

# Mapping forest vegetation patterns in an Atlantic–Mediterranean transitional area by integration of ordination and geostatistical techniques

Adriana E. Olthoff<sup>1</sup>, Cristina Gómez<sup>2,3</sup>, Josu G. Alday<sup>4,5</sup> and Carolina Martínez-Ruiz<sup>1,5,\*</sup>

<sup>1</sup> Ecology Area, Agroforestry Sciences Department, E.T.S.II.AA. University of Valladolid, Avda. Madrid 50, 34071 Palencia, Spain

<sup>2</sup> Department of Geography and Environment, School of Geoscience, University of Aberdeen, Aberdeen AB24 3UE, UK

<sup>3</sup> Department of Forest Systems and Resources, CIFOR-INIA, Ctra. de La Coruña km 7.5, 28040 Madrid, Spain

<sup>4</sup> Department of Crop and Forest Sciences-AGROTECNIO Center, University of Lleida, 25198 Lleida, Spain

<sup>5</sup> Sustainable Forest Management Research Institute, University of Valladolid-INIA, E.T.S.II.AA., Avda. Madrid 50, 34071 Palencia, Spain

\*Correspondence address. Ecology Area, Agroforestry Sciences Department, E.T.S.II.AA. University of Valladolid, Avda. Madrid 50, 34071 Palencia, Spain. Tel: +34-9-79-10-83-21; Fax: +34-9-79-10-84-40; E-mail: caromar@agro.uva.es

## Abstract

### Aims

Forest vegetation variability may be explained by the complex interplay among several spatial structuring factors, including climate and topography. We modelled the spatial variability of forest vegetation assemblages and significant environmental variables along a complex environmental gradient or coenocline to produce a detailed cartographic database portraying the distribution of forests along it.

### Methods

We combined an analysis of ordination coenoclines with kriging over 772 field data plots from the third Spanish National Forest Inventory in an Atlantic–Mediterranean transitional area (northern Spain).

### Important Findings

The best fitted empirical semivariogram revealed a strong spatial structure of forest species composition along the complex environmental gradient considered (the climatic–topographic gradient from north to south). The steady and gradual increase of semivariance with a marked lag distance indicates a gradual turnover of

forest assemblages according to the climatic–topographic variations (regional or local). Two changes in the slope of the semi-variogram suggest the existence of two different scales of spatial variation. The interpolation map by Kriging of forest vegetation assemblages along the main coenocline shows a clear spatial distribution pattern of trees and shrubs in accordance with the spatial variation of significant environmental variables. We concluded that the multivariate geostatistical approach is a suitable technique for spatial analysis of forest systems employing data from national forest inventories based on a regular network of field plots. The development of an assortment of maps describing changes in vegetation assemblages and variation in environmental variables is expected to be a suitable tool for an integrated forest management and planning.

**Keywords:** coenocline, Kriging, National Forest Inventory, ordination, variogram

Received: 22 April 2016, Revised: 16 September 2016, Accepted: 13 October 2016

## INTRODUCTION

Describing spatial patterns of ecological structures, such as species distributions, composition or diversity, has long been

a major concern in ecology (Dray *et al.* 2012; Legendre 1993). The spatial structure in plant communities arises from a variety of factors that fall into three broad groups (see Wagner 2003): (i) morphological factors, such as plant size or dispersal

mechanism, which influence the spatial aggregation within a population; (ii) interspecific interactions within a community; and (iii) the response to environmental factors, which are themselves spatially arranged (Dale 2000; Kershaw 1964; Koenig 1999). Hence the composition of species assemblages is an integrative response variable in that the variability may be explained by the complex interplay among several structuring factors. These factors are in turn variables, particularly in areas of heterogeneous environmental conditions (Dray *et al.* 2012). Although determining the specific processes involved in the creation of a given plant pattern is not possible by just observing its spatial distribution (Maestre *et al.* 2005), a thorough analysis of the spatial variation in the assemblage of species may help inferring the processes that shape the ecological communities (Dray *et al.* 2012; Kienel and Kumke 2002; Öpik *et al.* 2014; Park and Lee 2014). It is not surprising that spatial pattern analysis has received substantial attention by plant ecologists in the last decades (Cressie 1993; Dale 2000; Goovaerts 1999; Legendre and Fortin 1989; Maestre *et al.* 2005; Wagner 2003) and that numerous methods for quantifying spatial patterns have been developed (e.g. de la Cruz and Maestre 2013; Dray *et al.* 2012; Goovaerts 1997; Perry *et al.* 2002).

Although plant community composition and diversity result from processes at different levels of biological organization operating in a spatially structured environment (Wagner 2003), the majority of the ecological literature has treated, in the past, spatial structure as a problem rather than as information (see Maestre *et al.* 2005; Wagner 2003). Existing problems for significance tests due to spatial autocorrelation have been stressed, whereas the additional insights that could be gained from spatial analysis have not been given sufficient relevance (Koenig 1999; Legendre 1993; Peres-Neto and Legendre 2010). When ecologists analysed variation in plant communities, they either focused on plant–environment interactions using constrained and unconstrained ordination methods that ignore spatial structure or described the pattern in one or two species without taking into account the spatial structure caused by environmental factors or community-level processes (Wagner 2003). Nowadays, the availability of high-performance computers and the development of permutation procedures permit the use of ordination methods for evaluating ecological hypotheses in an inferential framework (see Manly 1997). More specifically, the incorporation of space as an explicit component in these analyses represents a further step towards the objective of uncovering unmeasured or unmeasurable ecological processes (McIntire and Fajardo 2009).

Ecological studies would benefit from the combined use of the classical statistical models of community composition data, such as constrained or unconstrained multivariate analyses of site-by-species abundance tables, with rapidly emerging and diversifying methods of spatial pattern analysis (Dray *et al.* 2012; Kienel and Kumke 2002; Maestre *et al.* 2005; Park and Lee 2014; Wagner 2003). Hence an integrated methodological

approach is needed to understand what determines community structure and, ultimately, why species coexist (Wagner 2003). More specifically, an integration of geostatistics with general multivariate statistical methods may enable ecologists from a broad range of fields to incorporate spatial structure and processes into their research and to integrate analyses across different levels of biological organization (Dray *et al.* 2012; Wagner 2003).

