Abrupt fragmentation thresholds of eight zonal forest

types in mainland Spain

Gabriel del Barrio^a*, Helios Sainz^b, Maria E. Sanjuán^a, Rut Sánchez de Dios^c, Jaime Martínez-Valderrama^{a,d} and Alberto Ruiz^a

^a Estación Experimental de Zonas Áridas (CSIC)
 Sacramento Rd.
 04120 La Cañada de San Urbano
 Almería, Spain
 <u>marieta@eeza.csic.es</u> (M.E.S.)
 <u>jaimonides@eeza.csic.es</u> (J.M.V.)
 <u>aruiz@eeza.csic.es</u> (A.R.)

^b Departamento de Biología (Universidad Autónoma de Madrid)
 Darwin St.
 28049 Cantoblanco
 Madrid, Spain
 helios.sainz@uam.es

^c Departamento de Biodiversidad, Ecología y Evolución (Universidad Complutense de Madrid)
 12 José Antonio Novais St.
 28040 Madrid, Spain
 <u>rut.sanchez@ucm.es</u>

^d Present address:
 Instituto Multidisciplinar para el Estudio del Medio "Ramón Margalef" (Universidad de Alicante)
 San Vicente del Raspeig Rd.
 03690 San Vicente del Raspeig
 Alicante, Spain

Corresponding author:
 gabriel@eeza.csic.es
 Phone: +34 696 214 827

Author contributions

Gabriel del Barrio: Conceptualization, Methodology, Formal analysis, Writing—original draft preparation, Writing—review and editing, Supervision, Project administration, Funding acquisition. Helios Sainz: Validation, Investigation, Writing—original draft preparation, Supervision, Project administration, Funding acquisition. María E. Sanjuán: Methodology, Validation, Data curation, Visualization. Rut Sánchez de Dios: Validation, Investigation, Data curation. Jaime Martínez-Valderrama: Investigation, Writing—original draft preparation, Writing—review and editing. Alberto Ruiz: Software, Resources, Visualization.

Declaration of interest

Declarations of interest: none

Funding

This work was supported by Tragsatec (Grupo Tragsa) [contract numbers 39.916 and 38.454]; and the Spanish National R&D&I Plan (including ERDF funds) [grant number CGL2016-78075-P]. The funders had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; nor in the decision to submit the article for publication.

Highlights

We measured fragmentation of eight zonally distributed forest habitat types in mainland
 Spain by fitting power laws of cumulative patch frequency over patch size.

2. The Korcak exponent, if determined by segmented regression of log-log transformations of the raw data, enables scale breakpoints in a spatial structure to be found. These area thresholds become new meaningful indicators of the underlying causes of fragmentation.

3. Unexpectedly, all eight forest types had abrupt breakpoints in their respective fitted functions, showing that relatively large patches become very rare beyond a certain, always small (27-101 ha) area threshold.

4. General interpretations coincide on landscape structures as the spatial memory of current and past events of human pressure on forest resources. The method and its resulting indicators can potentially produce forest-specific hypotheses based on patch size.

Abstract:

This study quantifies patchiness of eight types of zonal forests in three biogeographic regions of mainland Spain (Atlantic, Alpine and Mediterranean) which together occupy 1,726,578 ha. Their dominant species and European Habitat Type codes (EU Directive 92/43 EEC) are: Fagus sylvatica (9120, 9130 and 9150), Quercus robur and Q. pyrenaica (9230), Q. suber (9330), Pinus uncinata (9430), P. nigra ssp. salzmannii (9530) and P. pinea (subset of 9540). We applied the Korcak's exponent B, which describes a hyperbolic relationship between the cumulative frequency of the number of patches and their sizes. The objectives were: 1) detect possible patch size intervals in which B varies significantly, explicitly identifying area thresholds, and 2) contribute to development of a robust forest mass fragmentation indicator. Exponent B was found by segmented regression analysis. The vector data were extracted from a filtered version of the Spanish Forest Map 1:50,000 (1997-2006). After validating the procedure by applying it to a previously published dataset, we found that in all cases the patch size range could be split into two significant intervals around relatively small threshold areas of 27-101 ha. In the one on the left, the rate at which relatively large patches become less abundant was always very slow (B =0.017-0.094). After this threshold had been passed, the rate increased abruptly (B = 1.100-2.590). Both this high fragmentation and its lack of parsimony were unexpected in zonal forest types. General interpretations converge to the coexistence of forest patches of different ages due to human pressure events.

Graphical abstract



Keywords:

Ecological thresholds; Forest patchiness; Habitat types of European interest; Korcak exponent;

Natura 2000 network; Segmented regression.

1. Introduction

It is known that European forests are highly fragmented (Estreguil et al., 2013). The Spanish landscape, in particular, has been constantly occupied, modified and managed by human populations during the entire recent Holocene (Stevenson and Harrison, 1992). Natural vegetation has been relegated to marginal zones, and the intensity of these modifications has fluctuated over the course of time depending on the climate and socio-economic factors (Valbuena-Carabaña et al., 2010). Considering the magnitudes of time that it takes for these forest formations to be generated, deviations in the current landscape from their potential without any human influence collect the memory of recent millennia.

Some general transformation factors include the importance of grazing and firewood collection by Berber communities in the Betic Mountains (10th-16th centuries), the intensity of transhumance of merino sheep and Mesta privileges (13th-18th centuries) (Klein, 1920), use of fire in military tactics (8th-15th centuries), especially during the *Reconquista* Wars (Deacon, 1999), and naval construction (15th-17th centuries) (Merino Navarro, 1981). More recently, stages of intense exploitation, such as during the peak of rural population at the end of the 19th century, or after the Civil War in the first half of the 20th century, have alternated with stages of relative abandonment, such as in the transition from the 20th to 21st century after joining the European Union, which has led to an evident recovery of forested areas (del Barrio, Puigdefabregas, Sanjuán, Stellmes, & Ruiz, 2010; Hill, Stellmes, Udelhoven, Röder, & Sommer, 2008).

Forest habitats referred to in this article are zonal, and are therefore defined on relatively small scales in which abiotic aspects predominate, especially climate, soil and meso-topography (Breckle, 2002). Therefore, as they occupy the space available, they are located in optimum intervals in those environmental gradients (Thuiller et al., 2003), and by doing so, it may be expected for their spatial structure to be in consonance with them. That spatial structure is reflected in, among other properties, the frequency distribution of patch sizes.

