displayed in Fig. 3.



**Figure 3.** Schematic diagram of the stages involved in the delineation of the zones (HI, Huglin Index; CI, Cool Night Index; DI, Dryness Index; GSP, Growing Season Precipitation; AP, Annual Precipitation; SP, Summer Precipitation; ISR, Incoming Solar Radiation).

## Results and discussion

After processing all climatic and topographic data, a preliminary characterization of the Extremadura region was achieved. Different homogeneous zones (which can be used later to delimit smaller subzones or terroir units) were established.

From the PCA, all nine components were retained and they indicated the axes of the data space. Table 2 lists the percentage of variation explained by each component. The first principal component explained ~80% of the total variance, the second principal component ~8%, and the third principal component ~5%, that is, the first three principal components explained ~93% of the total variance. The eigenvectors from the PCA indicate the amount that each initial input band contributed to each PCA band (Table 3). A striking result was that contributions of all input layers to the

**Table 2.** Results of the principal component analysis for the nine layers.

| Principal component | Eigenvalue | Component loading | Cumulative loading |
|---------------------|------------|-------------------|--------------------|
| PC1                 | 4.258      | 79.71             | 79.71              |
| PC2                 | 0.448      | 8.39              | 88.10              |
| PC3                 | 0.262      | 4.90              | 93.00              |
| PC4                 | 0.185      | 3.45              | 96.45              |
| PC5                 | 0.099      | 1.86              | 98.31              |
| PC6                 | 0.048      | 0.87              | 99.18              |
| PC7                 | 0.026      | 0.48              | 99.66              |
| PC8                 | 0.016      | 0.29              | 99.95              |
| PC9                 | 0.003      | 0.05              | 100.00             |

**Table 3.** Eigenvectors. PC loadings for each variable.

|     | Elevation | Slope  | ISR   | HI     | CI     | DI     | GSP    | AP     | SP     |
|-----|-----------|--------|-------|--------|--------|--------|--------|--------|--------|
| PC1 | 0.343     | 0.286  | 0.317 | -0.352 | -0.347 | 0.326  | 0.339  | 0.321  | 0.362  |
| PC2 | 0.386     | -0.203 | 0.523 | -0.180 | -0.189 | -0.108 | -0.427 | -0.519 | -0.065 |
| PC3 | -0.037    | -0.925 | 0.087 | -0.057 | -0.066 | 0.045  | 0.224  | 0.261  | 0.089  |

ISR, Incoming Solar Radiation; HI, Huglin Index; CI, Cool Night Index; DI, Dryness Index; GSP, Growing Season Precipitation; AP, Annual Precipitation; SP, Summer Precipitation.

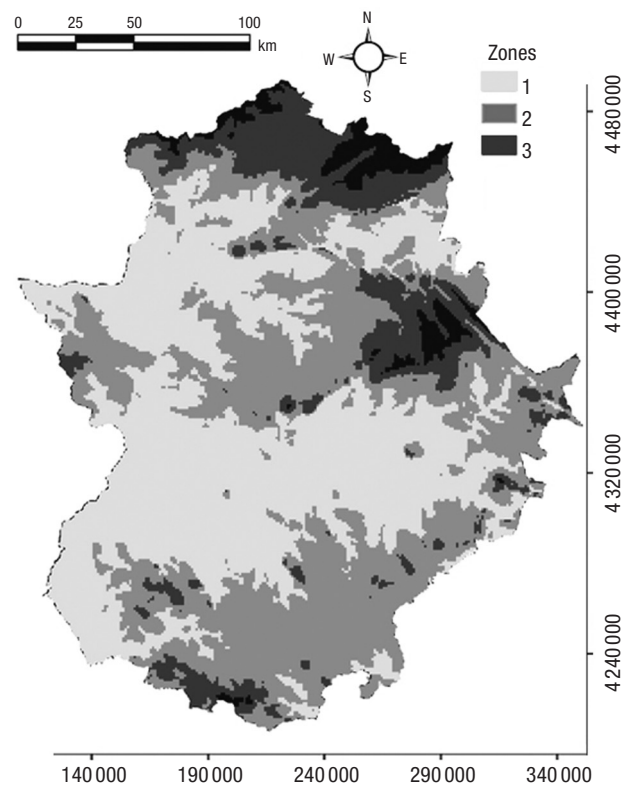
first principal component were similar. Thus, in this case the use of PCA is not for reduction of dimensionality; its use is for considering all the variability of the data to discriminate homogeneous zones, and for transforming all raster cells back into geographic space with environmentally relevant units. The second principal component has ISR, AP, GSP, and elevation as the highest loading parameters, and the third principal component has slope as the main loading parameter. But, as previously indicated, the greatest amount of variance was explained by the first principal component, so all input layers were important in this region for zonation. The large area of the region and its differences in climate and topography prevent removal any of initial layer to accurately take into account the spatial variability, in contrast with other zonation studies carried out in smaller areas (*e.g.*, Herrera-Nuñez *et al.*, 2011).