Geostatistics facilitate inference of the spatial structure of any variable of interest for which some measurements exist, estimating through a modelled variogram (or semivariogram) the scale and direction of spatial variability. Interpolation techniques such as Kriging enable detailed mapping of the spatial distribution of values, providing an estimation of the error committed at each punctual estimation (Gallardo and Maestre 2008; Goovaerts 1999; Park and Lee 2014). On the down side, geostatistical techniques have limited capacity to handle large multivariate data sets and to assess the causes of variation quantitatively (Kienel and Kumke 2002). Considering the main axes of some ordination approach rather than individual variables may to some extent lessen the limitation to deal with large multivariate data sets (Legendre and Fortin 1989; Wackernagel 1995). For quantitative analysis of the causes of variation, partial canonical analysis and variation partitioning may be used (Borcard *et al.* 1992).

There is an increasing awareness of the importance of the spatial distribution of variables and processes in ecology (Helm *et al.* 2014). The capacity to acquire measurements with precise spatial location provided by global positioning systems, as well as to store, analyse, visualize and interpret those data within a geographical information system (GIS), provide for an ever increasing number of data sets spatially referenced and susceptible of spatial analysis (Wagner 2003). This is the case of data sets derived from national forest inventories in many countries. The Spanish National Forest Inventory (SNFI) provides quantitative data of forest species abundance and local environmental variables in a network of geo-referenced field plots across the entire country. Plot data can be used, among other things, for describing and explaining spatial forest vegetation compositional changes along broad heterogeneous environmental gradients. The Palencia province, located in northern Spain, provides an ideal case study for this purpose, as it comprises different biogeographic regions (Atlantic and Mediterranean) and geomorphological units (Cantabrian Range and Castilian plateau) that result in a high variety of landscapes and environmental conditions (Tejero de la Cuesta 1988) and, in turn, in a considerable vegetation diversity (Olthoff *et al.* 2016).

Previous works have identified the main environmental factors responsible for forest species assemblages along the transitional area between the Atlantic and Mediterranean biographic regions in northern Spain (Olthoff *et al.* 2016). However, in this work, we combine the analysis of ordination coenoclines with ordinary Kriging, a powerful geostatistical tool, for description and mapping of the spatial variation of

forests using data from 772 field plots from the third Spanish National Forest Inventory (3SNFI). The aims are (i) to assess the spatial variability of the forest vegetation assemblages by modelling the main complex environmental gradient or coenocline and (ii) to model the main significant environmental variables along the Atlantic–Mediterranean gradient.

## MATERIALS AND METHODS

### Study area

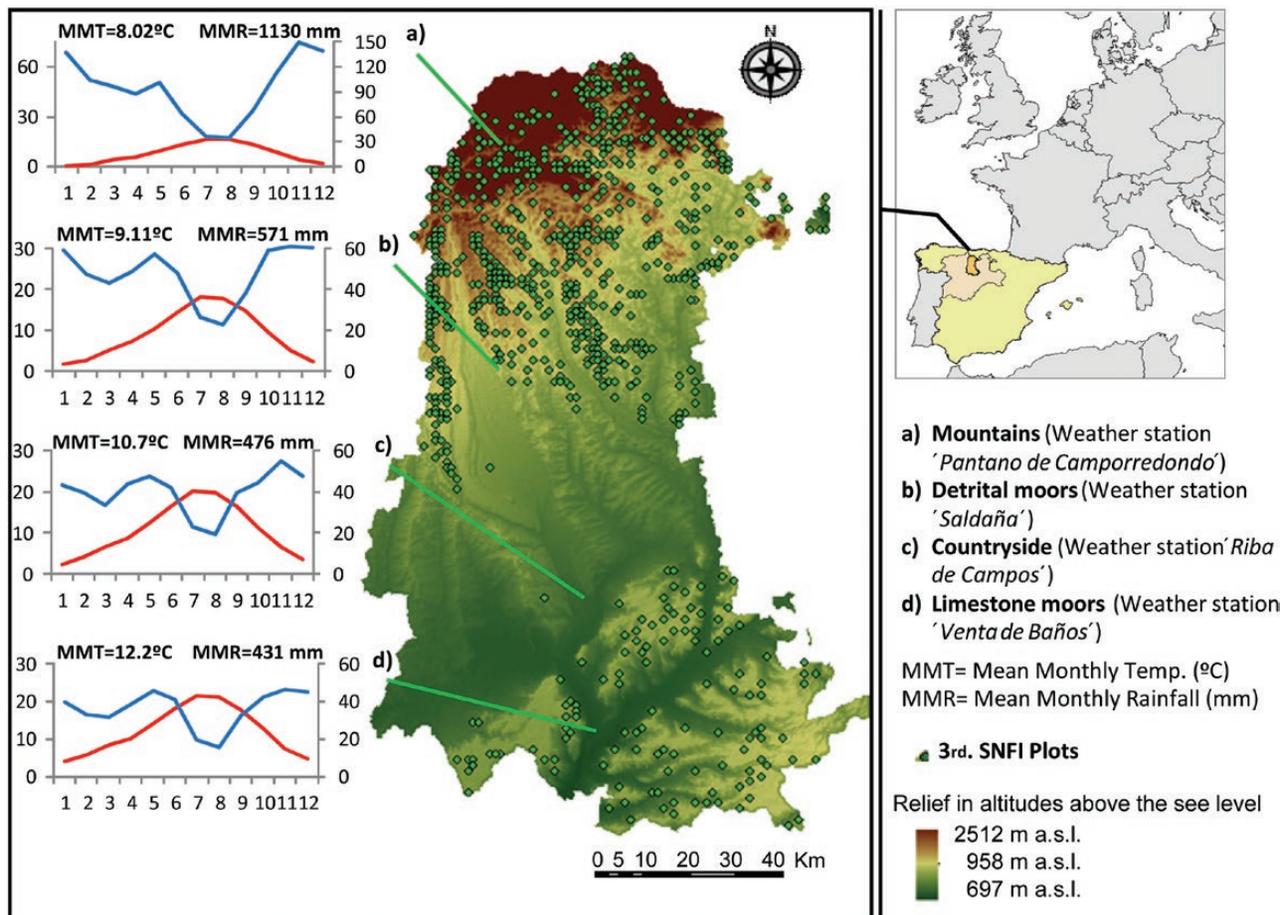
This study was conducted in the Palencia province (43°04'N to 41°46'N latitude and 3°53'W to 5°02'W longitude), which represents a transitional area between the Atlantic and Mediterranean biogeographic regions in northern Spain (Fig. 1). There is a strong climate variation from north to south, along a gradient of 140 km in length, which relates to a complex topographic gradient determined by the presence of the Cantabrian Range in the north (altitude up to 2540 m) and the Castilian plateau (mean altitude ~800 m) in the middle and southern parts of the province. From north to south, there is a notorious increase in temperature as well as a

decrease in precipitation (higher xericity), i.e. an increase in the continental nature of the climatic conditions (see Fig. 1).

