Journal of Biogeography (J. Biogeogr.) (2016) 43, 2502–2512

METHODOLOGICAL APPLICATIONS

Assessing relative variable importance across different spatial scales: a two-dimensional wavelet analysis

Gudrun Carl[1]*, Daniel Doktor[2], Oliver Schweiger[1] and Ingolf Kühn[1,3,4]

[1]Department of Community Ecology, Helmholtz Centre for Environmental Research - UFZ, 06120 Halle, Germany, [2]Department of Computational Landscape Ecology, Helmholtz Centre for Environmental Research - UFZ, 04318 Leipzig, Germany, [3]Geobotany & Botanical Garden, Martin Luther University Halle-Wittenberg, 06108 Halle, Germany, [4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany

*Correspondence: Gudrun Carl, Department of Community Ecology, Helmholtz Centre for Environmental Research - UFZ, Theodor-Lieser-Str. 4, 06120 Halle, Germany.
E-mail: gudrun.carl@ufz.de

ABSTRACT

Aim Assessing the relationship between a spatial process and environmental variables as a function of spatial scale is a challenging problem. Therefore, there is a need for a valid and reliable tool to examine and evaluate scale dependencies in biogeography, macroecology and other earth sciences.

Location Central Europe (latitude 43.99°–54.22° N, longitude 4.79°–15.02° E).

Methods We present a method for applying two-dimensional wavelet analysis to a generalized linear model. This scale-specific regression is combined with a multimodel inference approach evaluating the relative importance of several environmental variables across different spatial scales. We apply this method to data of climate, topographic and land cover variables to explain variation in annual greening of vegetation (i.e. phenology) in Central Europe.

Results Land use is more important to explain the variation in greening than climate at smaller resolution while climate is more important at larger resolution with a shift at *c.* 1000 km^2.

Main conclusions To the best of our knowledge, this is the first study analysing the scale dependency of an ecosystem process, clearly distinguishing between the different components of scale, namely grain, focus and extent. The obtained results demonstrate that our newly proposed method is particularly suitable for studying scale dependencies of various spatial processes on environmental drivers keeping grain and extent constant and changing focus (i.e. resolution).

Keywords
discrete wavelet transform, generalized linear model, multimodel inference, remote-sensing signal, spatial scales, vegetation period

INTRODUCTION

The importance of spatial patterns and spatial scales has often been cited as a key issue in biogeography, macroecology, and beyond that, all earth sciences (Levin, 1992; Dale, 1999; Wu & Hobbs, 2002; Fortin & Dale, 2005; Schröder & Seppelt, 2006). Data collection for biogeographical and environmental data is frequently carried out with reference to a gridded map of a specific resolution. A statistical model based on these data will provide statistical inferences at this specific spatial scale. Because different (e.g. biological) processes act at different scales, multiple relationships are scale-specific as well (Pearson & Dawson, 2003; Pearson *et al.*, 2004; Guisan & Thuiller, 2005; Keil *et al.*, 2012). Hence, the selection of scale for data collection and inference is crucial in statistical modelling. In general, however, different scales will be relevant in such multiple relationships and some of them will be different from the pre-specified collection unit (i.e. focus and extent of analysis *sensu* Scheiner *et al.* (2000); see below). As a consequence, conclusions based on regressions of these data, that is, its parameter estimates, hypotheses tests and *P*-values, may be misleading and can result in incorrect inferences. At least, this is the case if we ignore that these conclusions are restricted to a particular scale and disregard the complexity and multi-scaled structure of the problem. Therefore, there is a need for a valid and reliable

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12781

Journal of Biogeography

tool to examine and evaluate scale dependencies (Wu & Hobbs, 2002; Borcard *et al.*, 2004). The principal coordinates of neighbour matrices (PCNM) analysis (e.g. Borcard *et al.*, 2004) as well as its generalization, the Moran's eigenvector maps (MEM) approach (Dray *et al.*, 2006), provide a spectral decomposition of the spatial relationships. Both methods have in common that the eigenvectors used for spatial filtering purposes are sinusoidal waves of increasing frequency in case of regular sampling. Hence, the methods are basically statistical versions of Fourier analysis. Accordingly, $n - 1$ eigenvectors are generally needed in order to decompose a centred series of n observations and to completely capture its variability. In the general case of a large number n, however, the question arises as to how to identify the main scales of spatial variation and to select appropriate subsets of eigenvectors (Jombart *et al.*, 2009).

To overcome these problems, recent publications recommend the use of wavelet transforms as a tool for scale-specific regression analysis (Dong *et al.*, 2008; Lookingbill *et al.*, 2011), which is expected to be useful to identify scale-specific relationships between predictor and response variables, and thus to provide deeper insights in multiple scale dependencies. Wavelet analysis is an extension and refinement of Fourier analysis (Percival & Walden, 2000; Stark, 2005). Like Fourier analysis, it can be used to detect scale-specific (or frequency-specific) features of a signal. However, different from Fourier analysis, it is able to make any necessary local adjustments, providing different coefficients for different positions (or times). In fact, a Fourier analysis provides frequency or scale components whose amplitudes are the same at all positions (or times), whereas a wavelet analysis is able to provide additional information about which component is present at which spatial (or time) interval. This is because Fourier analyses are based on sinusoidal waves, whereas wavelet analyses use so-called wavelets, that is, small waves visualizable as localized oscillations (Daubechies, 1992; Torrence & Compo, 1998; Cazelles *et al.*, 2008). Due to their much better local adaptation, such a wavelet analysis requires only a few frequency components to completely capture the variability of a signal.

Different approaches have been proposed for applying wavelet transforms to multiple linear regressions. On the one hand, Keitt & Urban (2005) developed a scale-dependent regression and found evidence for scale-specific relationships and inferences regarding predictor variables and the response variable. However, their approach is limited in its application to one-dimensional data analyses and response vectors of Gaussian distribution. On the other hand, wavelets have been used to remove spatial autocorrelation in multiple regressions affected by correlated errors (Carl & Kühn, 2008, 2010). Our method (Carl & Kühn, 2010) allows regular two-dimensional (2-D) sampling grids as well as different distributions (e.g. binomial or Poisson). In both cases, it has proved fruitful to insert wavelet transforms into the regression analysis of spatial data. Most recently, Ma & Zhang (2015) as well as Ye *et al.* (2015) followed the idea and performed a regression analysis using 2-D wavelet transforms to describe scale-specific patterns. Their results have demonstrated that such regressions are appropriate tools for exploring spatial variations at multiple spatial scales. However, the approaches described by Ma & Zhang (2015) and Ye *et al.* (2015) are only applicable to Gaussian response models and therefore exclude, for example, logistic regressions. Moreover, the fact alone that different slopes at different scales can be discovered by means of wavelets is not sufficient. Ma & Zhang (2015) ranked explanatory variables at a given spatial scale in terms of the magnitude of the standardized coefficients. Ranking without any rank order weights is, however, a rather poor method. Instead, the regression analysis should be followed by any assessment, that is, the calculation of an appropriate index for the strength of evidence. Also, Ma & Zhang (2015) scaled species richness. Since species richness does not scale additively and cannot be averaged, wavelets are inappropriate to scale such data. Therefore, the major objective of this study is to combine the advantages of the previous methods and to develop a 2-D wavelet regression applicable to various distributions. Moreover, our wavelet multiresolution regression will lead to scale-dependent inferences by means of rank order weights.

