Spatio-temporal regression on compositional covariates: modeling vegetation in a gypsum outcrop

SPATIO-TEMPORAL REGRESSION ON COMPOSITIONAL COVARIATES: MODELING VEGETATION IN A GYPSUM OUTCROP

Francesca Bruno, Fedele Greco, Massimo Ventrucci

Dipartimento di Scienze Statistiche "P. Fortunati",

University of Bologna, Italy

SUMMARY

Investigating the relationship between vegetation cover and substrate typologies is important for habitat conservation. To study these relationships, common practice in modern ecological surveys is to collect information regarding vegetation cover and substrate typology over fine regular lattices, as derived from digital ground photos. Information on substrate typologies is often available as compositional measures, e.g., the area proportion occupied by a certain substrate. Two primary issues are of interest for ecologists: first, how much substrate typologies differ in terms of relative suitability for vegetation cover and, second, whether suitability varies over time. This paper develops a novel procedure for managing compositional covariates within a Bayesian hierarchical framework to effectively address the aforementioned issues. A spatio-temporal model is adopted to estimate the temporal pattern characterizing substrate relative suitability for vegetation cover and, at the same time, to account for spatio-temporal correlation. Relative suitability is modeled by time-varying regression coefficients, and spatial, temporal and spatio-temporal random effects are modeled using Gaussian Markov Random Field models.

KEYWORDS: binomial data; compositional covariates; Intrinsic Gaussian Markov Random Fields; hierarchical Bayesian model; vegetation cover; suitability

1 INTRODUCTION

In ecological applications, over the past years, there has been an increasing interest in collecting and analyzing datasets with both spatial and temporal dimensions. The increasing use of imaging tools able to collect and process digital ground photos managed by Geographic Information System (GIS) (Bennett et al. 2000; Richardson et al. 2001; Booth et al. 2005) has introduced ecologists to a new framework in which the analysis of ecological outcomes, such as species richness, abundance and vegetation cover, is based on high-resolution spatially referenced data. Vegetation cover is one of the most important indicators of rangeland condition and plays a key role in habitat management and conservation policies (Booth and Tuller 2003).

In this paper, we investigate the relationship between vegetation cover and substrate typologies within a rupicolous basophilic habitat defined as priority by the European Commission (Council Directive 92/43/CEE). Available data consist of post-processed ground photos taken at several points throughout the study period, which provide information about vegetation cover and substrate typology in a fine regular lattice.

Modeling vegetation cover in terms of substrate typologies is crucial to evaluating substrate suitability, i.e., substrates' natural ability to support vegetation. Knowledge concerning substrate suitability is strategic in predicting future developments and reactions to possible environmental changes in the considered habitat. The ecological behavior of vegetation in arid and rocky environments has been investigated in several studies (Kuntz and Larson 2006; Pueyo and Alados 2007).

Regression models, often including spatio-temporal components, have been largely adopted in the ecological literature to investigate the relationships between ecological outcomes and environmental factors (Guisan and Zimmermann 2000; Fortin et al. 2012;

Kneib et al. 2008). A flexible framework for ecological data is represented by hierarchical Bayesian models (Wikle 2003). The hierarchical structure is useful for accounting for multiple sources of uncertainty and allows for model complexity to be handled in a suitable manner, namely by factorizing the joint distribution into simple conditional models. The role of hierarchical modeling in ecological studies has been discussed in Latimer et al. (2006).

In this paper, a hierarchical Bayesian approach is adopted to investigate the temporal pattern of substrate suitability for vegetation, considering four types of substrate that are found in the habitat under examination, namely moss, litter, soil and bare rock. To properly capture substrate suitability, spatio-temporal correlation must be taken into account: when handling lattice data, Gaussian Markov Random Fields (GMRFs, Rue and Held 2005) are popular tools for modeling structured effects in space and time or both.

A peculiar feature of the case study is that data concerning substrate typology are compositional, i.e., at a given time, substrate data are expressed as the proportion of cell grid occupied by each type of substrate. The compositional nature of substrate information implies an additional challenge for statistical modeling: regression on compositional covariates requires suitable techniques to obtain meaningful regression coefficients indicating the effect of each single compositional part. Analyses involving compositional data begin by mapping the compositional values belonging to the simplex onto real space by any of several possible transformations (Aitchison 1986; Egozcue and Pawlowsky-Glahn 2006). A popular transformation for compositional data is the isometric log-ratio (*ilr*) transformation (Egozcue et al. 2003), which has the advantage of isolating the contribution of a given compositional part with respect to all others. In the context of linear regression on compositional covariates, *ilr* transformation was

adopted by Hron et al. (2012), who proposed an inferential procedure that requires the estimation of as many models as the number of compositions.