Due to the confluence of two biogeographic regions and two geomorphological units, the Palencia province comprises a high variety of landscapes and environmental conditions (Tejero de la Cuesta 1988; Sainz Ollero and Sánchez de Dios 2011), resulting in a great vegetation diversity, particularly of forest systems (Ruiz de la Torre 2002). It should be also noted that this area includes tree species located at its southernmost distribution limit (*Fagus sylvatica* L. and *Quercus petraea* (Matt.) Liebl.) or particularly relevant such as *Quercus pyrenaica* Willd., a typical Mediterranean species almost endemic of the Iberian Peninsula (Ruiz de la Torre 2002). All these circumstances make the climatic and environmental gradient throughout the Palencia province particularly interesting for investigating novel techniques for description and characterization of spatial forest vegetation compositional changes along broad heterogeneous environmental gradients.

### Data source

The entire forest area of Palencia is periodically inventoried with 903 SNFI permanent plots. Data of the woody species



**Figure 1:** Location of the 772 field plots measured by the 3SNFI and considered in the study (full dots). Four morphostructural units are identified in the Palencia province (mountains, detrital moors, countryside and limestone moors); the north to south climatic gradient description is portrayed (a, b, c and d).

measured in 772 permanent field plots as part of the 3SNFI (1997–2007) were considered for our analysis. Plots measured by the 3SNFI in riparian forests (106) were excluded from analysis (see [Olthoff et al. 2016](#)) as their floristic composition is mainly related to the water regime and do not meet the objectives of our study. An additional 25 plots with incomplete measurements of environmental variables were also excluded (see [Olthoff et al. 2016](#)). The SNFI measure circular plots of variable radius (5, 10, 15 and 25 m) systematically distributed and located on the intersection nodes of a 1 × 1 km Universal Transverse Mercator grid; only plots located inside forest areas are measured ([Ministerio de Medio Ambiente 1997–2007](#)). Trees with diameter at breast height of 75, 125, 225 and 425 mm are measured, respectively, in each one of the concentric circumferences. Shrubs are inventoried within a fixed plot radius of 10 m ([Torrás et al. 2012](#)). Tree and shrub species are differentiated on the basis of a predefined list available in SNFI publications ([Ministerio de Medio Ambiente 1997–2007](#)), which was elaborated based on the growth form and on the minimum height reached by each species at maturity (this height needs to be at least 5 m for the species to be considered a tree). The cover (%) of a total of 124 woody species (31% trees and 69% shrubs) from 25 families was registered for the selected set of plots ([Olthoff et al. 2016](#)).

A preliminary indirect ordination technique (detrended correspondence analysis [DCA]) was applied, using CANOCO 4.5 ([Ter Braak and Šmilauer 2002](#)), on the cover (%) matrix of all woody species identified in each of the 772 selected plots, to extract the primary ordination axes (see [Olthoff et al. 2016](#)). The ordination axes obtained may be considered as standardized, abstract complex environmental gradients or coenoclines ([Lawesson and Oksanen 2002](#)). Moreover, it was possible to conclude that forest vegetation compositional changes along the gradient were mainly determined by the north–south topographic–climatic differences (DCA1, primary coenocline), although a secondary coenocline also showed a turnover of species in relation to the slope gradient (DCA2, steepness); particularly, rainfall and altitude increasing and temperature decreasing towards DCA1 left end and steepness increasing towards DCA2 upper end (see [Table 1](#) and [Olthoff et al. 2016](#)).

### Geostatistical analysis

Once the intrinsic gradients in species composition (coenoclines) through the Atlantic–Mediterranean gradient were

**Table 1:** significant environmental variables scores for DCA1 and DCA2 axes and the goodness-of-fit statistic

Variables	DCA1	DCA2	$r^2$	$P$
Rainfall	−0.633	0.775	0.51	0.001
Temperature	0.718	−0.696	0.58	0.001
Altitude	−0.675	0.738	0.46	0.001
Steepness	−0.217	0.976	0.30	0.001

identified, the spatial variation of the first ordination axis (DCA1) was analysed by modelling the variogram of the 772 field plots site scores as well as the values of the most significant environmental variables (temperature, precipitation, altitude and steepness). The aim was to describe the patterns of forest vegetation systems along the primary coenocline ([Legendre 1993](#); [Wagner 2003](#)). An empirical semivariogram with spherical model was applied in all cases. The empirical semivariogram was fitted to spherical model because it reaches a true ‘sill’ (i.e. maximum variance between pairs of points) at a distance equivalent to its range, and it was best fitted to the actual data, providing higher reliability. The range of spatial correlation for the variogram computation was the range of the target gradient (see [Nanos et al. 2005](#)), i.e. 140 km (the length of gradient described from north to south). We considered adequate number of 20 ‘lags’ to calculate the average nearest neighbour, i.e. the ‘lag’ size necessary to build the semivariogram, which in our case was 1.5 km. The absence of anisotropy was previously explored by the computation of four semivariograms with spherical model in the directions N–S (0°), NE–SO (45°), E–O (90°) and NO–SE (135°) with angular tolerance of 22.5° ([Gallardo and Maestre 2008](#)); since ranges and nuggets were similar in the four directional semivariograms, our data can be considered isotropic.