Analysing scale dependency, one has to be very clear about the four different components of scale (Scheiner *et al.*, 2000): (1) sample unit, (2) grain, (3) focus and (4) extent. Sample unit refers to the spatial dimension of the collection unit (e.g. sampling plot). Grain is the smallest unit to which all sample units are standardized for a specific analysis (i.e. finest resolution). The units of grain can then be aggregated to coarser units of analysis, that is, focus (i.e. coarser resolution). Extent in this context is the complete geographical area sampled. The main advantage of scale-specific wavelet regression is that it differs from previous methods, which simply upscale data by averaging aggregated cells and thus regress the variables of enlarged grain size. Instead, wavelet analysis is able to extract scale-specific variations of both dependent and independent variables. Therefore, a wavelet regression can measure how a change in environmental variables at a given resolution (i.e. focus) influences change in the response variable at the same resolution (Ye *et al.*, 2015). To illustrate our new up-scaling method, it is necessary to use data at medium to large extent and fine sample unit because sample unit acts as a preset for the grain (i.e. finest resolution) in the analysis. Scale dependency is then studied by leaving extent and grain constant and aggregating $2^j \times 2^j$ (with j being a level of analysis) grains to coarser resolutions (i.e. foci). Hence, to discuss scale dependency, that is, alterations in the relative importance of different environmental factors caused by increasingly coarser resolutions (foci), we need data collected over a regular grid consisting of sufficient grid cells. Therefore, in our case study, we examine data sampled on a map of 1024×1024 grid cells at $0.01° \times 0.01°$ resolution (grain), that is, *c.* 1×1 km^2, in Central Europe. We focus on the relationship of vegetation greening to climate, topography and land

use. Remote-sensing vegetation indices based on satellite observations indicating the vegetation activity (Yang *et al.*, 2012) were used to estimate the vegetation period per year (White *et al.*, 2003). We use vegetation period as a response variable that is regressed on climate, topographic and land use data. To the best of our knowledge, this is the first study at intermediate extent and fine grain (and hence large sample size), which covers a very large range of different foci, clearly differentiating among the different components of scale. In principle, it can be assumed that: (1) The impact of climate, topographic as well as land use variables on vegetation period is scale-dependent. (2) Land use variables are more relevant in relation to vegetation period than climate and topographic ones for models with fine resolution, and vice versa for those with coarse resolution (cf. Pearson *et al.*, 2004). Hence, our goal is to demonstrate how variations or fluctuations at multiple spatial scales can be systematically analysed, and to draw specific conclusions regarding these assumptions.

METHODS

Wavelets

The crucial idea behind wavelet analysis can be formulated as follows: wavelets are small waves, that is, localized oscillating functions (Daubechies, 1992; Ma & Zhang, 2015). In a one-dimensional spatial context, one can imagine that such a brief oscillation is locally aligned with a segment of the given transect, thus enabling a comparison between template (wavelet) and original (transect). If there is high similarity, then the absolute value of the corresponding wavelet coefficient is high. If there is low similarity, it is low (Dale & Mah, 1998; Csillag & Kabos, 2002; Ye *et al.*, 2015). By translation, that is, shifting the wavelet along the transect line, one is able to stepwise evaluate the whole transect. Moreover, this transect can be scanned several times by gradually stretched or compressed wavelets, and thus varying in width and oscillating behaviour, which corresponds to scale or resolution. Based on a set of wavelets derived from a prototype (i.e. mother wavelet) and generated by scaling and translation of this original, it is possible to capture the complete information of any transect. Such wavelets of different dilations and locations and their associated scaling functions constitute the so-called wavelet family (Daubechies, 1992). Each wavelet acts as both a window and a filter. One can show that the information of any discrete function f is codable by a wavelet transform, that is, it can be captured by coefficients belonging to a certain wavelet family. If the used wavelet family is a family of orthogonal wavelets, then there exists a minimal set of wavelets, enabling a complete information transfer (Mallat, 1989; Percival & Walden, 2000). The number and kind of coefficients in discrete wavelet transforms (DWT) depends on the number and kind of wavelets used in the analysis, and thus not only on the number of observations, but also on a pre-specified number of resolution levels (Bruce & Gao, 1996). There are two kinds of coefficients: detail and smooth ones, reflecting different oscillating behaviour of mother wavelets and scaling functions and representing the highly varying (detailed) and slowly varying (smooth) parts of function f respectively (Bruce & Gao, 1996; Ma & Zhang, 2015). Subsequently, it is possible to reconstruct the original function f by applying the back transform, that is, the inverse wavelet transform. Moreover, by means of wavelet transform and back transform, one is able to decompose a function into orthogonal components at different scales. These components can be visualized as parts of the function at different resolutions. Therefore, this method is called multiresolution analysis (MRA) (Mallat, 1989; Dong *et al.*, 2008). The MRA algorithm always provides detail components at levels gradually incremented up to a preset limit (D_1, D_2, ..., D_J) and one smooth component at the upper level (S_J).

For illustration, we present the results of a wavelet MRA decomposition stopped at level 3 (Fig. 1b) in comparison to a PCNM analysis limited to four components (Fig. 1a). Both analyses are performed on the same signal vector f, which is a time series (or spatial transect). The example illustrates that PCNM and wavelet decompositions differ in their ability to detect local variations. Only in case of wavelet analysis, the components reveal that signal variability increases with time (or spatial variable). As a consequence, this method yields a perfect reconstruction of the signal from just four components. In general, wavelet analysis is locally more accurate compared to Fourier analysis, which requires many more components.