In the present paper, starting from the proposal of Hron et al. (2012), a new approach for handling compositional covariates in regression models is developed within a Bayesian hierarchical framework. The proposed strategy provides regression coefficients that are easy to interpret as relative suitability values.

This paper is organized as follows. A motivating example describing the structure of the data concerning an ecological application is outlined in section 2. In section 3, the modeling approach is discussed, and the procedure for handling compositional covariates is proposed. The application results are discussed in section 4. The concluding remarks are provided in section 5.

2 MOTIVATING EXAMPLE

Our focus is on a study designed within the framework of the priority defined by the European Commission (Council Directive 92/43/CEE), with the aim of investigating habitat "6110* Rupicolous calcareous or basophilic grasslands of the *Alysso-Sedion albi*". A sampling campaign was performed on a gypsum outcrop within the Site of Community Importance "IT4050001 Gessi Bolognesi e Calanchi dell'Abbadessa" located in the Emilia Romagna Region, Italy. For a detailed description of the case study, see Velli (2014).

Data were collected over a study period going from April 2012 to late March 2013. Sampling campaigns were adjourned during the dryness period of August and September, when low vegetation cover is expected, and on January because of snow cover. A total of nine sampling campaigns were conducted at unequally spaced times,

denoted as $t = 1,...,9$. The study area consists of a $1.5 \times 1.5m$ *quadrat*, structured as a regular lattice of dimension 30x30, containing 900 grid cells, indexed by $s = 1,...,900$, which are denoted as *plots* in the ecological literature . At each time, a ground-photo of the study area was taken and then processed via GIS algorithms to produce a digital image for vegetation cover and ground composition. Namely, ground composition considers $D = 4$ substrates: *moss*, *litter*, *soil* and *bare rock*. . Each grid cell (plot) in the digital image provides information collected over $n = 100$ pixels. At time t and plot s, the number of pixels covered by a plant is denoted as y_{st}, and the proportions of substrate typologies are collected in the vector $\boldsymbol{z}_{st} = (z_{1st}, z_{2st}, z_{3st}, z_{4st})$.As an example, z_{1st} is the count, or the proportion, of pixels covered by moss in plot s at time t; analogously we obtain the substrate proportions for litter (z_{2st}), soil (z_{3st}) and bare rock (z_{4st}). It is worth noticing that, at any time t and plot s, $z_{1st} + z_{2st} + z_{3st} + z_{4st} = n$, thus substrate typology $\boldsymbol{z}_{st}$ is a compositional variable.

Some descriptive statistics concerning pixel-specific vegetation cover (expressed as y_{st}/n) are summarized in Table 1, where means and quartiles over the study area are reported for each time.

T	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Mean	0.158	0.127	0.102	0.046	0.195	0.258	0.338	0.340	0.393
Q1	0.010	0.000	0.000	0.000	0.000	0.000	0.040	0.060	0.110
Median	0.060	0.040	0.000	0.000	0.080	0.140	0.210	0.220	0.310
Q3	0.210	0.150	0.110	0.000	0.310	0.420	0.600	0.560	0.640
Moran's *I*	0.433	0.569	0.510	0.469	0.534	0.600	0.612	0.615	0.614

Table 1 Descriptive statistics of vegetation cover (y_{st}/n) over time

The average vegetation cover changes substantially across time. A decreasing trend is observed until $t = 4$; at the fourth time point, vegetation is almost absent, later it

increases in terms of both the number of pixels where vegetation is present (i.e., Q1 is greater than zero only for $t \geq 7$) and average coverage. Last row in Table 1 reports the Moran's I statistics measuring spatial correlation. Moran's I index (Moran 1948) measures the degree of correlation between observations at neighboring plots. This index takes continuous values from -1 (maximum negative spatial correlation) to 1 (maximum positive spatial correlation) with 0 indicating absence of spatial correlation.

A spatio-temporal representation of the vegetation cover is reported by maps in Figure 1.The spatial correlation characterizing the vegetation cover measured at the pixel level is considerable throughout the entire study period and becomes more evident from the fifth to the last time point, as evidenced by the Moran's *I* values in Table 1. Overall, the data show both spatial and temporal correlation.