Ordinary Kriging spatial interpolation method for subsequent mapping and analysing of spatial patterns was then applied ([Gallardo and Maestre 2008](#)). Ordinary Kriging is a frequently used interpolation technique that assures the estimation is unbiased and has minimum variance ([Cressie 1993](#)). Ordinary Kriging allows accounting for local variation of the mean, by limiting its domain stationarity to the local neighbourhood centred on the point being estimated. Also, ordinary Kriging estimates provide a good representation of the data fluctuations ([Li and Heap 2008](#)) and vary proportionally to the local data average ([Goovaerts 1997](#)).

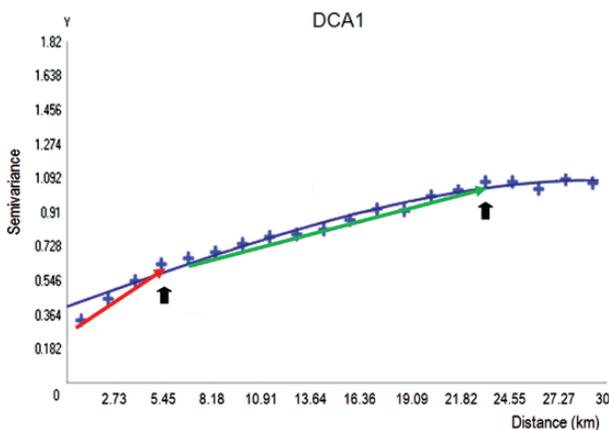
Finally, a cross-validation procedure, a jackknife method to check the model assumptions used in the Kriging, was applied to test the effect of all decisions taken during the estimation of the variable in the space, i.e. to examine the variogram models ([Gallardo and Maestre 2008](#)). Geostatistical analyses were conducted in ArcGIS10 (ESRI, Redlands, CA, USA).

## RESULTS AND DISCUSSION

The empirical semivariogram ([Fig. 2](#)) revealed a strong spatial structure for forest species composition along the complex environmental gradient studied, i.e. the climatic–topographic gradient (primary coenocline). In general, there is a steady gradual increase of semivariance with lag distance that suggests a gradual turnover of forest species assemblages from north to south according to the climatic–topographic variations (DCA1). This result is consistent with the idea that boundaries between ecosystems are diffuse since vegetation responds in a non-linear manner to environmental factors ([Danz et al. 2013](#)). Nevertheless, the semivariogram also

showed two main changes in the slope of the semivariance: the first at around 5 km of lag distance and the second at ~23 km (Fig. 2), which could indicate the existence of two scales of spatial variation (see Nanos et al. 2002). The first slope change can be attributed to the highest similarity between plots belonging to the same type of forest with respect to others, particularly into the mountains of the north, whereas the second slope change (at 23 km) might represent variability at regional scale, particularly through differences in forest floristic composition between the mountains of the north and the middle and southern parts of the gradient.

The semivariogram also exhibits an apparent discontinuity at the origin (Fig. 2), the nugget or semivariance at zero lag distance is not zero (see Table 2). Although in practice the semivariance is rarely zero at zero lag distance (Cressie 1993), the magnitude of this discontinuity in our variogram (the so-called ‘nugget effect’) suggests the existence of spatial processes operating at distances shorter than the smallest lag observed in the data set (see Kienel and Kumke 2002; Park and Lee 2014; Zhao et al. 2009). In particular, the nugget effect found in our variogram suggests that the spatial



**Figure 2:** Empirical semivariogram of the DCA axis 1 scores (primary coenocline). Regionalization fitted by spherical model (solid line) with 99.23% of reliability according with its average standard error prediction (MSE = 0.77). Black arrows indicate the two main changes in the slope of the semivariance, at around 5 and 23 km of lag distance.

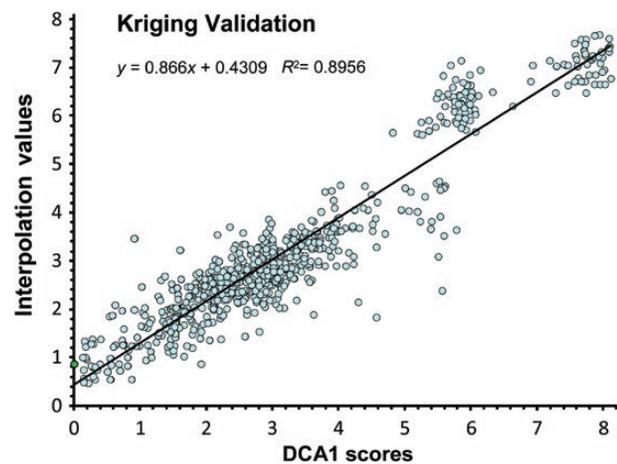
**Table 2:** variogram model parameters, RMS, RMSS and mean standard error for DCA axis 1 (main coenocline)

	DCA axis 1
Model	Spherical
Range (km)	140
Sill	2.21
Nugget	0.48
RMS	0.77
RMSS	0.98
MSE	0.77