The first three principal components were introduced in the MGC analysis. The final result was a classification map of Extremadura, with three differentiated zones (Fig. 4). An initial analysis after considering four zones showed how two of them had the same characteristics, that is, their means for all variables did not differ. Visually, an apparent relationship between zones and topography can be appreciated. Zone 1 is found along the valleys, mainly along the valleys of the Tagus and Guadiana rivers and covers ~45% of the region. The areas with higher elevations, and hilly areas near the rivers, are mainly in zone 2, and covers ~39% of the region. The remaining area, zone 3, covers ~16% of Extremadura, and represents the highest elevations and slopes (Table 4).

Summary statistics for all climate indices and topographical data within each zone are shown in Table 4. Taking into account the multicriteria climatic classification system proposed by Tonietto & Carbonneau (2004), and according to the mean values of HI, CI, and DI, zones 1 and 2 are classified as HI+2, CI-1, and DI+2, and zone 3 is classified as HI+2, CI+1, and DI+2. Thus, zones 1 and 2 correspond to the warm class of viticultural climate (HI+2), which is characterized by a potential exceeding the heliothermal needs to ripen the varieties and some risk of stress, temperate nights (CI-1), so later varieties ripen under lower night tem-

peratures than the early ones, and very dry areas (DI+2), in which potential dryness is very high, generating frequent stress effects and needing irrigation. Zone 3 is similar but differs in terms of cool nights class (CI+1), with cooler conditions than in zones 1 and 2.

Mean HI, CI, and ISR values are similar between zones 1 and 2, but in zone 3 HI and CI are lower while ISR is slightly higher. Potential energy is highest in zones 1 and 2, which are flatter areas with less elevation. DI values are predominantly negative in all zones, due to dry to very dry climatic conditions and leading to severe plant stress in those areas where potential dryness is very pronounced, *i.e.*,  $DI < -200$  mm; an adequate irrigation is mandatory in



**Figure 4.** Zonation map resulting from principal component analysis and geographic clustering based on the kriged maps of climate indices and topographical properties. Coordinate system is ETRS89/UTM zone 30N. Black areas are zones in which elevation is higher than 800 m.