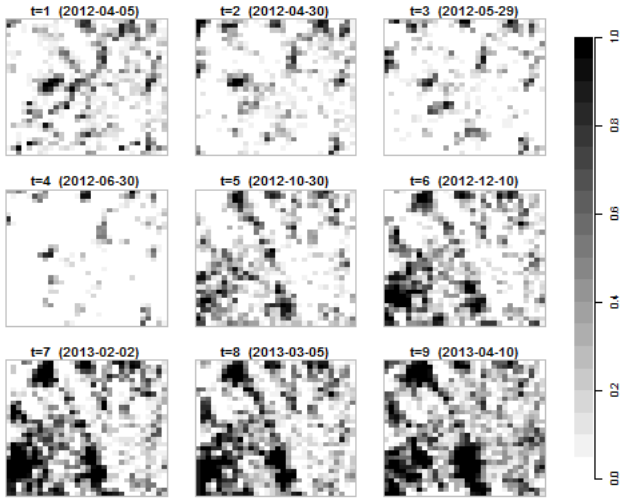


Fig. 1 Maps of vegetation cover at each collection time (indicated within brackets)

In Table 2, the pixel-specific shares z_{st} of substrate typologies are summarized, demonstrating the existence of temporal variability both within and between substrates.

T	1	2	3	4	5	6	7	8	9
Moss	0.46	0.30	0.35	0.35	0.42	0.50	0.57	0.54	0.49
Litter	0.29	0.15	0.14	0.19	0.06	0.01	0.03	0.01	0.01
Soil	0.09	0.30	0.26	0.14	0.22	0.23	0.17	0.20	0.18
Bare rock	0.16	0.25	0.25	0.32	0.30	0.26	0.23	0.25	0.32

Table 2 Share of substrate typology averaged over the study region at each collection time

To investigate the effect of substrate typology on vegetation cover, the marginal average cover over time (see the first row of Table 1) is considered as a benchmark to be compared to the substrate-specific conditional average. These syntheses are shown in the four panels of Figure 2 as solid and dashed lines, respectively. Noisy behavior characterizes the pattern observed on litter substrates, particularly at the end of the study period, when such a typology is scarcely represented in the region (see Table 2). Vegetation cover on moss, soil and bare rock substrates shows a decreasing trend until summer, followed by an increasing trend; such a pattern reflects the general behavior of vegetation over the study period due to the natural vegetation life cycle in dry grasslands.

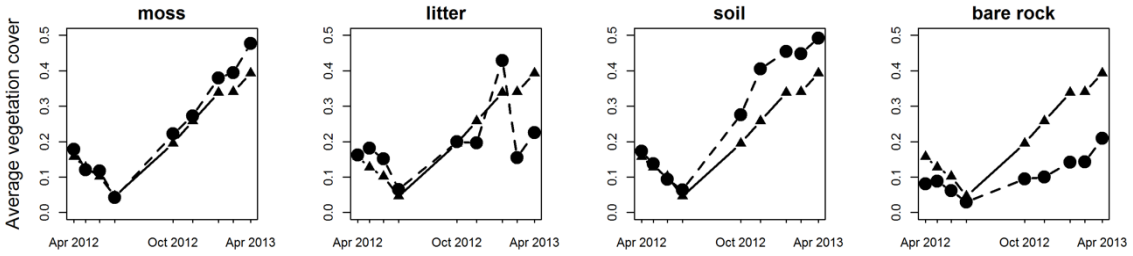


Fig. 2 Comparison of trends for average vegetation cover (solid line) and conditional vegetation cover (dashed lines) for each substrate

The aim of the model proposed in the following section is both to capture the marginal temporal trend for vegetation, basically due to climate, and to quantify the substrate-specific effect on vegetation, i.e., the ability of a substrate typology to favor/inhibit vegetation cover . Because of the compositional nature of the substrate data, such an ability can only be evaluated comparatively among substrates. In fact, if a substrate typology shows a higher-than-average vegetation cover, necessarily some other substrates will show lower-than-average cover; for this reason, substrate effects will be interpreted as "relative suitability" measures. Two main questions arise: How much do substrate typologies differ in terms of relative suitability? Does this difference vary over time? From a statistical point of view, substrate suitability estimates can be obtained as regression coefficient estimates of a regression model where vegetation cover is the response variable and substrate typology is aa compositional covariate. A spatio-temporal modelling framework is seek in this work for two reasons. First, spatial and temporal dependencies observed in the descriptive plots for vegetation cover need to be modelled for reliable estimates of the regression coefficients to be obtained (ref?). Second, it is of interest to model the temporal structure in the regression coefficients to predict suitability of all substrates at unobserved times and investigate their temporal variations across the study period.