Two-dimensional wavelet analysis

The use of wavelets in the fields of geophysics, biology, ecology and agriculture is rapidly developing (Kumar & Foufoula-Georgiou, 1997; Torrence & Compo, 1998; Dong *et al.*, 2008). However, most of the concepts for wavelet analysis apply to either time signals (Cazelles *et al.*, 2008) or one-dimensional spatial data, which are much like time series (Dale & Mah, 1998). The need to explore and assess images and landscapes requires a more comprehensive, 2-D wavelet analysis (Csillag & Kabos, 2002). The 2-D DWT enables us to transform a data matrix into a matrix of wavelet coefficients. Therefore, 2-D wavelet analysis allows us to analyse data such as discrete images or geographical patterns of ecological or environmental variables (Csillag & Kabos, 2002). Note that the increased dimensionality results in newly formed wavelets. Four shapes are formed out of the two (detailed and smooth) ones, which are used in case of one dimension (i.e. mother wavelet and scaling function). These new four types of 2-D wavelets are three mother wavelets (applied in different directions: vertically, horizontally and diagonally) and one scaling function. The scaling procedure is dyadic, that is, it is a stepwise enlargement of wavelets by scale factor 2^j in both dimensions: $2^j \times 2^j$, $j = 1, 2, \ldots$If a data matrix has size $2^K \times 2^K$, then level j can range from 1

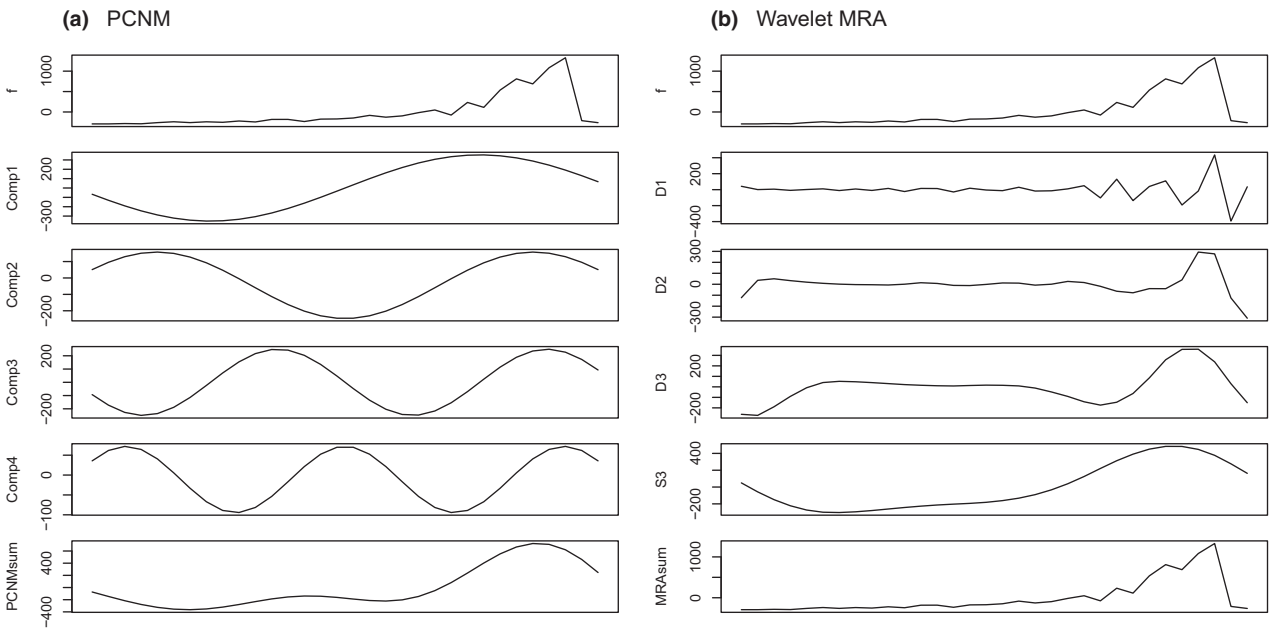

Figure 1 Comparison of signal decomposition and reconstruction obtained for two different methods and constructed for a time series (or spatial transect) containing 32 observations. (a) PCNM is used to decompose the signal f (top panel) into its first four components (mid panels) and to reconstruct the signal by the sum of these four components (lower panel). (b) Wavelet MRA is used to decompose the same signal f (top panel) into four components (mid panels) and to reconstruct the signal by the sum of these four components (lower panel).

to K without any edge effects. Therefore, the 2-D MRA decomposition of a matrix F is

$$F = \sum_{j=1}^{J} D_j^v + \sum_{j=1}^{J} D_j^h + \sum_{j=1}^{J} D_j^d + S_J \qquad (1)$$

This means that it is an additive decomposition into $3J + 1$ components, where the matrices D_j represent the detail (namely high-frequency) parts and the matrix S_J represents the smooth (namely low-frequency) part of matrix F (cf. also Fig. 1b). The matrices D_j^v, D_j^h, D_j^d, and S_J are linear combinations of corresponding 2-D wavelets. The limit parameter J is used to constrain the number of multiresolution components. Note that the smooth matrix S_J exclusively emerges at the coarsest resolution, whereas the matrices D_j accumulate over all scales up to resolution level J. This means that, as a scale-by-scale decomposition, the (particularly resulting) smooth matrices can be decomposed again and again. This finally resulting matrix S_J reflects the averaged or smoothed part of matrix F at maximum resolution, whereas the detail matrices D_j^v, D_j^h, D_j^d arising at every resolution level represent its multiple spatial variations or fluctuations. The three directions of 2-D wavelets are indexed by v, h and d for vertical, horizontal and diagonal respectively. If index m corresponds to these different spatial directions, equation (1) can be written as follows

$$F = \sum_{m=1}^{3} \sum_{j=1}^{J} D_j^m + S_J \qquad (2)$$

For our purpose, the application of a modified version of the DWT is more appropriate. It is called the maximal overlap DWT (MODWT) (Percival & Walden, 2000). The MODWT is a redundant non-orthogonal transform but has the advantage that it provides as many wavelet coefficients per scale and wavelet-type as elements of F are given. Therefore, the wavelet coefficients give information about which frequencies are dominant at which positions in matrix F. The described properties (Eqs. 1 and 2) hold for DWT as well as for MODWT.

Wavelet multiresolution regression

Based on the 2-D MRA decomposition (Eq. 2), that is, the decomposition of matrices into scale-specific subcomponents, we are able to develop a regression technique, which allows scale-specific regressions. Different from other methods such as PCNM or MEM, our approach is applied to the response variable as well as all explanatory variables in a multiple regression. This is possible as the components of all these variables occur in a spatial context as they were sampled in a plane. Thus, we first must convert these vectors into matrices reflecting the spatial layout, that is, where all the components were sampled. Then the 2-D MRA decomposition (2) can be applied to each matrix built in this way. Subsequently, all scales which are not to be analysed in the model have to be eliminated. Therefore, for instance, transform P_{D_j}

$$F \rightarrow P_{D_j} F = \sum_{m=1}^{3} D_j^m \qquad (3)$$

provides a tool keeping only detail matrices of level j.

Finally, we revert to vectors (i.e. convert each matrix of specific scale into a vector) that allow us to continue as usual

in a linear regression where y is a vector of responses and X is a matrix of predictors. According to the above-mentioned interpretation of detail matrices, this means that this regression, keeping only detail matrices of level j, accounts for fluctuations or spatial variations at a specific spatial resolution (i.e. focus).