**Table 4.** Summary statistics of all variables used in the study.

| Variables <sup>[1]</sup>  | Zone 1  |         |           |                   | Zone 2  |         |           |       | Zone 3  |         |           |        |
|---------------------------|---------|---------|-----------|-------------------|---------|---------|-----------|-------|---------|---------|-----------|--------|
|                           | Min     | Max     | Mean      | SD <sup>[2]</sup> | Min     | Max     | Mean      | SD    | Min     | Max     | Mean      | SD     |
| Elevation (m)             | 136     | 504     | 303.64 c  | 63.71             | 223     | 749     | 464.58 b  | 82.37 | 212     | 2091    | 673.53 a  | 269.84 |
| Slope (%)                 | 0.01    | 4.54    | 0.83 c    | 0.54              | 0.01    | 6.27    | 1.25 b    | 0.87  | 0.03    | 16.60   | 3.78 a    | 30.20  |
| ISR (kWh/m <sup>2</sup> ) | 931.95  | 982.74  | 956.09 a  | 7.01              | 936.81  | 1008.88 | 972.85 a  | 10.24 | 906.63  | 1184.95 | 991.80 a  | 33.49  |
| HI (°C units)             | 2728.79 | 3188.43 | 2948.73 a | 74.21             | 2471.63 | 3081.09 | 2812.21 b | 85.37 | 798.09  | 3019.91 | 2486.22 c | 322.18 |
| CI (°C)                   | 13.97   | 16.59   | 15.61 a   | 0.33              | 13.34   | 16.30   | 14.92 a   | 0.42  | 5.81    | 16.15   | 13.55b    | 1.47   |
| DI (mm)                   | -294.51 | -139.86 | -192.66 b | 19.65             | -281.77 | -110.27 | -175.63 b | 24.33 | -234.53 | 179.43  | -115.52 a | 58.46  |
| AP (mm)                   | 389.18  | 832.34  | 525.66 b  | 74.73             | 422.92  | 883.75  | 594.13 b  | 92.58 | 472.61  | 1687.03 | 893.75 a  | 221.51 |
| SP (mm)                   | 16.53   | 35.35   | 25.14 b   | 2.54              | 18.88   | 41.36   | 29.98 b   | 3.09  | 25.24   | 104.43  | 43.38 a   | 11.95  |
| GSP (mm)                  | 169.09  | 326.81  | 231.64 b  | 28.20             | 192.51  | 362.32  | 264.66 b  | 31.19 | 225.64  | 773.27  | 390.18 a  | 96.36  |

<sup>[1]</sup>ISR, incoming solar radiation; HI, Huglin index; CI, cool night index; DI, dryness index; AP, annual precipitation; SP, summer precipitation; GSP, growing season precipitation <sup>[2]</sup>SD, standard deviation. Significant differences between means are indicated by different letters ( $p < 0.05$ ).

this case to avoid stress effects. With respect to the precipitation indices (AP, SP, and GSP), due to the direct relationship between precipitation and elevation (Moral, 2010), and the fact that each zone has different topographic characteristics, mean and range of precipitation values among zones are also different, being lowest in zone 1, intermediate in zone 2, and the highest in the more mountainous zone 3 (Table 4).

A first study to define homogeneous zones in Extremadura was conducted in Moral *et al.* (2016), only considering some bioclimatic indices to delimit viticultural climatic zones. In the present study, topographical information has also been incorporated and, consequently, its effects on the delineation of zones are apparent. However, some similarities exist as the fact that three zones were delimited in both cases and zone 3 in this study, which comprises the higher areas, has a similar pattern to a zone of the previous study, with similar climatic characteristics. The patterns of the other zones are very different, showing the influence of the topographical properties on the definition on the homogeneous zones.

As it was previously indicated, data for soil physico-chemical properties in Extremadura is lacking; thus, this important information cannot be used to generate a GIS, and later, to be incorporated in the zonation algorithm. However, we utilized a geology map due to the relationship between soil distribution

and rock weathering. The proportion of different subsoil classes of viticultural importance within each zone, obtained overlaying the geology map with the zonation layer, is shown in Table 5. Thus, sandstone or arenaceous rocks predominate within zone 3, and clay deposits, fluvial gravel and sand deposits in zone 1, encompassing ~37% of this zone. In general, a higher percentage of zone 1 has lithologic classes of interest for viticulture (Herrera-Núñez *et al.*, 2011), ~72% of its surface expanse, being lower in zone 3, accounting for only 60%.