If this is true, deviations with respect to the expected relationship in sizes and frequencies can be attributed to other ecosystems, which having similar environmental requirements, can compete with, or exploit, the forest in question, impeding its optimum distribution. This study assumes that the main competing or exploiting ecosystems are the most productive uses of the soil. Our hypothesis is that the effects of those interactions are detected as discontinuities in the relationship of sizes and frequencies of the forests affected.

The size of the individual patches of an ecosystem, and especially, the distribution of size frequencies, indicate their persistence over time (Hastings et al., 1982). Size is proportional to resilience and homeostasis, facilitating demographic and trophic structures with which the ecosystem can respond to disturbances. Adequate patch-size distribution fosters functional connections within the overall distribution and optimizes the balance between genetic diversity and sensitivity to the propagation of disturbances (Collinge, 1996).

In this study, we assume that abiotic gradients in zonal ecosystems are imposed by evolutionary conditions which have led, among other traits, to optimizing their relative patch-size frequency. The concept of structural fragmentation used here refers to the regularity in patch-size frequency distribution of an ecosystem ("patchiness"). Patches, in turn, refer to continuous pieces of an ecosystem showing own structural attributes within their bounds. That regularity reflects the extent of alteration in an interpretable area containing the geographic distribution of the ecosystem in question. The smaller it is, and in particular, the scarcer the large patches are in relation to the frequency of small patches, the harder it will be for the ecosystem to persist over time.

Other attributes in addition to size indicate similar properties of ecosystems. One of the best known is the perimeter to area ratio, which indicates, for example, the extent of interaction with surrounding ecosystems (Forman, 1995). The problem is that shape is much more sensitive than size to cartographic errors and to the spatial scale of the study (Levin, 1992), because of the problem inherent in drawing lines, sometimes arbitrary, which separate adjacent ecosystems. We therefore decided to use size for this study.

The procedure applied here originated in a generalized fact of nature, by which large events are less frequent that small ones. The relationships between event size and frequency can be

described by a power function with an exponent that varies in a small interval for the phenomenon in question. This method was tested for the first time by Korcak to describe the Aegean Islands (Imre and Novotný, 2016; Korcak, 1938), and has since then been applied to other problems, including such diverse fields as the magnitude of earthquakes (Lai, 2000), porosity of soil in forests (Menéndez et al., 2005) or pancreatic islets (Hastings et al., 1992). Hastings et al. (1982) made a pioneering application for finding changes in vitality of forest masses in the Okefenokee Swamp (Georgia-Florida, USA), and since then, the use of hyperbolic exponents has been generalized to measure fragmentation (Peralta and Mather, 2000; Scanlon et al., 2007) or degradation (Kéfi et al., 2007) in vegetation cover. Korcak's exponent is a particularly robust indicator which depends only weakly on the parameters of the system considered (Pascual et al., 2002).

Most of the applications cited above consist of finding exponents for datasets representing variations in a factor, such that the comparison of exponents indicates the factor's effect. This approach is useful for detecting even small changes when the factor is known, but is less applicable to large datasets for which a macroscopic view is necessary regardless of the factors operating on them. In such cases, a segmented fit of the power law over the whole size range is preferable, identifying significant intervals within the data by their different exponents. For example Imre (2005) applied a Korcak analysis to detect intrinsic classification of landmasses into islands, subcontinents and continents, based on their different exponents by area intervals. This was the approach followed in this study.

This study reports the results of applying segmented linear functions to describe the relationship between sizes and frequencies of patches of eight types of zonal forests in mainland Spain. These forests are catalogued by European conservation policy, and their selection followed two criteria: include diverse biogeographic regions, and operate on normalized ecosystems to facilitate method replication.

The objectives were: 1) detect any intervals in which the relationship between patch size and frequency of each forest type varied significantly, identifying explicit patch area thresholds, and 2)

contribute to development of a robust indicator of the extent of fragmentation of forest masses. The study did not attempt to explain the causes of any size intervals, but consolidated them as fact for further exploring its regularity and orienting future detailed studies by type of forest and locality.

The metric applied and data used are explained in the section below. Along with the Spanish forest data are the original data used by Hastings et al. (1982) to describe forest fragmentation using the Korcak exponent. Section 3 uses these to validate and demonstrate the method, and then presents the results for the eight target case studies. Finally, Section 4 examines possible general causes for the patchiness found. Section 5 presents the main conclusions.

2. Materials and Methods

2.1. Segmented regression analysis and fragmentation indicators

2.1.1. The Korcak exponent

The relationship between the size of an event and its frequency is inverse, such that large events are usually less frequent than small ones. The cumulative size distribution can be described by a power law (Hastings and Sugihara, 1993):

$$N(a) = c \cdot a^{-B} \tag{1}$$

where *N(a)* is the number of events larger than size *a*, *c* is a constant, and *B* is the Korcak exponent, which varies within a small interval for the type of event considered. Exponent *B* conveys the frequency change rate as a function of size, and is characteristic of the phenomenon under consideration, so any deviation in it has diagnostic properties. It further enables a fractal interpretation of the phenomenon in question (Jang and Jang, 2012), which has recently been placed in doubt (Imre et al., 2012), but that is not within the scope of this study and will not be developed here.

The simplest way to find the exponent *B* for a dataset is to fit a linear function to the logarithmic transformation of the equation above:

$$\log_e N(a) = \log_e(c) - B \cdot \log_e(a) \tag{2}$$

This is how it was done in the pioneer study cited above on regularity of natural vegetation patches in the Okefenokee Swamp, and other more recent applications (Imre et al., 2011). However, it should be borne in mind that this method for fitting power laws has some inherent problems: the standard error of the predictions once the inverse transformation has been made, the difficulty in rejecting the hyperbolic distribution hypothesis when the data do not follow it, and the fact that a regression line is not a valid distribution. It is therefore preferable to obtain the parameters using maximum likelihood estimators, checking the goodness of fit and comparing the distribution found with other possibilities by means of likelihood ratios (Clauset et al., 2009). Some recent studies on vegetation dynamics which attempted to predict the dependent variable have used this method (Kéfi et al., 2007).