Note that this wavelet analysis is applicable not only to normal linear models, but also to regressions in which the response variable has a non-normal distribution. In that case, the canonical link function is $\eta_i = g(\mu_i) = x_i'\beta$, $i = 1, 2, \cdots, n$, with the expected value of the response being $\mu_i = E(y_i)$, n is sample size, and β is a vector of unknown parameters. The matrix

$$W = \mathrm{diag}\left\{ v_{ii}^{-1} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \right\} \tag{4}$$

denotes a diagonal weight matrix, where the variance of the response is $A = \mathrm{diag}\{\mathrm{var}(y_i)\} = \mathrm{diag}\{v_{ii}\}$.

In case of generalized linear models, the process of decomposition and scale selection needs to be restarted within each step of the iterative procedure (Carl & Kühn, 2010). Therefore, the generalized iterative solution for parameter β is

$$b^{(m)} = \left(\left(P_{D_j} W^{\frac{1}{2}} X \right)' P_{D_j} W^{\frac{1}{2}} X \right)^{-1} \left(P_{D_j} W^{\frac{1}{2}} X \right)' P_{D_j} W^{\frac{1}{2}} z, \tag{5}$$

where

$$P_{D_j} W^{\frac{1}{2}} z = P_{D_j} W^{\frac{1}{2}} X b^{(m-1)} + P_{D_j} A^{-\frac{1}{2}} (y - \mu) \tag{6}$$

Quality Criteria

On the one hand, it has been suggested that wavelet covariance could provide a useful measure for scale-dependent interactions between, for example, an explanatory variable and the response variable in a regression (Kumar & Foufoula-Georgiou, 1997; White *et al.*, 2003). The scale-specific contributions to sample variance or covariance can be efficiently estimated by MODWT wavelet coefficients (Lark & Webster, 2004). The 2-D wavelet variance can be defined as an extension to the 1-D approach as

$$\mathrm{var}(j) = \frac{1}{n} \sum_{m=1}^{3} \sum_{i=1}^{n} \left(d_{im}^{F}(j) \right)^2 \tag{7}$$

where $d_{im}^{F}(j)$ are the detail wavelet coefficients of matrix F at level j and n is the sample size. Likewise, the 2-D wavelet covariance, which is a scale-dependent component of covariance for two matrices F and G, is given by

$$\mathrm{cov}(j) = \frac{1}{n} \sum_{m=1}^{3} \sum_{i=1}^{n} \left| d_{im}^{F}(j) \cdot d_{im}^{G}(j) \right| \tag{8}$$

Here the matrices F and G have to be produced from vectors f and g as described above. The vectors f and g have to be standardized, to ensure that the wavelet transforms are comparable to each other.

On the other hand, regression methods may give us deeper insight into the variety of factors and its relations to a response variable. Generally, regression methods do not only aim for estimation of slope coefficients, but also for hypothesis testing and P-values in order to decide whether the predictors are significant or which subset of predictors is relevant. However, problems can arise if we want to compare P-values of different models, in particular, models of different sample size. This is due to the fact that the power of the test depends on sample size (McDonald, 2009). When decreasing sample size, the power of the test declines. As we perform a multiscale analysis eliminating step by step scale-specific subcomponents and thus rendering certain resolutions ineffective, as a matter of fact the sample size does change. To avoid comparisons of significance tests across scales and instead to provide a consistently good quality criterion, we use the approach of model selection based on multimodel inference (MMI) developed by Burnham & Anderson (2002). Our scale-dependent regression analysis outlined above (Eqs. 5 and 6) allows us to separately apply the method at each scale. This is required because the set of candidate models must always be related to the same data set. Therefore, at each scale, MMI can make statistical inference via the full set of candidate models and model ranking by means of Akaike's information criterion (AIC). In order to estimate the relative importance of a variable, so-called Akaike weights, that is, normalized likelihoods of AIC differences, are introduced as model weights. The sum of Akaike weights over the subset of models that include a certain variable can then be considered as a measure for the importance of this variable. Only these values of relative variable importance, that is, relative instead of absolute values, are the measures that are eventually used for evaluations and comparisons across scales. In our application, it is necessary not only to estimate regression coefficients, but also to calculate an effective sample size for log-likelihoods and thus Akaike weights. That is because most of the information in the 1024×1024 grid is redundant due to the positive spatial dependence in the data. This means that individual observations include information already present in nearby observations, so that neighbouring grid cells are highly correlated and the effect or value of sample size is less than the number of observations (Dale & Fortin, 2009). The estimated degree of spatial autocorrelation can therefore be used to adjust the sample size, that is, to determine how much smaller this effective sample size is (Dale & Fortin, 2005). As shown by Dale & Fortin (2009), the effective sample size n_{eff} for correcting for autocorrelation can approximately be calculated by the following formula:

$$n_{\mathrm{eff}} = \frac{n^2}{n + 2\sum_{k=1}^{n-1} (n-k) r(k)} \tag{9}$$

where $r(k)$ is the autocorrelation at lag k calculated on $n - k$ pairs of observations. We use Moran's I values of the residuals of the full model as a 2-D, radially symmetric approximation for the function $r(k)$. For the redundant MODWT based on sample size n across all levels, n_{eff} is a good approximation for all levels.

All statistical analyses were performed using R x64 3.0.1 (R Development Core Team, 2013). The R code is given in

Appendix S2. The tools for calculating wavelet transforms are available in package *waveslim* (Whitcher, 2005). Except where explicitly noted otherwise, the results given in the following sections were calculated using d4 wavelets. We tested other types of wavelets as well, for example, Haar and d16 wavelets. These analyses yielded almost the same results (not presented here for brevity) as those for d4 wavelets. Differences are hardly detectable, except for the highest levels where the loss of information generally causes expanded uncertainties.

DATA

We extracted satellite data from the MEDOKADS NOAA AVHRR data archive provided by the Meteorological Institute of the Freie Universität Berlin. Signals from satellite observations are available as so-called normalized difference vegetation index *(NDVI)* values. *NDVI* based on the daily reflectance in the red (Red) and near infrared *(NIR)* AVHRR bands

$$NDVI = (NIR - Red)/(NIR + Red) \qquad (10)$$

detects the part of photosynthetically relevant radiation absorbed by plants. Thus, *NDVI* is accepted as a good indicator of the vitality and photosynthetic activity of the vegetation, that is, *NDVI* indirectly indicates seasonal changes in leaf and shoot growth and in the greenness of the vegetation. Therefore, it can be used to estimate the length of the vegetation period (White *et al.*, 2003). We computed estimates of vegetation period (see Appendix S1 in Supporting Information), averaged the values over the years 1989–2007, and provided a map representative of Central Europe (Fig. 2). This map has a resolution of $0.01^{\circ} \times 0.01^{\circ}$ (*c.* 1×1 km^2) and consists of $2^{10} \times 2^{10} = 1024 \times 1024$ grid cells, allowing a dyadic up-scaling from level 1 to level 10 (The original matrix can be indexed by level 0). The selected area ranges across 10.24 degrees of both latitude and longitude. The Alps as well as coastal regions are included. The altitude ranges from sea level to 4300 m. Thus, the extent of environmental variation in the region is remarkable. Vegetation period is generally highest for the land cover type grassland especially in southern Germany followed by forests. Agricultural areas generally display a shorter vegetation period particularly in areas with extensive irrigation or even controlled flooding such as rice fields in northern Italy. Here, the satellite cannot receive vegetation signals due to surfaces covered extensively by water until shortly before harvest. Elevated areas exhibit shorter vegetation periods, most