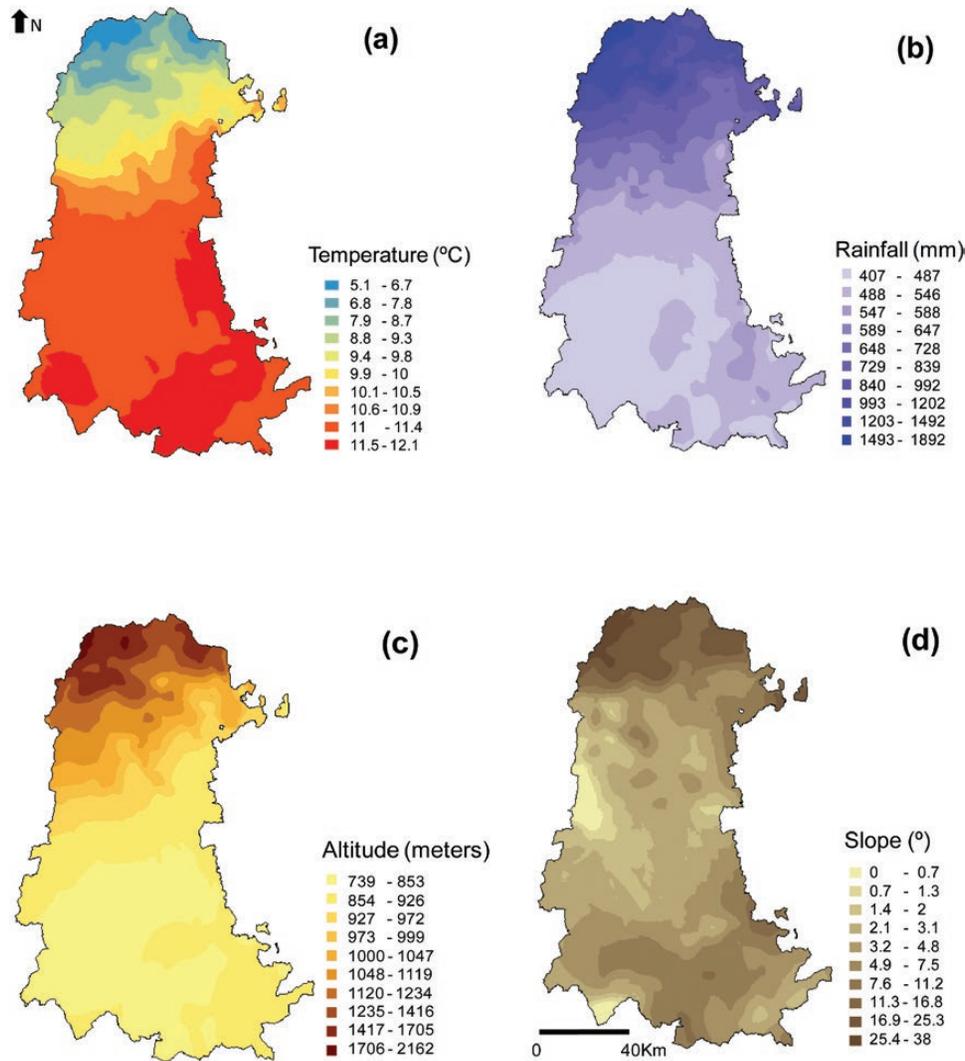
distribution of many tree and shrub species vary at a shorter distance than the 1 km field sampling interval of the SNFI plots or that species have a too low occurrence frequency (see Olthoff et al. 2016). Nevertheless, the local scale variation, nugget variance, calculated as the ratio of the ‘nugget’ to ‘sill’, is about 22% for site scores along DCA1, typical of data obtained from spatially distributed field measurements (see Kienel and Kumke 2002), and in agreement with the variation explained by finer scale spatial variables over the same data set (25%, see Olthoff et al. 2016). The niche partitioning processes that dominate this gradient are therefore well captured by the variogram.

The accuracy of the spatial model for site scores along DCA1 was examined using a cross-validation procedure (Fig. 3). Cross-validation yielded a correlation coefficient of 0.94 ( $P < 0.001$ ), indicating that the quality of the interpolation map by ordinary Kriging was significant. The interpolation map by Kriging for forest vegetation assemblages along DCA1 coenocline reached a root mean square (RMS) of 0.77 and a root mean square standardized (RMSS) of 0.98 (Table 2). It is known that an RMS ~1 indicates a good fit of the model and its parameters (Cerri et al. 2000) and an RMSS ~1 indicates that the prediction errors are valid. Both values provide a complete picture of the error distribution in the Kriging map (Chai and Draxler 2014).

The spatial variation of the most significant environmental variables through the Palencia province over 772 field data plots from the 3SNFI (Fig. 4) reflected the well-known north–south gradient in environmental conditions (Tejero de la Cuesta 1988). The interpolation map of forest vegetation along DCA1 coenocline (Fig. 5) showed a spatial distribution pattern of forest vegetation assemblages (trees and shrubs) from north to south of the gradient, in accordance with the spatial variation of the most significant environmental variables along it. A number of apparently different areas can be



**Figure 3:** correlation between the scores of 772 field plots along DCA1 coenocline (measured values) and the post-interpolation values (predicted values) using cross-validation.

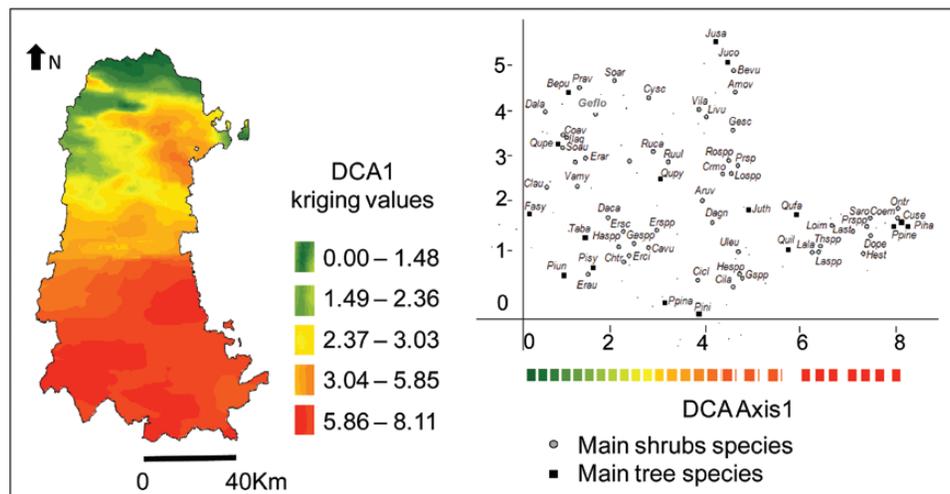


**Figure 4:** interpolation maps by ordinary Kriging for more significant environmental variables (temperature, rainfall, altitude and steepness) through the Palencia province, over 772 field data plots from the 3SNFI. Data of annual rainfall and mean temperature were obtained from the Agroclimatic Atlas of ‘Castilla y León’ (ITACyL-AEMET 2013) and data of altitude and steepness were obtained from the Digital Terrain Model (DTM 200 m) of the National Geographic Institute of Spain (<http://centrodedescargas.cnig.es/CentroDescargas/catalogo.do>), using in both cases the 3D analysis tool in ArcGIS (version 10, ESRI, Redlands, CA, USA).

distinguished analysing the distribution of the DCA1 krigged values. From 0 to 1.48 estimated values occurring in the northern mountains with altitude between 1000 and 2000 m and slope between 16 and 38° (Fig. 4c and d) and where annual rainfall is of 1000–1600 mm and annual average temperature of 4–8°C (Fig. 4a and b), the Atlantic deciduous broadleaf forests of *F. sylvatica* and *Q. petraea* dominate, with an understory composed by *Daphne laureola* L., *Erica arborea*, *Vaccinium myrtillus* L. and *Corylus avellana* L. This complex understory is also highly mixed with *Sorbus aucuparia* L., *Prunus avium* L. and *Ilex aquifolium* L. under the *Betula pubescens* Ehrh. forests located in the moist valleys of the mountains (DCA1 top left end, Fig. 5). Atlantic deciduous broadleaf forests also coexist with natural relicts of *Taxus baccata* L. and *Pinus sylvestris* in these northern mountains.