Here we only intended to find exponent *B* by anticipating the need to explore size intervals within the whole range available for each type of forest. That does not incur in the problems mentioned above, and the elementary steps in linear fit enable better control of predictor interval detection. We therefore decided to use this technique.

Size classes should be relevant on the logarithmic numerical scale, not the original one. Some options along this line include photographic f-stops (using the transformation $af_i = a_{min} \cdot \sqrt{2^{n_i}}$ where af_i is the upper limit of the class number n_i and a_{min} is the minimum area considered), regular intervals in the logarithmic transformation of a, or dividing in half successively beginning from the largest size a_{max} . All of them are valid, and we used the last one in this study. Therefore, the size classes were produced by generating a geometric progression with a common ratio of 0.5 and initial value equal to the maximum area rounded to the least multiple of 10. That led to n pairs of values (f_i , a_i), which would represent the number of patches (f) larger than a certain area (a) in pair i, respectively. This series was then entered in the linear fit. The pairs of values were transformed into their natural logarithm, and the resulting slope yielded exponent B.

Some preliminary tests led to the observation that, even when a significant linear regression was produced, the scatterplot seemed, in fact, to contain two different intervals that would benefit from individual fit.

We explored the separation of those intervals for each type of forest using segmented regression with a single and abrupt breakpoint (Muggeo, 2003; Toms and Lesperance, 2003). First, we formed all the possible pairs of intervals in the sequence of *n* size classes, such that each interval had at least three points (f_{ii} , a_i):

 $([a_1 \text{ through } a_3], [a_4 \text{ through } a_n]) \dots ([a_1 \text{ through } a_{n-3}], [a_{n-2} \text{ through } a_n])$ (3) Then we fitted a segmented regression to each pair of intervals, identifying the lower limit of the right-hand one as the breakpoint or threshold area between them. This method fits separate regressions to each interval, but also produces a residual standard error for the combined segmented regression. We selected the pair of intervals with the lowest standard error of all the pairs tested as the optimum, and retained the threshold area as an additional indicator. The decision of whether to accept the segmented regression in two intervals, versus the single regression fit to the total range, was made by an analysis of variance between the two regression models, with the null hypothesis that there was no significant difference between them. We set the threshold of significance for this test at 95% ($p \le 0.05$).

2.1.3. Interpretation criteria

In general, exponent *B* is negative, and its magnitude indicates the rate at which the large patches become scarcer. Theoretical fractal models restrict its absolute values to the $0.5 \le B \le 1$ interval (Hastings and Sugihara, 1993), but in practice, there may be results outside of this range, especially in altered ecosystems.

In turn, and in a wider scope, structural fragmentation indicated by *B* can be interpreted as inversely proportional to ecosystem persistence (Irvine et al., 2016; Wu and Loucks, 1995). This does not necessarily imply poor state or poor trend, but rather vulnerability and lower long-term viability to changing external conditions, for example, rapid climate change (Loarie et al., 2009).

Exponent *B* is thus useful for measuring the extent of structural fragmentation. However, the absence of theoretical references on optimum values suggests that it should be used in relative mode. Typically, the same ecosystem can be compared in two different zones. Alternatively, the whole size range available can be explored for size intervals with non-uniform variation of frequency, as we have done here. In particular, unsustainable use and degradation increase structural fragmentation, affecting mainly larger patches, which can become abnormally scarce and truncate the power law (Kéfi et al., 2007).

2.2. Test data of Okefenokee Swamp

As a preliminary step to its application to the Spanish forest types, we tried out the procedure described with a set of independent published data, to confirm its proper implementation and produce an interpretative base. In particular, we were interested in comparing the application of segmented regressions between the test and the study data to discard any artefacts in the results. These data come from the original article by Hastings et al. (1982) on the use of fractal models for estimating fragmentation of forest ecosystems, and were later revisited (Hastings and Sugihara, 1993), precisely for explaining Korcak's exponent *B* as a measure of fragmentation. That study describes the distribution of cypress (*Taxodium ascendens*) and perennial broadleaf forests (*Magnolia virginia, Persea palustris* and *Gordonia lasianthus*) in the Okefenokee Swamp.

2.3. Experimental data of Spanish forest types

The forest types studied are monospecific formations, dominant or at least typical, in their respective areas of distribution in mainland Spain. The set includes a sample of the functional diversity of Iberian woodlands: broadleaved deciduous (*Fagus sylvatica, Quercus robur*), marcescent (*Q. pyrenaica*) and sclerophyll (*Q. suber*) forests, and Euro-Siberian (*Pinus uncinata*) and Mediterranean (*P. pinea*) conifer forests. Although this sample is arbitrary, it may be considered representative for this fragmentation study for two reasons. First, its zonal nature implies a predictable distribution by climate and soil gradients, from which deviations are

interpretable. Second, the set is distributed over three biogeographic regions in mainland Spain (Atlantic, Alpine and Mediterranean). According to the data described below, the set covers 1,726,578 ha.

Those forest types are identified here by their code in the European Union Habitats Directive (European Commission, 1992), an instrument of continental scope for harmonizing conservation policy in the Member States. Its Annex I catalogues the types of habitats of European interest (HT), preservation of which requires special zones to be designated. The HT are essentially landscape types, in most cases associated with a dominant vegetation. The eight types of forest reported in this study are HT. Their dominant species and HT codes are: *F. sylvatica* (9120, 9130 y 9150), *Q. robur* and *Q. pyrenaica* (9230), *Q. suber* (9330), *P. uncinata* (9430), *P. nigra* ssp. *salzmannii* (9530) and *P. pinea* (a subset of 9540). As more detailed information on these and other HTs may be found in the corresponding manual and database (European Commission, 2013), these data are not repeated here. Figure 1 shows their distribution areas and official denominations.