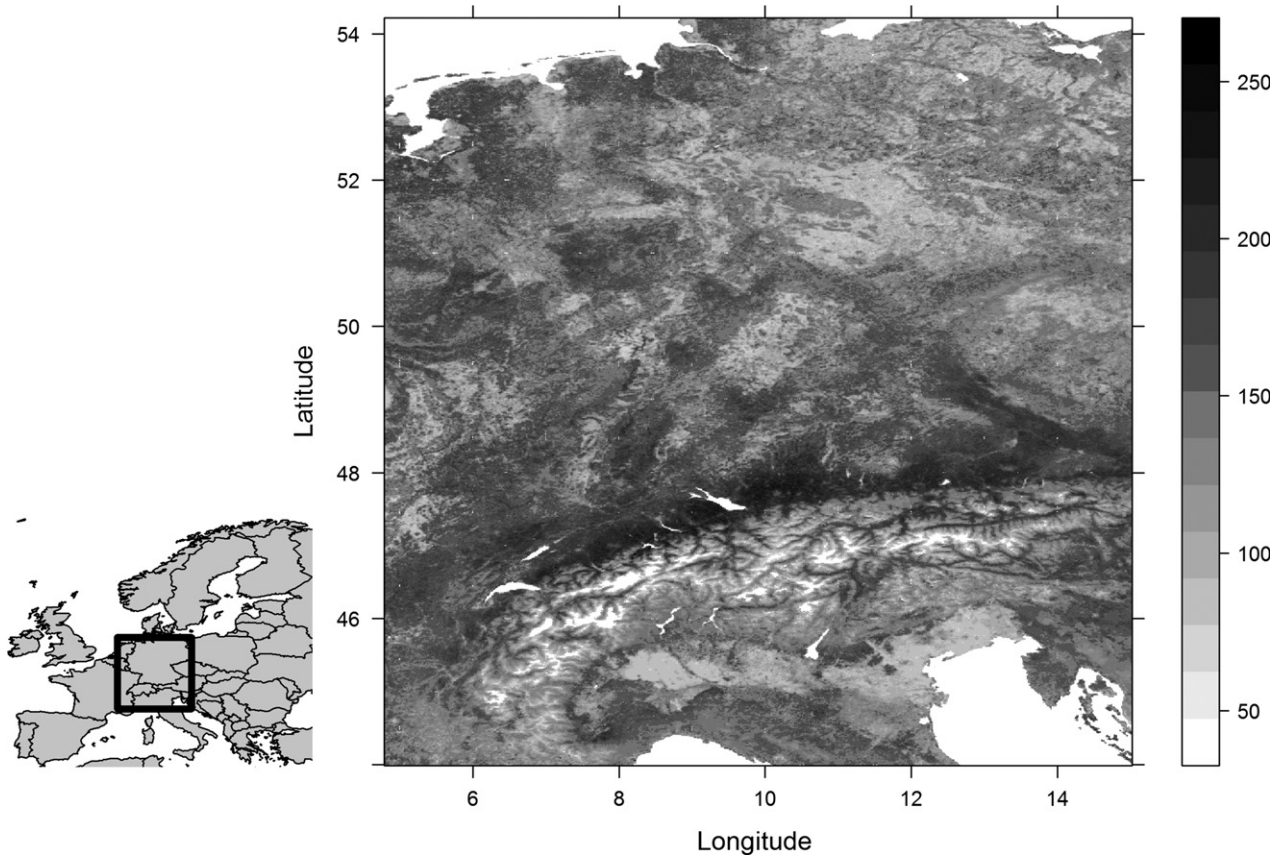

Figure 2 Map of vegetation periods in Central Europe (latitude 43.99°–54.22° N, longitude 4.79°–15.02° E). The data are estimates based on remote-sensing vegetation indices sampled on a 1024 ×1024 grid at a resolution of *c.* 1 ×1 km^2. The vegetation period given in days is an average over the years 1989–2007 (greyscale to the right of the map). A map of Europe displaying the selected map section for Central Europe is shown in the inset.

notably for the Alps. Lakes, glaciers and bare land obviously do not display any vegetation period.

Moreover, we extracted climate variables from the WorldClim database Version 1.4 (Hijmans *et al.*, 2005), elevation data from the WorldClim data base, and land cover data from Corine Land Cover 2006 vector data Version 17 (EEA, 2013) and assigned them to the vegetation grid (see Appendix S1 in Supporting Information). This enables us to use vegetation period as the response or outcome variable for further analyses. It is regressed on climate, topographic and land use data, that is, annual mean temperature (Bio1), annual precipitation (Bio12), altitude and the land cover categories artificial areas, agricultural areas, forests and grass/scrublands.

RESULTS

The extent of variation as a function of increasing spatial resolutions is represented as wavelet variance (Eq. 7, Fig. 3a). This procedure of up-scaling related to the resolution level can be imagined as a two-dimensionally gradual (i.e. dyadic) enlargement of sizes of grid cells. Roughly speaking, all variances of land use variables show decreasing values with increasing resolution level, that is, increasing cell size, while variances of other variables show opposite trends. This becomes particularly evident for the levels from 1 (*c.* 4 km^2) to 5 (*c.* $2^5 \times 2^5$ km$^2 \approx 1000$ km^2 resolution), whereas higher levels yield other results. Note that because of the loss of information with increasing level, the levels higher than 7 provide less reliable results than others.

Wavelet covariances (Eq. 8) evaluating the relationship between explanatory variables and response variables are informative as well (Fig. 3b). One can say at least that both Figs. 3a and 3b reveal that the contributions of climatic, topographic as well as land use variables vary considerably across resolutions. Moreover, the difference of land use variables compared to other ones is detectable in both cases. The relevance of land use variables compared to others switches approximately at level 5 or 6 (cell length *c.* $2^6 = 64$ km, cell size *c.* 4000 km^2).

A more accurate analysis of the true amount of relative variable importance can be based on 2-D MRA decompositions, scale-specific regressions and Akaike weights as described above. This relative variable importance shows that it indeed provides more detailed curves and thus a deeper insight into what variances and covariances roughly reflect (Fig. 4a). All land use variables appear clearly dominant at the levels 3 (cell size *c.* 64 km^2) to 5 (cell size *c.* 1000 km^2). The switch of importance between land use variables and other ones occurs between level 5 and 6.