Between 1.49 and 2.36 interpolated values, both naturally grown and plantations of *Pinus sylvestris* and a few *Pinus uncinata* plantations appear with an understory dominated by Atlantic shrubs such as *Corylus avellana*, *Daboecia cantabrica* (Huds.) K.Koch, *Chamaespartium tridentatum* (L.) P.E.Gibbs, and *Erica* L. and *Halimium* (Dunal) Spach species. These species constitute the substitution shrubby understory of more humid woodlands, occupy areas with annual rainfall ~700–1000 mm and annual average temperature of 8–9°C in the foothills and with 550–700 mm (annual rainfall) and 9–10°C (annual average temperature) in the detrital moors (DCA1 bottom left end, Fig. 5) where the altitude decreases about 800 m (Fig. 4c) with smooth slopes (1–7°, Fig. 4d).

From 2.37 to 3.03 interpolated values, *Pinus sylvestris* and *Pinus uncinata* woodlands coexist with *Q. pyrenaica* forests



**Figure 5:** interpolation map by ordinary Kriging for forest vegetation (trees and shrubs) assemblages along DCA1 coenocline (i.e. the climatic–topographic gradient from north to south of the Palencia province) and main tree and shrub species along the first two axes of a DCA. Species codes: Amov = *Amelanchier ovalis* Medik., Aruv = *Arctostaphylos uva-ursi* (L.) Spreng., Bepu = *Betula pubescens* Ehrh., Cavu = *Calluna vulgaris* (L.) Hull, Chtr = *Chamaespartium tridentatum* (L.) P.E.Gibbs, Cici = *Cistus clusii* Dunal, Cila = *Cistus ladanifer* L., Clau = *Cistus laurifolius* L., Coem = *Coronilla emerus* L., Coav = *Corylus avellana* L., Crmo = *Crataegus monogyna* Jacq., Cusc = *Cupressus sempervirens* L., Cysc = *Cytisus scoparius* (L.) Link, Daca = *Daboecia cantabrica* (Huds.) K.Koch, Dagn = *Daphne gnidium* (L.) Dala = *Daphne laureola* L., Dope = *Dorycnium pentaphyllum* Scop., Erar = *Erica arborea* L., Erau = *Erica australis* L., Erci = *Erica cinerea* L., Ersch = *Erica scoparia* L., Erspp = *Erica* spp., Fasy = *Fagus sylvatica* L., Geflo = *Genista florida* L., Gesc = *Genista scorpius* (L.) DC., Gespp = *Genista species* L., Gsp = *Genistella* spp., Haspp = *Halimium* spp. (Dunal) Spach, Hest = *Helichrysum stoechas* (L.) Moench, Ilaq = *Ilex aquifolium* L., Juco = *Juniperus communis* L., Jusa = *Juniperus sabina* L., Juth = *Juniperus thurifera* L., Lala = *Lavandula latifolia* Medik., Laspp = *Lavandula species* L., Last = *Lavandula stoechas* L., Livu = *Ligustrum vulgare* L., Loim = *Lonicera implexa* Aiton, Lospp = *Lonicera* spp. L., Ontr = *Ononis tridentata* L., Piha = *Pinus halepensis* Mill., Pini = *Pinus nigra* J.F.Arnold, Ppina = *Pinus pinaster* Aiton, Ppine = *Pinus pinea* L., Pisy = *Pinus sylvestris* L., Piun = *Pinus uncinata* Mill. ex Mirb., Prsp = *Prunus spinosa* L., Prspp = *Prunus* spp. L., Qufa = *Quercus faginea* Lam., Quil = *Quercus ilex* subsp. *ballota* (Desf.) Benth., Samp., Qupe = *Quercus petraea* (Matt.) Liebl., Qupy = *Quercus pyrenaica* Willd., Rospp = *Rosa* spp. L., Ruca = *Rubus caesius* L., Ruul = *Rubus ulmifolius* Schott., Saro = *Santolina rosmarinifolia* L., Soar = *Sorbus aria* (L.) Crantz, Soau = *Sorbus aucuparia* L., Taba = *Taxus baccata* L., Thspp = *Thymus* spp. L., Uleu = *Ulex europaeus* L., Vamy = *Vaccinium myrtillus* L., Vila = *Viburnum lantana* L.

that are dominant in the area defined between 3.04 and 5.85 DCA1 interpolated values. *Quercus pyrenaica* forests are representative of the Mediterranean–Iberian–Atlantic floristic domain (López Leiva et al. 2009), together with transitional vegetation between the Atlantic and Mediterranean biogeographical regions, such as scrub forest fringes and wet understorey. Annual rainfall is ~500–600 mm and annual average temperature of 11–12°C (Fig. 4a and b). The main shrubs are *Viburnum lantana* L., *Ligustrum vulgare* L., *Prunus spinosa* L., *Lonicera* L. species, *Crataegus monogyna* Jacq. and *Genista scorpius* Lam., among others (DCA1 middle part, Fig. 5). As regards the conservation of Mediterranean forests of *Q. pyrenaica*, Álvarez et al. (2009) suggest that management plans should aim to maintain a landscape with a mosaic of successional stages, since each stage has different structure and above all a distinct specific composition.