Further analysis of the Ribera del Guadiana DO wine-growing subzones in relation to the zonation layer showed that zone 1 dominates in two subzones, zone 2 dominates in three, and zone 3 dominates in one subzone (Table 6). Taking into account the characteristics of both zones 1 and 2, it is important to denote the suitability and potential for viticulture of almost the whole territory within the DO. Only in the Montánchez and Cañamero subzones, where ap-

**Table 5.** Percentage (%) of subsoil classes within each zone.

|        | Sandstone or arenaceous rock | Clay deposits | Fluvial gravel and sand deposits | Others |
|--------|------------------------------|---------------|----------------------------------|--------|
| Zone 1 | 34.92                        | 23.32         | 13.72                            | 28.04  |
| Zone 2 | 48.34                        | 15.21         | 1.52                             | 34.93  |
| Zone 3 | 54.09                        | 3.15          | 2.95                             | 39.81  |

**Table 6.** Percentage (%) of Ribera del Guadiana DO subzones within each zonation class.

|        | Cañamero | Matanegra | Montánchez | Ribera Alta | Ribera Baja | Tierra de Barros |
|--------|----------|-----------|------------|-------------|-------------|------------------|
| Zone 1 | 8.69     | 4.64      | 28.32      | 74.97       | 94.62       | 44.53            |
| Zone 2 | 36.82    | 95.36     | 52.47      | 23.82       | 5.38        | 52.70            |
| Zone 3 | 54.48    | 0         | 19.21      | 1.22        | 0           | 2.76             |



proximately a fifth and more than half of their areas respectively belong to zone 3, coinciding with higher elevations and more abrupt orography, a considerable part of their territories has less suitable characteristics for viticulture.

As conclusions, three homogeneous zones of viticultural interest have been delineated in Extremadura by integrating different environmental (climatic and topographical variables) information important for grapevine and wine production. PCA and MGC techniques were applied to spatial data layers in which the patterns of the variables are depicted. In a GIS, a DEM was used to derive the slope and incoming solar radiation throughout the region, and climate layers (temperature and precipitation indices) were generated by a multivariate geostatistical algorithm, regression kriging, with elevation used as an auxiliary variable. All layers exerted a significant influence on spatial variability, so all are relevant to discriminate the zones considering the full variability of the data. Although a wide part of the Extremaduran territory has warm climate characteristics, the zones have different viticultural potential. Moreover, the analysis of lithology distribution shows that a high proportion of the region lays on suitable substrate. The subzones belonging to the Ribera del Guadiana DO are located within zones in which a high potential for viticulture is apparent, excluding the more mountainous areas of the east. Benefits of this study include suggestions about sites for new vineyard sites, suitability of new varieties and cultivars, adoption of better resource conservation practices, and establishment of regional agricultural policy. Furthermore, zonation at a regional level is the basis for further zoning studies at more detailed field scale and the modeling of vineyard response to climate change.

## Acknowledgements

The authors are very grateful to the reviewers of this paper for providing constructive comments which have contributed to improve the final version.