DISCUSSION

The relative variable importance visualized in Fig. 4a shows clear and strong dependency on resolution. Thus, our first assumption, stated at the end of the Introduction, is supported: the impact of climate, topographic as well as land use variables on vegetation period is scale-dependent. Our second assumption was: land use variables are more relevant in relation to vegetation period than climate and topographic ones for models with fine resolution, and vice versa for those with coarse resolution. Although Fig. 3b supports this assumption and Fig. 4a also detects this switch of variable importance between level 5 and 6, this statement cannot be accepted without qualification. Figure 4a provides more details. It shows that, at resolution levels 1 and 2, temperature as climatic variable is more important than all the

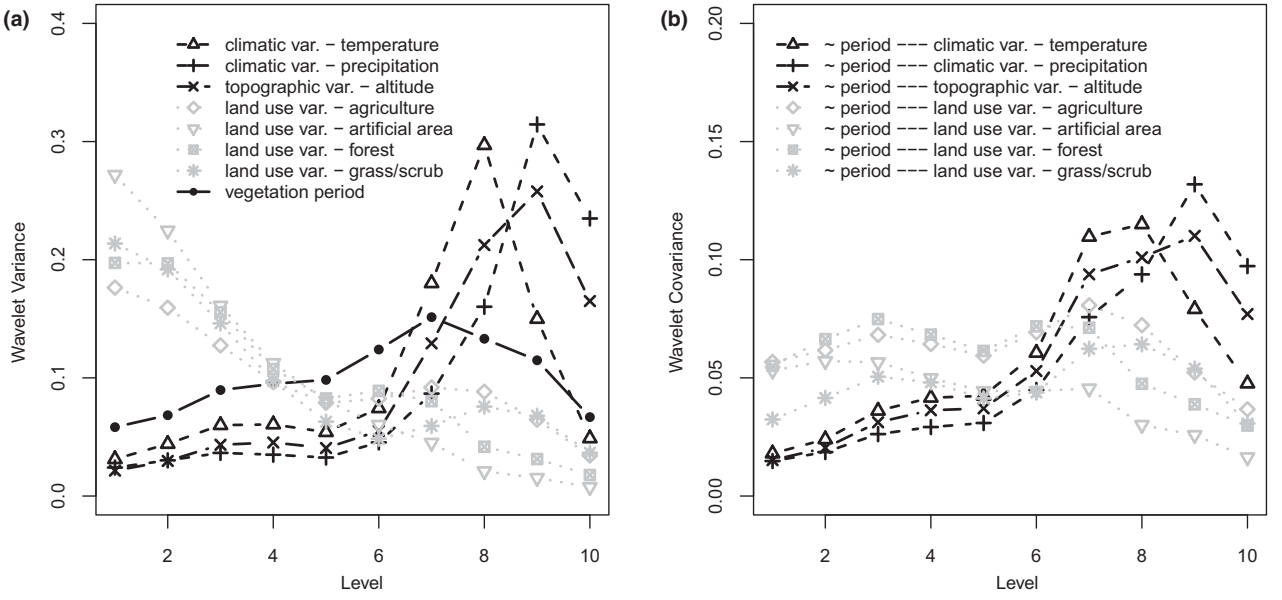

Figure 3 Wavelet variance (a) and covariance (b) for different variables. Levels range from 1 (cell size *c.* 4 km^2) to 10 (coarsest resolution, cell size *c.* 1,000,000 km^2). (Relationships with) Land use variables are indicated by grey lines and symbols, all others are indicated by black ones.

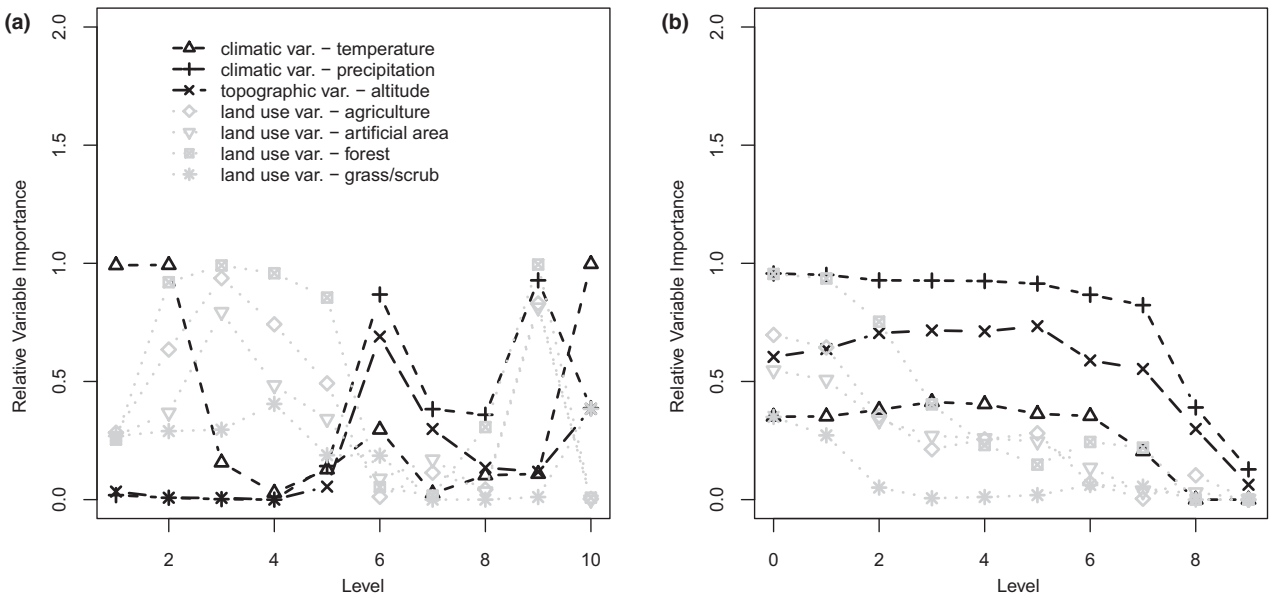

Figure 4 Relative variable importance (measured as Akaike weights) as a function of scale. Land use variables are indicated by grey lines and symbols, all the others are indicated by black ones. (a) The analysis is based on scale-specific regressions capturing only *detail* components as scale-specific subcomponents. Levels range from 1 (cell size *c.* 4 km^2) to 10 (coarsest resolution, cell size *c.* 1,000,000 km^2). (b) The analysis is based on scale-specific regressions capturing only *smooth* components as scale-specific subcomponents. Levels range from 0 (grain size *c.* 1 km^2, finest resolution, no wavelet transform) to 9 (grain size *c.* 250,000 km^2).

others, especially than land use variables. One reason for this could be that landscape structure and grid cell structure for remote-sensing images are different, that is, patches of certain land use are not spatially congruent with AVHRR pixels. This leads us to conclude that this difference between landscape and raster format becomes increasingly relevant at lower levels. Note that substituting sample size n with effective sample size n_{eff} in log-likelihoods has no impact on the ranking order in MMI, only the relative distance of weights is reduced. A change in sample size cannot inflate the importance of a single variable, enabling it to dominate and earn the top-ranking position. Therefore, the importance of temperature at fine scales (i.e. levels 1 and 2) is probably an artefact of misassigned land use/land cover.