Figure 1. Distribution in mainland Spain of the forests used in this study. Official codes and designations as in the EU Habitats Directive 92/43/EEC. 9120: Atlantic acidophilous beech forests with *llex* and sometimes also *Taxus* in the shrublayer (*Quercion robori-petraeae* or *llici-Fagenion*); 9130: *Asperulo-Fagetum* beech forests; 9150: Medio-European limestone beech forests of the *Cephalanthero-Fagion*; 9230: Galicio-Portuguese oak woods with *Quercus robur* and *Quercus pyrenaica*; 9330: *Quercus suber* forests; 9430: Subalpine and montane *Pinus uncinata* forests; 9530: (Sub-) Mediterranean pine forests with endemic black pines; 9540:

Mediterranean pine forests with endemic Mesogean pines (only *P. pinea* forests used in this study). Biogeographic regions: Atlantic (blue), Alpine (purple), Mediterranean (yellow). Distribution patches shown oversized for clarity. Coordinate reference system: EPSG 25830.

The geographic distributions and patch size attributes observed in these forest types were found from the dataset described in Sanchez de Dios et al. (n.d.). Both the data and R-scripts implementing the methods used in this study are available at Sanchez de Dios et al. (2020). These data use tesserae from the Spanish Forest Map 1:50,000 (MAPAMA, 2013) as elementary spatial units, each tessera differing in at least one relevant structural attribute from those neighbouring it. These attributes, such as physiognomy, dynamics, species composition and forest structure differentiate one tessera from its neighbours even if they are contiguous and belong to the same forest type.

The patches in this study are tesserae such as those, and are not lumped together in a homogeneous spatial distribution at a lower level of detail. Hence, the method described in Section 2.1 operated on the set of patches of each HT. The patch sizes were extracted from the field containing the area in the original vector coverages.

3. Results

The numerical results of fitting regressions to the complete data range for each type of forest are summarized in Table 1. These correspond to the null hypothesis of not partitioning the scatterplots, and serve as a reference for evaluating the segmented regressions shown in Table 2.

Table 1. Linear regressions fitted to log-log transformations of cumulative patch frequency as a function of patch size. 'Overall fit' refers to the regression fitted to the whole data range and reports the slope, or Korcak exponent (*B*), coefficient of determination (r^2), degrees of freedom (*df*) and probability (*p*). 'Interval separation' refers to the division of the patch size range in two optimum intervals and reports the lower limit of the right-hand interval (*Threshold area*), its position in the corresponding patch size distribution (*Percentile*), the ANOVA made to decide whether the segmented regression is different from the overall fit (*F*) and its probability (*p*). Rows correspond to the forest types considered in this study: published test data of Okefenokee cypress (*Taxodium ascendens*), and eight European habitat types identified by their code.

Forest type	Overall fit				Interval separation					
	В	r ²	df	p	Threshold	Percentile	F	р		
					area (ha)					
Okefenokee	0.625	0.97	11	4E-10	91.46	50%	21.84	4E-04		
cypress										
9120	0.357	0.52	16	7E-04	101.09	76%	182.45	9E-11		
9130	0.339	0.64	17	4E-05	39.19	66%	641.52	3E-15		
9150	0.374	0.65	19	9E-06	26.67	80%	176.79	4E-12		
9230	0.273	0.46	25	1E-04	55.29	77%	582.35	2E-20		
9330	0.489	0.67	17	2E-05	67.87	76%	320.83	5E-13		
9430	0.139	0.45	31	2E-05	26.61	73%	588.17	3E-24		
9530	0.397	0.54	22	4E-05	43.64	55%	201.30	6E-14		
9540	0.299	0.50	19	3E-04	40.42	66%	293.07	7E-14		

Table 2. Segmented linear regressions fitted to left- and right-hand optimum intervals around a threshold area, of log-log transformations of cumulative patch frequency as a function of patch size. The slope, or Korcak exponent (*B*), coefficient of determination (r^2), degrees of freedom (*df*) and probability (*p*) are reported for each interval. Rows correspond to the forest types considered in this study: published test data of Okefeokee cypress (*Taxodium ascendens*), and eight European habitat types identified by their code

Forest type	Left-ha			and interval			Right-hand interval			
	В	r²	df	p		В	r ²	df	p	
Okefenokee	0.433	0.99	3	3E-04		0.781	0.99	6	5E-07	
cypress										
9120	0.071	0.57	12	2E-03		2.590	0.96	2	2E-02	
9130	0.043	0.56	9	8E-03		1.100	1.00	6	3E-08	
9150	0.094	0.64	12	6E-04		1.436	0.97	5	4E-05	
9230	0.028	0.34	19	5E-03		2.129	0.98	4	2E-04	
9330	0.075	0.56	11	3E-03		1.741	0.99	4	7E-05	
9430	0.028	0.56	25	8E-06		1.376	0.98	4	2E-04	
9530	0.017	0.36	14	1E-02		1.805	0.94	6	7E-05	
9540	0.030	0.40	13	1E-02		1.763	0.97	4	4E-04	

3.1. Fragmentation in Okefenokee Swamp

The following explanation can help understand the application of segmented regressions to explore forest fragmentation, and to assess the validity of the results for Spanish forest habitats. The dataset in the Northeast quadrant of the Okefenokee Swamp consists of 13 pairs of values (f_i , a_i), from (3, 2560) to (43, 40), where patch size a is expressed in acres. Patches smaller than 40 acres were excluded from the original study.



Figure 2. Linear regression fitted to the log-log transformation of cumulative patch frequency (*N*(*a*), number of patches larger than size *a*) as a function of size (*a*). The slope (*B*) is the Korcak exponent of the corresponding power law. Regression equation: $\log N(a) = 6.268 - 0.625 \cdot \log (a) [df = 11, p = 4.37E-10]$. Data from cypress of the Okefenokee Swamp in the NE quadrant, taken from Hastings et al. (1982).

Figure 2 shows the linear regression fitted to the double logarithmic transformation. The B=0.625 computed here practically coincides with the one published (B=624) and the difference can be attributed to rounding off. Those authors interpreted this value as relatively less fragmented than the B =0.799 found in the SE quadrant of the zone (not computed here), suggesting less persistence of the cypress forest in the latter. Examination of the field conditions demonstrated that a dike separated the two quadrants and retained the NE-to-SE flow of water, which increased pooling and could be interpreted as a more favorable habitat for cypresses in the NE. Those authors also found values for broadleaved forests in the same zone (B =0.453 and B =0.491 for NE and SE respectively). Their explanation for lesser fragmentation, and therefore, greater persistence, was that these forests belong to a more mature state in the ecological succession than the cypress.