Also in the middle part of DCA1 *Juniperus* spp., as representative of oro-Mediterranean vegetation, are located on high plateaus and Mediterranean mountains with continental climate, with understorey of *Amelanchier ovalis* and *Ligustrum vulgare* on locations with large rocky outcrops and *G. scorpius* in the foothills of the slopes, whereas *Pinus nigra* and *Pinus pinaster* Aiton woodlands dominate in flat areas,

and the more xerophytic *Helianthemum* species constitute the understorey of *Pinus nigra* woodlands.

In limestone moors (southern area), the interpolated values range between 5.86 and 8.11): here, typical Mediterranean conditions (<400 mm of annual rainfall and 15°C of annual average temperature, Fig. 4a and b), with a more intense summer drought determines the presence of the *Quercus ilex* subsp. *ballota* (Desf.) Benth., sclerophyllous species typical of the Mediterranean region. This species coexists with *Quercus faginea*, more adapted than *Q. ilex* subsp. *ballota* to loamy calcareous soils and that in colder sites with shallow soils is replaced by *Juniperus thurifera* forests with limited presence and relict. Conifer woodlands consist of *Pinus halepensis* Mill., *Pinus pinea* L. and *Cupressus sempervirens* L. plantations, with an understorey dominated by heliophyllous and summer drought shrub species such as *Prunus* spp., *Santolina rosmarinifolia* and *Ononis tridentata* (DCA1 right end), being *O. tridentata* typical of gypsum soils. Understorey shrub species typical of calcareous soils under *Q. ilex* subsp. *ballota* forests are *Genista scorpius*, *Dorycnium pentaphyllum*, *Lavandula latifolia* Medik. and *Thymus* L. spp., whereas under *Q. faginea* forests dominated *Rosa* L. spp., *Prunus spinosa* and also *G. scorpius*.

Our results clearly show the benefits of combining ordination techniques and geostatistics in contrast with other (multivariate) approaches to model species assemblages distribution (see Olthoff *et al.* 2016). The use of Kriging techniques on main axis of DCA ordination approach instead of using single variables to model forests distribution along it reduces the limited capacity of geostatistical techniques to handle large multivariate data sets (Legendre and Fortin 1989; Wackernagel 1995) and enables the incorporation of space as an explicit component in the analysis (McIntire and Fajardo 2009), which is not possible with ordination methods only since they ignore spatial structure, taking into account the spatial structure caused by environmental factors and/or community-level processes (Wagner 2003).

### Management implications

Transitional areas between different biogeographic regions are especially relevant in terms of plant diversity and sensitivity to climate change (Cadenasso *et al.* 2003; Danz *et al.* 2013). Therefore, thorough information on current vegetation patterns in these areas is especially important to track the effects of global change and human impact on ecosystems (Park and Lee 2014) and to preserve plant biodiversity in the future (Amici *et al.* 2015; Helm *et al.* 2014; Hernández-Stefanoni *et al.* 2012).

The increasing amount of information on climate, soil properties and human influence enables building better models to reveal the underlying factors that shape vegetation patterns (Helm *et al.* 2014; Saura *et al.* 2011), and geostatistics represents a suitable method for understanding spatial variability and interpolation of vegetation distribution maps (Park and Lee 2014). Taking into account that many of the experimental forest variables (e.g. basal area, trunk volume and leaf area index) are spatially structured, it is obvious that the empirical value of geostatistics in resource planning and other problems encountered in forest management (Kraus and Krumm 2013) could be addressed in future research from forest inventory data (Chica-Olmo 2005). Moreover, geostatistics can also profit from GIS capabilities in data acquisition and manipulation (Rivera *et al.* 2011), and thematic maps derived with data from different sources, including aerial photographs and satellite imagery (Maestre *et al.* 2005), can be integrated and used for the spatial analysis of vegetation over large geographical areas.

Therefore, thematic maps provide explicit spatially arranged information, summarising ecological and environmental data of the territory under study. These maps are very useful and practical for planning environmental research, resource management and conservation. Here, the forest vegetation and environmental variables distribution maps along an Atlantic–Mediterranean gradient might inform practitioners in forest management decision-making process.

### FUNDING

J.G.A. was supported by the program ‘Juan de la Cierva-Incorporación’ fellowships (IJCI-2014-21393).

### ACKNOWLEDGEMENTS

We thank Dr David Green (Geosciences School, University of Aberdeen, UK) for facilitating the use of GIS Laboratory and for his ideas in the first steps of this work.

*Conflict of interest statement.* None declared.

### REFERENCES

- Álvarez R, Muñóz A, Pesqueira XM, *et al.* (2009) Spatial and temporal patterns in structure and diversity of Mediterranean forest of *Quercus pyrenaica* in relation to fire. *For Ecol Manage* **257**:1596–602.
- Amici V, Rocchini D, Filibeck G, *et al.* (2015) Landscape structure effects on forest plant diversity at local scale: exploring the role of spatial extent. *Ecol Complex* **21**:44–52.
- Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* **73**:1045–55.
- Cadenasso ML, Pickett STA, Weathers KC, *et al.* (2003) An interdisciplinary and synthetic approach to ecological boundaries. *BioScience* **53**:717–22.
- Cerri C, Bernoux M, Arrouays D, *et al.* (2000) Carbon stocks in soils of the Brazilian Amazon. In Lal R, Kimble J, Stewart B (eds). *Global Climate Change and Tropical Ecosystems*. Boca Raton, FL: CRC-Lewis Publishers, 33–50.
- Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci Model Dev* **7**:1247–50.
- Chica-Olmo M (2005) La geoestadística como herramienta de análisis espacial de datos del inventario forestal. *Cuad Soc Esp Cien For* **19**:47–55.
- Cressie N (1993) *Statistics for Spatial Data*. New York, NY: Wiley.
- Dale M (2000) *Spatial Pattern Analysis in Plant Ecology*. Cambridge: Cambridge University Press.
- Danz N, Frelich L, Reich P, *et al.* (2013) Do vegetation boundaries display smooth or abrupt spatial transitions along environmental gradients? Evidence from the prairie-forest biome boundary of historic Minnesota, USA. *J Veg Sci* **24**:1129–40.
- de la Cruz M, Maestre FT (2013) *Avances en el Análisis Espacial de Datos Ecológicos: Aspectos Metodológicos y Aplicados*. Móstoles, Spain: Asociación Española de Ecología Terrestre.
- Dray S, Péliissier R, Couteron P, *et al.* (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecol Monogr* **82**:257–75.
- Gallardo A, Maestre F (2008) Métodos geoestadísticos para el análisis de datos ecológicos espacialmente explícitos. In Maestre F, Escudero A, Bonet A (eds). *Introducción al Análisis Espacial de Datos en Ecología y Ciencias Ambientales: Métodos y Aplicaciones*, Madrid, Spain: Universidad Rey Juan Carlos & Dykinson S.L., 216–72.
- Goovaerts P (1997) *Geostatistics for Natural Resources Evaluation*. New York, NY: Oxford University Press.
- Goovaerts P (1999) Geostatistics in soil science: state of the art and perspectives. *Geoderma* **89**:1–45.
- Helm A, Kalamees R, Zobel M (2014) Vegetation patterns and their underlying processes: where are we now? *J Veg Sci* **25**:113–6.
- Hernández-Stefanoni JL, Gallardo-Cruz JA, Meave J, *et al.* (2012) Modeling  $\alpha$ - and  $\beta$ -diversity in a tropical forest from remotely sensed and spatial data. *Int J App Earth Observ Geoinf* **19**:359–68.