On the whole, our examination of resolution-specific variability has shown that there is a marked change in the importance of drivers of the ecosystem process of vegetation greening at a cell length slightly above 32 km. Considering previous knowledge (e.g. Willis & Whittaker, 2002; Pearson & Dawson, 2003) this does not seem unexpected. Similarly to our result, Luoto *et al.* (2007) found that species distribution models of birds in Finland improved notably by including land use data in addition to climate data at resolutions of 10 km and 20 km while at resolutions of 40 km and 80 km climate was sufficient. In this context, it is important to note that our analyses are examinations of the scale-specific variability. Such analyses capture and evaluate resolution-specific variations of variables and their relations, that is, relations of local fluctuations at a given scale. This is because the previously used wavelet analysis captures only the detail (i.e. high-frequency) components.

However, it seems that many previous studies did not clearly differentiate between resolution-specific (i.e. focus-specific) variation analysis and upscaling by averaging of aggregated cells. As explained above, one has to distinguish between four different components of scale (Scheiner *et al.*, 2000): sample unit, grain, focus and extent. Willis & Whittaker (2002), inter alia, simply discussed the importance, very generally, of 'scale', whereas Pearson & Dawson (2003) were more concrete and translating them into different extents. Still, it is clear that the importance of specific ecosystem processes does not only depend on extent but also on grain and focus. Analysing a process with a grain of 1 km^2 at an extent of 100 km^2 will quite likely yield a different result than having the same grain and continental extent. Hence, we hypothesize that not only varying focus (i.e. resolution level) and keeping grain and extent constant (as we did) will have an effect but also varying grain and keeping extent constant or keeping grain constant and varying extent will have an effect on the hierarchy of drivers of ecosystem processes (see also Rahbek, 2004). It hence is clear that studies using the same data sets can come to different conclusions (Thuiller *et al.*, 2004; Pompe *et al.*, 2008). Previous studies using cross-scale analyses frequently employ several generalized additive or linear models with intermediate (Luoto *et al.*, 2007) or large resolutions (Rahbek & Graves, 2001) but are often unclear on how the scaling precisely was done. We can assume that many authors simply averaged or lumped data from finer to coarser units.

It is worth noting that data averaging of aggregated cells changes grain size, but not focus. It can be visualized as a smoothing. Based on this interpretation, one might ask the

question, what if matrix S_J is always included in scale-specific regressions. The matrix S_J is the smooth or low-frequency part the MRA decomposition of matrix F (see Eq. 2). The interpretation that can be given in this case is as follows: As a consequence of the relation

$$\sum_{m=1}^{3} D_J^m + S_J = S_{J-1}, \quad (11)$$

only smooth components would be under consideration. A scaling procedure based on smooth components can be imagined as a smoothing over gradually enlarged 2-D grid cells. In particular, smooth MRA components of Haar wavelets can be seen as a series of averaging operations. The results for scale-specific regressions for such smooth components are given in Fig. 4b. (Note that the level shift is a consequence of the index shift in Eq. 11). These results reflect what we have to expect when the analysis is carried out for data sampled on increasingly larger grid cells, that is, it is an analysis quantifying the effect of increasing grain. It can clearly be seen that if the map is split up into increasingly larger grains, the relative importance of land use variables decreases, whereas precipitation as the variable with the smoothest curve is dominant across all levels and also the relative importance of temperature and altitude remains stable across all levels except for the highest, most uncertain ones.

CONCLUSIONS

To the best of our knowledge, this is the first study analysing the scale dependency of an ecosystem process, clearly distinguishing between the different components of scale, namely extent, grain and focus, having an extremely large sample size (n = 1,048,576), and covering a large range of different resolutions (*c.* 1 km^2 to *c.* 1,000,000 km^2). In summary, our method has the advantage that all calculations were done in a single framework. Firstly, the wavelet approach is carried out by means of multiresolution analysis, which is able to decompose gridded data (maps or images) into components at different resolutions. This data decomposition is embedded into the framework of a multiple regression analysis (Keitt & Urban, 2005; Carl & Kühn, 2008). This wavelet multiresolution regression (WMRR) also allows response vectors of binary or Poisson distribution. Therefore, our WMRR approach is a method for applying 2-D wavelet analysis to generalized linear models. Secondly, applying all regressions in a multimodel inference approach circumvents a common problem: Using separate regressions for each scale will result in multiple testing. Due to decreasing sample size, hypothesis tests have declining power. Therefore, results cannot be compared by means of hypothesis testing. The multimodel inference approach does not suffer from this problem. It calculates variable importance by using an information-theoretic approach based on Akaike weights (Burnham & Anderson, 2002). As results obtained from finer to coarser scaled data can then be compared, one is able to examine the effect of scale dependencies and to evaluate the relative importance of several environmental variables across different spatial scales. Therefore, we provided an answer to the key question whether similar mechanisms act at different spatial scales. We applied our method to data on climate variables and land cover data to explain variation in vegetation greening as an example of an ecosystem process. Our results indicated that the relative variable importance detectable by scale-specific regressions is strongly scale-dependent. Moreover, for two different approaches, (1) leaving grain and extent constant and changing focus and (2) leaving extent constant and changing grain, we were able to demonstrate how 2-D scale dependencies can be systematically analysed. It was shown at which 'scale' the turning point is where drivers change in importance.

Finally, we believe that our newly proposed method is particularly suitable for studying scale dependencies of various spatial processes on environmental drivers having gridded data with sufficiently large sample size. Roughly speaking, this means that not less than five levels should be analysed to check whether scale-dependent changes occur in variable importance. Therefore, a quadratic matrix of at least $2^5 \times 2^5 = 1024$ elements is needed for each of the predictor and response variables.

ACKNOWLEDGEMENTS

We acknowledge support by the EU infrastructure project 'Distributed Infrastructure for EXPErimentation in Ecosystem Research – EXPEER' (project reference 262060) and the Integrated Projects SCALES (grant 226 852) funded under the European Commission Framework Programme (FP) 7. We are grateful to Richard Pearson, Jonathan Belmaker and two referees for constructive critiques of the manuscript.

REFERENCES

Borcard, D., Legendre, P., Avois-Jacquet, C. & Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, **85**, 1826–1832.

Bruce, A. & Gao, H.Y. (1996) *Applied wavelet analysis with S-PLUS*. Springer, New York.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference*. Springer, New York.