It may be observed in Figure 2 that the regression residuals are not distributed entirely at random, as positive residuals concentrate in the middle of the line. This occurs approximately between log-transformed values log (a) = 5 and log (a) = 6, suggesting that the fit might be better if it were done separately for each of the two sides of this zone. Applying the segmented regression procedure described in Section 2.1.2, we found that the standard error of residuals reached a minimum of 0.066 when the two intervals separated around the log (a) = 5.4 threshold, hence a equals to 226 acres (Figure 3).



Figure 3. Variation of residual standard error of segmented linear regressions fitted to all possible pairs of patch size intervals within the data range. Abscissas represent the threshold area separating each pair, which corresponds to the lower limit of the right-hand interval. Optimum separation was found for log (a) = 5.4. Data from cypress in the NE quadrant of the Okefenokee Swamp, taken from Hastings et al. (1982).

The combined model using the regressions fitted to both sides of that threshold area was significantly different from the overall linear fit (Table 2, Interval separation), and we decided to use it because of its improved description of the scatterplot (Figure 4). Both the left-hand and right-hand regression increased their coefficient of determination and their significance with respect to the overall regression in Figure 2, and after that, visual inspection of the residuals did not detect any obvious pattern.



Figure 4. Segmented regressions fitted respectively to left- and right-hand intervals of patch size around a threshold area of 226 acres. Compare with Figure 2. Data from cypress of the Okefenokee Swamp in the NE quadrant, taken from Hastings et al. (1982).

Figure 4 shows that the decrease in frequency of large patches relative to that of small ones is not monotonous. It is relatively slow while the patches are less than 226 acres (91.46 ha), but beyond it, the frequency of patches larger than this threshold decreases more rapidly with their size.

3.2. Fragmentation of Spanish forest types

The regressions fitted to the overall data range were significant in all cases, enabling initial reference values for *B* to be found (Table 1, Overall fit). But also in all cases, the segmented regression led to identification of two optimum intervals whose combined regressions improved significantly the overall linear fit (Table 1, Interval separation). Table 2 and Figure 5 show the fit by intervals for each HT.



Figure 5. Segmented linear regressions fitted to log-log transformations of cumulative patch frequency N(a) as a function of patch size a, for eight forest types in Spain identified by their HT code. A vertical line

indicates separation in two optimum intervals. HT codes and dominant forest species: 9120: *Fagus sylvatica*; 9130: *F. sylvatica*; 9150: *F. sylvatica*; 9230: *Quercus pyrenaica*; 9330: *Q. suber*; 9430: *Pinus uncinata*; 9530: *P. nigra*; 9540: *P. pinea*.

Figure 5 shows a general pattern. The variation of cumulative patch-size frequencies is abruptly truncated near the threshold, so that the slope in the left-hand interval is very weak, while on the right it is very strong. This explains the differences between the coefficients of determination r^2 fitted to the overall data range (Table 1), and those fitted to the corresponding intervals (Figure 4, Table 2). In general, the coefficient of determination refers to the proportion of variance in the dependent variable explained by the predictor. If the slope is small, relatively large changes in the predictor cause relatively small changes in the dependent variable. This is why, in most cases here (except HT 9120 and 9430), r^2 decreases in the left-hand interval with respect to the overall fit, while in all cases r^2 for the right-hand interval strongly increases with respect to the overall fit. In parallel, the 95% confidence intervals are wider in the latter.

The threshold areas determining the separation between intervals (Table 2, Interval separation) are generally very small, and only in HT 9120 exceeds 1 km² by very little. Furthermore, all of them (except HT 9530, discussed below) surpass the median of the corresponding size distribution (Figure A1), which means that large patch sizes are not only rare relative to small ones (as illustrated by *B* values), but also in absolute terms.

4. Discussion

All of the results were found by the same method, and we believe its application to Spanish forest habitats is validated by the case of the Okefenokee Swamp. Therefore, these results may be considered findings subject to interpretation.

Examination of the results for the Okefenokee Swamp helps establish a line of interpretation for Spanish forest habitats. That zone has been a protected area since 1974, but from 1909 to 1927 it was an important center of logging concentrating precisely on cypress, a situation which lasted until 1942 (Malcolm et al., 1994). If it is admitted that forests could still be recovering at the time of the publication of reference (1982), the threshold of 226 acres (91.5 ha) could help distinguish between young, and relatively small patches undergoing expansion, and old patches which have survived that exploitation. If it is assumed, on the contrary, that the forest was steady in 1982, that threshold could identify a critical size above which cypress patches become unstable and begin to fragment. One way or the other, the message is that, probably, the threshold area detected identifies a trait in the landscape memory.

In the Spanish HTs, at simple visual inspection, the significant statistical results reinforce all the graphs, which show much more contrast between intervals than the cypress forests of the Okefenokee Swamp. This pattern indicated by their respective *B* exponents, along with the threshold areas separating the two intervals, are the main elements for describing HT structural fragmentation.

The effect of fragmentation on forest species needs to be found within a context (Aparicio et al., 2012). The description of this fragmentation would therefore require assembling a scenario for each HT in which the history of use and abandonment in recent centuries, documented climate variation, and their current dendrometry according to the National Forest Inventory (Alberdi et al., 2016), could be considered. Such a task is outside the scope of this study. Nevertheless, just as in the example of the Okefenokee Swamp, it might be possible to perform such explorations from a common perspective in view of certain convergences in the results. Below we propose some guiding hypotheses, first, from an abiotic perspective, then, incorporating historical elements We analyzed three types of beech forests: oligotrophous on acid soils (9120); eutrophous on non-acid soils, either typical Euro-Siberian or Medio-European (9130) and sub-Mediterranean (9150). In Spain, the HT 9120 beech forests occupy mountainous zones with generally Paleozoic substrates. They are territories with scant population or agricultural interest in which the forests are rather well conserved. It is logical in this case for the patch-size threshold (101 ha) to be rather larger than in the rest of the forests included, because of their relative environmental homogeneity, lower agricultural and general human pressure, and intense dynamics, with strong

capacity for recovery after charcoal production, logging and extensive livestock grazing practiced in the past were abandoned (García et al., 2005).