- Kershaw KA (1964) *Quantitative and Dynamic Ecology*. London, UK: Edward Arnold.
- Kienel U, Kumke T (2002) Combining ordination techniques and geostatistics to determine the patterns of diatom distributions at Lake Lama, Central Siberia. *J Paleolimnol* **28**:181–94.
- Koenig WD (1999) Spatial autocorrelation of ecological phenomena. *Trends Ecol Evol* **14**:22–6.
- Kraus D, Krumm F (2013) *Integrative Approaches as An Opportunity for the Conservation of Forest Biodiversity*. Germany: European Forest Institute.
- Lawesson J, Oksanen J (2002) Niche characteristics of Danish woody species as derived from coenoclines. *J Veg Sci* **13**:279–90.
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**:1659–73.
- Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. *Vegetatio* **80**:107–38.
- Li J, Heap A (2008) *A Review of Spatial Interpolation Methods for Environmental Scientists*. Record 2008/23. Australia: Geoscience Australia.
- López Leiva C, Espinosa Rincón J, Bengoa Mtz de Mandojana J (2009) *Mapa de Vegetación de Castilla y León. Síntesis 1:400.000*. Spain: Junta de Castilla y León, Consejería de Medio Ambiente.
- Maestre F, Rodríguez F, Bautista S, et al. (2005) Spatial associations and patterns of perennial vegetation in semi-arid steppe: a multivariate geostatistics approach. *Plant Ecol* **179**:133–47.
- Manly BFJ (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology*. London, UK: Chapman and Hall.
- McIntire EJ, Fajardo A (2009) Beyond description: the active and effective way to infer processes from spatial patterns. *Ecology* **90**:46–56.
- Ministerio de Medio Ambiente (1997–2007) *Tercer Inventario Forestal Nacional*. Madrid, Spain: Dirección General para la Biodiversidad.
- Nanos N, Gil L, Montero G (2002) Análisis espacial de los datos del inventario forestal nacional utilizando técnicas geoestadísticas. In Bravo F, del Río M, del Peso C (eds). *El Inventario Forestal Nacional, Elemento Clave para la Gestión Forestal Sostenible*, Spain: Fundación General Universidad de Valladolid, 149–158.
- Nanos N, Pardo F, Alonso Nager J, et al. (2005) Using multivariate factorial kriging for multiscale ordination: a case study. *Can J Forest Res* **35**:2860–74.
- Olthoff A, Martínez-Ruiz C, Alday J (2016) Distribution patterns of forest species along an Atlantic-Mediterranean environmental gradient: an approach from forest inventory data. *Forestry* **86**: 46–54.
- Olthoff A, Martínez-Ruiz C, Alday J (2016) Distribution patterns of forest species along an Atlantic-Mediterranean environmental gradient: an approach from the third Spanish National forest inventory data. *Forestry* **201**:46–54.
- Öpik M, de Bello F, Price JN, et al. (2014) New insights into vegetation patterns and processes. *New Phytol* **201**:383–7.
- Park J, Lee E (2014) Geostatistical analyses and spatial distribution patterns of tundra vegetation in Council, Alaska. *J Ecol Environ* **37**:53–60.
- Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecol Biogeogr* **19**:174–84.
- Perry JN, Liebhold AM, Rosenberg MS, et al. (2002) Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* **25**:578–600.
- Rivera S, Shultz L, Hernández A, et al. (2011) GIS ordination approach to model distribution of shrubs species in Northern Utah. *Natural Resources and Environmental Issues* **17**:263–74.
- Ruiz de la Torre J (2002) Vegetación forestal española. In Pineda F, de Miguel J, Casado M, et al. (eds). *La Diversidad Biológica de España*, Madrid: Prentice Hall, 65–79.
- Sainz Ollero H, Sánchez de Dios R (2011) Diversidad de paisajes españoles. *Memorias de la Real Sociedad Española de Historia Natural* **9**:109–55.
- Saura S, Vogt P, Velázquez J, et al. (2011) Key structural forest connectors can be identified by combining landscape spatial pattern and network analyses. *For Ecol Manage* **262**:150–60.
- Tejero de la Cuesta J (1988) *Análisis del Medio Físico de Palencia: Delimitación de Unidades y Estructura Territorial*. Junta de Castilla y León, Consejería de Fomento, Spain
- Ter Braak CJF, Šmilauer P (2002) *CANOCO Reference Manual and User's Guide to Canoco for Windows: Software for Canonical Community Ordination (Version 4.5)*. Ithaca, NY: Microcomputer Power.
- Torras O, Gil A, Saura S (2012) Changes in biodiversity indicators in managed and unmanaged forests in NE Spain. *J For Res* **17**:19–29.
- Wackernagel H (1995) *Multivariate Geostatistics: An Introduction With Applications*. Berlin: Springer.
- Wagner H (2003) Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. *Ecology* **84**:1045–57.
- Zhao X, Wang Q, Kakubari Y (2009) Stand-scale spatial patterns of soil microbial biomass in natural cold-temperate beech forests along an elevation gradient. *Soil Biol Biochem* **41**:1466–74.