Carl, G. & Kühn, I. (2008) Analyzing spatial ecological data using linear regression and wavelet analysis. *Stochastic Environmental Research and Risk Assessment*, **22**, 315–324.

Carl, G. & Kühn, I. (2010) A wavelet-based extension of generalized linear models to remove the effect of spatial autocorrelation. *Geographical Analysis*, **42**, 323–337.

Cazelles, B., Chavez, M., Berteaux, D., Menard, F., Vik, J.O., Jenouvrier, S. & Stenseth, N.C. (2008) Wavelet analysis of ecological time series. *Oecologia*, **156**, 287–304.

Csillag, F. & Kabos, S. (2002) Wavelets, boundaries, and the spatial analysis of landscape pattern. *Ecoscience*, **9**, 177–190.

Dale, M.R.T. (1999) *Spatial pattern analysis in plant ecology*. Cambridge Univ. Press, Cambridge.

Dale, M.R.T. & Fortin, M.J. (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, Cambridge.

Dale, M.R.T. & Fortin, M.J. (2009) Spatial autocorrelation and statistical tests: some solutions. *Journal of Agricultural, Biological, and Environmental Statistics*, **14**, 188–206.

Dale, M.R.T. & Mah, M. (1998) The use of wavelets for spatial pattern analysis in ecology. *Journal of Vegetation Science*, **9**, 805–814.

Daubechies, I. (1992) *Ten lectures on wavelets. CSBM-NSF series application mathematics*, vol. 61. SIAM Publication, Philadelphia.

Dong, X., Nyren, P., Patton, B., Nyren, A., Richardson, J. & Maresca, T. (2008) Wavelets for agriculture and biology: a tutorial with applications and outlook. *BioScience*, **58**, 445–453.

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

EEA (2013) Corine Land Cover 2006 seamless vector data. Available at: http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector-data-version-3#tab-gis-data.

Fortin, M.J. & Dale, M. (2005) *Spatial analysis*. A guide for ecologists. Cambridge Univ. Press, Cambridge.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Jombart, T., Dray, S. & Dufour, A.B. (2009) Finding essential scales of spatial variation in ecological data: a multivariate approach. *Ecography*, **32**, 161–168.

Keil, P., Schweiger, O., Kühn, I., Kunin, W.E., Kuussaari, M., Settele, J., Henle, K., Brotons, L., Pe'er, G., Lengyel, S., Moustakas, A., Steinicke, H. & Storch, D. (2012) Patterns of beta diversity in Europe: the role of climate, land cover and distance across scales. *Journal of Biogeography*, **39**, 1473–1486.

Keitt, T.H. & Urban, D.L. (2005) Scale-specific inference using wavelets. *Ecology*, **86**, 2497–2504.

Kumar, P. & Foufoula-Georgiou, E. (1997) Wavelet analysis for geophysical applications. *Reviews of Geophysics*, **35**, 385–412.

Lark, R.M. & Webster, R. (2004) Analysing soil variation in two dimensions with the discrete wavelet transform. *European Journal of Soil Science*, **55**, 777–797.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–1067.

Lookingbill, T.R., Rocca, M.E. & Urban, D.L. (2011) Focused assessment of scale-dependent vegetation pattern. *Predictive species and habitat modeling in landscape ecology* (ed. by C.A. Drew, Y.F. Wiersma and F. Huettmann), pp. 111–138. Springer, New York.

Luoto, M., Virkkala, R. & Heikkinen, R.K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography*, **16**, 34–42.

Ma, Z. & Zhang, L. (2015) Modeling bird species richness at multiple spatial scales using two-dimensional wavelet analysis. *Forest Science*, **61**, 1–16.

Mallat, S.G. (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674–693.

McDonald, J.H. (2009) *Handbook of biological statistics*. Sparky House Publishing, Baltimore, Maryland.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

Pearson, R.G., Dawson, T.P. & Liu, C. (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, **27**, 285–298.

Percival, D.B. & Walden, T. (2000) *Wavelet methods for time series analysis*. Cambridge University Press, Cambridge.

Pompe, S., Hanspach, J., Badeck, F., Klotz, S., Thuiller, W. & Kühn, I. (2008) Climate and land use change impacts on plant distributions in Germany. *Biology Letters*, **4**, 564–567.

R Development Core Team (2013) *R: a language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, {ISBN} 3-900051-07-0, Available at: http://www.R-project.org.

Rahbek, C. (2004) The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, **8**, 224–239.

Rahbek, C. & Graves, G.R. (2001) Multiscale assessment of patterns of avian species richness. *Proc Natl Acad Sci U S A*, **98**, 4534–4539.

Scheiner, S.M., Cox, S.B., Willig, M., Mittelbach, G.G., Osenberg, C. & Kaspari, M. (2000) Species richness, species-area curves and Simpson's paradox. *Evolutionary Ecology Research*, **2**, 791–802.

Schröder, B. & Seppelt, R. (2006) Analysis of pattern-process-interactions based on landscape models – overview, general concepts, methodological issues. *Ecological Modelling*, **199**, 505–516.

Stark, H.G. (2005) *Wavelets and signal processing. An application-based introduction*. Springer, Berlin.

Thuiller, W., Araújo, M.B. & Lavorel, S. (2004) Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography*, **31**, 353–361.

Torrence, C. & Compo, G.P. (1998) A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, **79**, 61–78.

Whitcher, B. (2005) Waveslim: basic wavelet routines for one-, two- and three-dimensional signal processing. *R package version*, **1**, 5.

White, M.A., Brunsell, N. & Schwartz, M.D. (2003) Vegetation phenology in global change studies. *Phenology: an integrative environmental science* (ed. by M.D. Schwartz), pp. 453–466. Kluwer Academic Publishers, Dodrecht.

Willis, K.J. & Whittaker, R.J. (2002) Ecology. Species diversity–scale matters. *Science*, **295**, 1245–1248.

Wu, J. & Hobbs, R. (2002) Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. *Landscape Ecology*, **17**, 355–365.

Yang, Y., Xu, J., Hong, Y. & Lv, G. (2012) The dynamic of vegetation coverage and its response to climate factors in Inner Mongolia, China. *Stochastic Environmental Research and Risk Assessment*, **26**, 357–373.

Ye, X., Wang, T., Skidmore, A.K., Fortin, D., Bastille-Rousseau, G. & Parrott, L. (2015) A wavelet-based approach to evaluate the roles of structural and functional landscape heterogeneity in animal space use at multiple scales. *Ecography*, **38**, 740–750.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Additional information about data sets.
Appendix S2 R code for calculating scale-specific regressions.

BIOSKETCH

Gudrun Carl received her PhD in theoretical physics and gained experience in the field of mathematics. Her recent field of research is the development of methods for spatial and temporal analysis of environmental data.

Author contributions: G.C. and I.K. conceived the ideas; D.D. and O.S. provided the data; G.C. analysed the data and led the writing with all authors contributing to the text.

Editor: Richard Pearson