The HT 9130 beech forests are rather more uncommon and occupy sometimes extra-zonal locations scattered around the Cantabrian and Pyrenees Mountains. These Iberian beech forests are more like the Medio-European ones, and usually appear in mosaics with other types of beech forests, occupying the most Atlantic and eutrophic localities, often on non-acid soils. They therefore tend to appear naturally in small areas with deep soil and gentle slope or unique substrates. Regardless of the human influence which they may have been subjected to, it seems logical that they would be naturally more fragmented (threshold 39 ha), because they are microclimatically related to the less Mediterranean enclaves in Iberian mountains.

The beech forests in HT 9150 are more unique, as they represent the adaptation of these Euro-Siberian forests to the Mediterranean environment. They occupy the most southerly locations in the beech distribution areas. These forests, which include sub-Mediterranean species, such as boxwood (*Buxus sempervirens*), are found on non-acid substrates and respond only to favorable geomorphic or microclimatic conditions. These facts seem to be mainly responsible for the strong fragmentation (threshold 27 ha) found in the analysis.

A similar situation of strong natural fragmentation (threshold 27 ha) occurs in the subalpine *Pinus uncinata* forests (HT 9430). In this case, natural environmental stress at the timberline, with windy areas and unstable landforms are the main drivers.

The results of fragmentation for the Mediterranean forests show some intermediate thresholds similar to each other. Thresholds of broadleaved HTs (*Quercus pyrenaica*, 9230: 68 ha and *Q. suber*, 9330: 55 ha) are slightly higher than those of conifers (*P. nigra*, 9530: 44 ha and *P. pinea*, 9540: 40 ha). This fragmentation may primarily be related to historic use of the territory. However, the higher *Quercus* threshold areas could also be associated with their being a successful Holocene species, while the pine forests have remained relegated to singular substrates, where they persist with a certain relict nature under the present climate, favorable to broadleaved trees. The capacity

of pine forests for acting as substitution stages of broadleaf Quercus forests could act in the opposite direction, facilitating the continuity of masses in historically altered territories. From a historical perspective, Spain has gone through several processes that could be considered Global Change events, and which have had documented impacts on forest cover. Among those affecting the forest types studied are: the expulsion in the 15th century of the Islamic population, followed by occupation of their lands by immigrants, who in general came from territories located in the North; Spanish confiscation of lands belonging to religious orders from the 18th to 20th centuries by expropriation and public auctioning to amortize the national public debt; the rural demographic maximum reached at the end of the 19th century, right before the industrial revolution; or the entry of Spain in the European Economic Community in 1985, which caused a growing trend toward agricultural intensification and progressive abandonment of large land areas. All those processes have an age commensurate with the demographic dynamics of a forest, and each of them could lead to coexistence with earlier or later forest patches. For example, the history of forest patches has been recognized as the most important variable for explaining the current pine and oak forest cover in Catalonia (NE Spain) (Guirado et al., 2008). In addition, all the forests included here are natural resources exploited by the local human population. This pressure could be reflected in the relative rarity of patches over a certain threshold area.

Alternatively, many forest ecosystems are, in fact, marginal areas originating from abandonment of lands, and therefore, may not have any functional relationship with land use other than buffering abandonment and recovery (Costa et al., 2011). This could explain why the frequency of patches would be almost independent of their size up to a certain threshold, and hence the *B* value near zero in the left-hand interval.

At the same time as above, some instability could be inherent after a certain size threshold. For example, a rapid climate change (Loarie et al., 2009) could cause growing uncoupling between the spatial structure of a certain ecosystem and its abiotic limiting factors. In zonal forests, such decoupling would be more important on the edges of its areas of distribution, where they would in fact become extra-zonal (del Barrio et al., 2006). This may be relevant in forests such as those dealt with here, especially in a territory extensively subjected to land uses and transactions, in which their current distribution has been forced since historical times.

5. Conclusions

This study quantified structural fragmentation (patchiness) of eight types of forests in mainland Spain designated as habitats of European interest by the main pan-European conservation policy (Directive 92/43 EEC). These forests occupy a total of 1,726,578 ha in the three biogeographic regions present in the study area, and are dominated by zonal species: *Fagus sylvatica* (9120, 9130 y 9150), *Quercus robur* and *Q. pyrenaica* (9230), *Q. suber* (9330), *Pinus uncinata* (9430), *P. nigra* ssp. *salzmannii* (9530) and *P. pinea* (subset of 9540). The results may therefore be considered representative of other natural forests, and in any case, their incidence is important on a nationwide scale.

The hypothesis was that frequency of patches decreases gradually with their size, following a general trend observed in many natural phenomena. To test it, we fit power laws to the data observed in each type of forest, and used the Korcak exponent *B* as an indicator of the rate of frequency variation with patch size.

That hypothesis was rejected in all cases. In general, the function that describes the change in frequency by size was truncated into two intervals around relatively small threshold areas of 27 to 101 ha. The rate of variation on the left was always very low (B = 0.017 - 0.094), although significant. When the corresponding threshold area is reached, the variation rate increases abruptly (B = 1.100 - 2.590). The conclusion is that the eight types of forest analyzed are made up of relatively small patches whose frequency decreases very slowly with their size until a certain threshold, after which relatively large patches rapidly become very scarce.

The numerical experiment was applied in parallel to the original published data which described the use of the Korcak exponent in forests (Hastings et al., 1982). Those data are for cypress (*Taxodium ascendens*) forests in the Okefenokee Swamp. The *B* exponent values found for the overall fit using the whole range of patch sizes were equivalent to those published (*B*=0.625), thus confirming the numerical procedure. The additional exploration of intervals using segmented regressions also provided a breakpoint of 91.5 ha, although the exponents on each side (*B*=0.433 and *B*=0.781, respectively) were less disparate, which suggests greater monotony in the loss of relatively large patches.

This study only intended to detect patterns in patch-size distribution, not explain them, which would have required incorporating many details on the ecological history in each case. However, the general hypotheses that explain the patterns observed (and also in Okefenokee) agree on the existence of patches of different ages due to human pressure events on forests. This pressure on Spanish forests is currently maintained, and that would explain the strongly contrasting patterns around the threshold areas.