## References

- Almarza C, 1984. Fichas hídricas normalizadas y otros parámetros hidrometeorológicos, tomo II. National Meteorological Institute, INM. Madrid, Spain.
- Bishop TFA, McBratney AB, 2001. A comparison of prediction methods for the creation of field-extend soil property maps. *Geoderma* 103: 149-160. [http://dx.doi.org/10.1016/S0016-7061\(01\)00074-X](http://dx.doi.org/10.1016/S0016-7061(01)00074-X).
- BOE, 1999. Order of 16 April that ratified the regulation of the regulatory board of the Ribera del Guadiana Denomination of Origin. *Boletín Oficial del Estado* No. 105, 03/05/99.
- Goovaerts P, 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena* 34 (3-4): 227-242. [http://dx.doi.org/10.1016/S0341-8162\(98\)00116-7](http://dx.doi.org/10.1016/S0341-8162(98)00116-7).
- Hall A, Jones GV, 2010. Spatial analysis of climate in winegrape-growing regions in Australia. *Aust J Grape Wine R* 16 (3): 389-404. <http://dx.doi.org/10.1111/j.1755-0238.2010.00100.x>.
- Herrera-Núñez JC, Ramazzotti S, Stagnari F, Pisante M, 2011. A multivariate clustering approach for characterization of the Montepulciano d'Abruzzo Colline Teramane area. *Am J Enol Viticult* 62 (2): 239-244. <http://dx.doi.org/10.5344/ajev.2010.10008>.
- Huglin P, 1978. Nouveau mode d'évaluation des possibilités héliothermiques d'un milieu viticole. *CR Acad Agr France* 64: 1117-1126.
- Liu M, Samal A, 2002. A fuzzy clustering approach to delineate agroecozones. *Ecol Model* 149: 215-228. [http://dx.doi.org/10.1016/S0304-3800\(01\)00446-X](http://dx.doi.org/10.1016/S0304-3800(01)00446-X).
- Mackenzie DE, Christy A, 2005. The role of soil chemistry in wine grape quality and sustainable soil management in vineyards. *Water Sci Technol* 51: 27-37.
- MAGRAMA, 2015. Estadísticas agrarias: Agricultura. Ministerio de Agricultura, Alimentación y Medio Ambiente, Gobierno de España. <http://www.magrama.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/default.aspx>. [April 7 2015].
- Moral FJ, 2010. Comparison of different geostatistical approaches to map climate variables: application to precipitation. *Int J Climatol* 30: 620-631. DOI: 10.1002/joc.1913.
- Moral FJ, Rebollo FJ, Paniagua LL, García A, Honorio F, 2016. Integration of climatic indices in an objective probabilistic model for establishing and mapping viticultural climatic zones in a region. *Theor Appl Climatol* 124: 1033-1043. <http://dx.doi.org/10.1007/s00704-015-1484-0>.
- Munier B, Birr-Pedersen K, Schou JS, 2004. Combined ecological and economic modelling in agricultural land use scenarios. *Ecol Model* 174: 5-18. <http://dx.doi.org/10.1016/j.ecolmodel.2003.12.040>.
- Odeh I, McBratney A, Chittleborough D, 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma* 63 (3-4): 197-214. [http://dx.doi.org/10.1016/0016-7061\(94\)90063-9](http://dx.doi.org/10.1016/0016-7061(94)90063-9).
- Odeh I, McBratney A, Chittleborough D, 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* 67 (3-4): 215-226. [http://dx.doi.org/10.1016/0016-7061\(95\)00007-B](http://dx.doi.org/10.1016/0016-7061(95)00007-B).
- OEMV, 2012. Estadísticas. Observatorio Español del Mercado del Vino. <http://www.oemv.es/esp/estadisticas-3p.php>. [May 23, 2015].
- OIV, 2010. Summary of resolutions adopted by the Eight General Assembly of the Organisation Internationale de la Vigne et du Vin. June 2010. Tbilisi, Georgia.

- Tonietto J, Carbonneau A, 2004. A multicriteria climatic classification system for grape-growing regions worldwide. *Agr Forest Meteorol* 124: 214-251. <http://dx.doi.org/10.1016/j.agrformet.2003.06.001>.
- Williams CL, Hargrove WW, Liebman M, James DE, 2008. Agro-ecoregionalization of Iowa using multivariate geographical clustering. *Agr Ecosyst Environ* 123: 161-174. <http://dx.doi.org/10.1016/j.agee.2007.06.006>.
- Zhou Y, Narumalani S, Waltman WJ, Waltman SW, Palecki MA, 2003. A GIS-based spatial pattern analysis model for ecoregion mapping and characterization. *Int J Geophys Inform Sci* 17: 445-462. <http://dx.doi.org/10.1080/1365881031000086983>.