We believe that these results will help diagnose the fragmentation of the eight types of forests included. The threshold area is an explicit indicator for spatial patch selection, enabling the land uses that affect them to be known, and exponent *B* is an accepted indicator of the intensity of that pressure. The segmented regression technique is simple and works on easily available data, so it has an important potential as a tool for the assessment and surveillance of fragmentation of other forest masses.

Appendix



Fig. A1. Patch size statistical distribution of eight forest types in Spain identified by their HT code. Boxes indicate median (notch) and the 25th and 75th percentiles. Whiskers show the non-outlier data range, defined as 1.5 times the interquartile range. Area thresholds found in this study are indicated by grey circles.

Acknowledgements

Rafael Hidalgo, of the Spanish Ministry for the Ecological Transition and the Demographic Challenge, provided with useful end-user insights on the applicability of the approach developed in this study. Elena Bermejo, of Tragsatec, reviewed and harmonized the results to fit in the broader task to develop a surveillance system for the Spanish habitat types. Juan Carlos Simón, of Tragsatec, coordinated the overall project to ensure consistent interaction between its components. We would also like to thank the two anonymous reviewers who reviewed the original manuscript for their helpful assessment.

References

Please note that reference Sanchez de Dios et al. (n.d.) below corresponds to the manuscript for Data in Brief describing the data used in this study, which was submitted along with this manuscript.

- Alberdi, I., Sandoval, V., Condes, S., Cañellas, I., Vallejo, R., 2016. The Spanish National Forest Inventory, a tool for the knowledge, management and conservation of forest ecosystems. Ecosistemas 25, 88–97. https://doi.org/10.7818/ECOS.2016.25-3.10
- Aparicio, A., Hampe, A., Fernández-Carrillo, L., Albaladejo, R.G., 2012. Fragmentation and comparative genetic structure of four mediterranean woody species: Complex interactions between life history traits and the landscape context. Divers. Distrib. 18, 226–235. https://doi.org/10.1111/j.1472-4642.2011.00823.x
- Breckle, S.-W., 2002. Walter's Vegetation of the Earth, 4th ed. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-56164-1
- Clauset, A., Rohilla Shalizi, C., J Newman, M.E., 2009. Power-Law Distributions in Empirical Data. SIAM Rev. 51, 661–703. https://doi.org/10.1214/13-AOAS710
- Collinge, S.K., 1996. Ecological consequences of habitat fragmentation: Implications for landscape architecture and planning. Landsc. Urban Plan. 36, 59–77. https://doi.org/10.1016/S0169-2046(96)00341-6
- Costa, A., Madeira, M., Lima Santos, J., Oliveira, Â., 2011. Change and dynamics in Mediterranean evergreen oak woodlands landscapes of Southwestern Iberian Peninsula. Landsc. Urban Plan. 102, 164–176. https://doi.org/10.1016/j.landurbplan.2011.04.002

Deacon, R.T., 1999. Deforestation and ownership: Evidence from historical accounts and contemporary data.

Land Econ. 75, 341–359. https://doi.org/10.2307/3147182

- del Barrio, G., Harrison, P.A., Berry, P.M., Butt, N., Sanjuán, M.E., Pearson, R.G., Dawson, T., 2006.
 Integrating multiple modelling approaches to predict the potential impacts of climate change on species' distributions in contrasting regions: comparison and implications for policy. Environ. Sci. Policy 9, 129–147. https://doi.org/10.1016/j.envsci.2005.11.005
- del Barrio, G., Puigdefabregas, J., Sanjuán, M.E., Stellmes, M., Ruiz, A., 2010. Assessment and monitoring of land condition in the Iberian Peninsula, 1989-2000. Remote Sens. Environ. 114, 1817–1832. https://doi.org/10.1016/j.rse.2010.03.009
- Estreguil, C., Caudullo, G., de Rigo, D., San-Miguel-Ayanz, J., 2013. Forest Landscape in Europe: Pattern, Fragmentation and Connectivity. JRC 77295. Publications Office of the European Union, Luxembourg. https://doi.org/10.2788/77842
- European Commission, 2013. Interpretation Manual of European Union Habitats, version EUR 28, April. European Commission, DG-ENV, Brussels.
- European Commission, 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora [WWW Document]. Off. J. Eur. Union. URL https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31992L0043&from=EN (accessed 5.7.20).
- Forman, R.T., 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, Cambridge.
- García, D., Quevedo, M., Obeso, J.R., Abajo, A., 2005. Fragmentation patterns and protection of montane forest in the Cantabrian range (NW Spain). For. Ecol. Manage. 208, 29–43. https://doi.org/10.1016/j.foreco.2004.10.071

- Guirado, M., Pino, J., Rodà, F., Basnou, C., 2008. Quercus and Pinus cover are determined by landscape structure and dynamics in peri-urban Mediterranean forest patches. Plant Ecol. 194, 109–119. https://doi.org/10.1007/s11258-007-9278-9
- Hastings, H.M., Pekelney, R., Monticciolo, R., Vun Kannon, D., Del Monte, D., 1982. Time scales, persistence and patchiness. BioSystems 15. https://doi.org/10.1016/0303-2647(82)90043-0
- Hastings, H.M., Schneider, B.S., Schreiber, M.A., Gorray, K., Maytal, G., Maimon, J., 1992. Statistical geometry of pancreatic islets. Proc Biol Sci 250, 257–261. https://doi.org/10.1098/rspb.1992.0157
- Hastings, H.M., Sugihara, G., 1993. Fractals: a user's guide for the natural sciences. Oxford University Press, Oxford.
- Hill, J., Stellmes, M., Udelhoven, T., Röder, A., Sommer, S., 2008. Mediterranean desertification and land degradation. Mapping related land use change syndromes based on satellite observations. Glob.
 Planet. Change 64, 146–157. https://doi.org/10.1016/j.gloplacha.2008.10.005
- Imre, A.R., 2015. Description of the area distribution of landmasses by Korcak exponent—the importance of the Arabic and Indian subcontinents in proper classification. Arab. J. Geosci. 8, 3615–3619. https://doi.org/10.1007/s12517-014-1481-0
- Imre, A.R., Cseh, D., Neteler, M., Rocchini, D., 2011. Korcak dimension as a novel indicator of landscape fragmentation and re-forestation. Ecol. Indic. https://doi.org/10.1016/j.ecolind.2010.12.013
- Imre, A.R., Novotný, J., 2016. Fractals and the Korcak-law: a history and a correction. Eur. Phys. J. H. https://doi.org/10.1140/epjh/e2016-60039-8
- Imre, A.R., Novotný, J., Rocchini, D., 2012. The Korcak-exponent: A non-fractal descriptor for landscape patchiness. Ecol. Complex. https://doi.org/10.1016/j.ecocom.2012.10.001
- Irvine, M.A., Jackson, E.L., Kenyon, E.J., Cook, K.J., Keeling, M.J., Bull, J.C., 2016. Fractal measures of spatial

pattern as a heuristic for return rate in vegetative systems. R. Soc. Open Sci. 3. https://doi.org/10.1098/rsos.150519

- Jang, J., Jang, Y.H., 2012. Spatial distributions of islands in fractal surfaces and natural surfaces. Chaos, Solitons and Fractals 45, 1453–1459. https://doi.org/10.1016/j.chaos.2012.09.003
- Kéfi, S., Rietkerk, M., Alados, C.L., Pueyo, Y., Papanastasis, V.P., ElAich, A., de Ruiter, P.C., Kefi, S., Rietkerk, M., Alados, C.L., Pueyo, Y., Papanastasis, V.P., ElAich, A., de Ruiter, P.C., 2007. Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. Nature 449, 213–7. https://doi.org/10.1038/nature06111
- Klein, J., 1920. The Mesta. A study in Spanish economic history, 1273–1836. Cambridge, Harvard University Press, London.

Korcak, J., 1938. Deux types fondamentaux de distribution statistique. Bull. l'Institut Int. Stat. 3, 295–299.

- Lai, D., 2000. Statistical and fractal analysis of Chinese earthquake time series. Fractals 8, 123–127. https://doi.org/10.1142/S0218348X00000135
- Levin, S.A., 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture Author(s): Simon A. Levin Source: Ecology. https://doi.org/doi:10.2307/1941447
- Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B., Ackerly, D.D., 2009. The velocity of climate change. Nature 462, 1052–1055. https://doi.org/10.1038/nature08649
- Malcolm, R.L., McKnight, D.M., Averett, R.C., 1994. History and description of the Okefenokee Swamp -Origin of the Suwannee River, in: Averett, R.C., Leenheer, J.A., McKnight, L.M., Thorn, K.A. (Eds.), Humic Substances in the Suwannee River, Georgia: Interactions, Properties, and Proposed Structures. US Geological Survey Water -Supply Paper 2373. United States Government Printing Office, Denver,

pp. 1–12.

- MAPAMA, 2013. Mapa Forestal de España. Escala 1:50.000 [WWW Document]. Minist. Agric. Aliment. y Medio Ambient. URL https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponibl
- Menéndez, I., Caniego, J., Gallardo, J.F., Olechko, K., 2005. Use of fractal scaling to discriminate between and macro- and meso-pore sizes in forest soils. Ecol. Modell. 182, 323–335. https://doi.org/10.1016/j.ecolmodel.2004.04.009

e/mfe50.aspx (accessed 5.5.20).

- Merino Navarro, J.P., 1981. La armada española en el siglo XVIII. Fundación Universitaria Española, Madrid, Spain.
- Muggeo, V.M.R., 2003. Estimating regression models with unknown break-points. Stat. Med. 22, 3055–3071. https://doi.org/10.1002/sim.1545
- Pascual, M., Roy, M., Guichard, F., Flierl, G., 2002. Cluster size distributions: Signatures of self-organization in spatial ecologies. Philos. Trans. R. Soc. B Biol. Sci. 357, 657–666.
 https://doi.org/10.1098/rstb.2001.0983
- Peralta, P., Mather, P., 2000. An analysis of deforestation patterns in the extractive reserves of Acre, Amazonia from satellite imagery: A landscape ecological approach. Int. J. Remote Sens. 21, 2555–2570. https://doi.org/10.1080/01431160050110179
- Sanchez de Dios, R., Sainz, H., Sanjuan, M.E., Martinez-Valderrama, J., Ruiz, A., del Barrio, G., 2020. Distribution of eight forest habitats of European Interest in mainland Spain, and R code for Korcak fragmentation analysis [WWW Document]. Mendeley Data, v2. https://doi.org/10.17632/bm59f4nddy.2

Sánchez de Dios, R., Sanjuán, M.E., Sainz-Ollero, H., Ruiz, A., Martínez-Valderrama, J., del Barrio, G., n.d. A

dataset with complete geographic distributions of eight zonal monospecific forest types in mainland Spain. Data Br.

- Scanlon, T.M., Caylor, K.K., Levin, S. a, Rodriguez-Iturbe, I., 2007. Positive feedbacks promote power-law clustering of Kalahari vegetation. Nature 449, 209–212. https://doi.org/10.1038/nature06060
- Stevenson, A.C., Harrison, R.J., 1992. Ancient Forests in Spain: A Model for Land-use and Dry Forest Management in South-west Spain from 4000 BC to 1900 AD. Proc. Prehist. Soc. 58, 227–247. https://doi.org/DOI: 10.1017/S0079497X00004175
- Thuiller, W., Vayreda, J., Pino, J., Sabate, S., Lavorel, S., Gracia, C., 2003. Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). Glob. Ecol. Biogeogr. 12, 313–325. https://doi.org/10.1046/j.1466-822X.2003.00033.x
- Toms, J.D., Lesperance, M.L., 2003. Piecewise regression: A tool for identifying ecological thresholds. Ecology 84, 2034–2041. https://doi.org/10.1890/02-0472
- Valbuena-Carabaña, M., de Heredia, U.L., Fuentes-Utrilla, P., González-Doncel, I., Gil, L., 2010. Historical and recent changes in the Spanish forests: A socio-economic process. Rev. Palaeobot. Palynol. 162, 492–506. https://doi.org/https://doi.org/10.1016/j.revpalbo.2009.11.003
- Wu, J., Loucks, O.L., 1995. From balance of nature to hierarchical patch dynamics: A paradigm shift in ecology. Q. Rev. Biol. 70, 439–466. https://doi.org/10.1086/419172