3 MODEL

The relationship between substrate typology and vegetation cover is modeled by the hierarchical Bayesian model described in what follows; the strategy for handling compositional covariates is discussed in the next subsection. Let y_{st} denote the count of

185 pixels covered by vegetation at plot *s* and time *t*. Conditional on the vegetation
186 occupancy probability π_{st}, counts are assumed to follow a Binomial distribution, i.e.,
187 the binary response at each pixel within a plot is considered as the realization of an
188 independent trial. Spatio-temporal variability and dependence on covariates are
189 managed at higher levels of the hierarchy. The likelihood is specified as follows:
190
191 $$y_{st} \left| \pi_{st} \sim Binomial\left(\pi_{st}, n\right) \right. \qquad s = 1,...,S;\ t = 1,...,T$$
192
193 where *n* denotes the constant number of pixels belonging to each plot. The model can be
194 easily generalized to the case of a variable number of trials; nevertheless, having a
195 constant number of trials simplifies the algebraic representation of the model presented
196 in section 3.2 and in Appendix A. To model the vegetation occupancy probability as a
197 function of covariates, we adopt the *probit* link; the linear predictor is specified as
198 follows:
199
200 $$\Phi^{-1}\left(\pi_{st}\right) = \eta_{st} = \boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st} \qquad s = 1,...,S;\ t = 1,...,T \qquad (0)$$
201
202 where $\boldsymbol{x}_{st}$ is a P-dimensional covariate vector and $\boldsymbol{\beta}_t = \left(\beta_{1t},...,\beta_{pt},...,\beta_{Pt}\right)$ denotes the
203 vector of time-varying regression coefficients. Postulating time-varying coefficients
204 implies that time is an effect modifier; checking this assumption in our case study is
205 crucial because the aim is to evaluate the evolution (if any) of substrate suitability over
206 time. Three sets of random effects are included to account for spatio-temporal
207 variability not explained by the dependence on the covariates: $\boldsymbol{\alpha} = \left(\alpha_1,...,\alpha_t,...,\alpha_T\right)$ and

$\boldsymbol{\theta} = (\theta_1,...,\theta_s,...,\theta_S)$ are vectors of temporal and spatial main effects, respectively, whereas vector $\boldsymbol{\delta} = (\delta_{11},...,\delta_{st},...,\delta_{ST})$ contains space-time interactions.

Modeling of both regression coefficients and random effects is based on the theory concerning Intrinsic Gaussian Markov Random Fields (IGMRFs, see Rue and Held 2005, for a comprehensive review), which are very popular tools for spatio-temporal modeling of lattice data. Because of the Markov property, which implies the sparseness of the precision matrix, IGMRFs allow for extremely fast computations. IGMRFs have also been extended to cope with irregularly spaced time locations by Lindgren and Rue (2008); this is particularly useful in our case study, in which measurements have been collected at unequally spaced times, as often occurs in ecological studies.

In the following, $IGMRF_J(\tau \boldsymbol{K})$ denotes the (multivariate normal) distribution of a J-dimensional random vector with $J \times J$-dimensional structure matrix $\boldsymbol{K}$ and precision parameter τ; the IGMRFs are improper, i.e., $\boldsymbol{K}$ is rank-deficient. Structure matrix $\boldsymbol{K}$ describes conditional dependence relationships derived from assumptions concerning the spatial and temporal structure of the phenomenon.

Let $\boldsymbol{\beta}_p = (\beta_{p1},...,\beta_{pt},...,\beta_{pT})$ denote the T-dimensional vector of regression coefficients referring to the p-th covariate obtained by re-arranging the p-th elements of vectors $\boldsymbol{\beta}_t = (\beta_{1t},...,\beta_{pt},...,\beta_{Pt})$ for all $t = 1,...,T$. We introduce a temporal dependence among these coefficients, assuming a smooth variation of the covariate effect over the study period, by specifying the prior $\boldsymbol{\beta}_p \sim IGMRF_T(\tau_{\beta_p} \boldsymbol{K}_{\beta_p})$; $p = 1,...,P$. Following the modeling strategy first proposed in Knorr-Held (2000) in the context of spatio-temporal disease mapping, prior distributions for spatio-temporal random effects are specified as follows: $\boldsymbol{\alpha} \sim IGMRF_T(\tau_\alpha \boldsymbol{K}_\alpha)$, $\boldsymbol{\theta} \sim IGMRF_S(\tau_\theta \boldsymbol{K}_\theta)$ and $\boldsymbol{\delta} \sim IGMRF_{ST}(\tau_\delta \boldsymbol{K}_\delta)$.

Structure matrices $\boldsymbol{K}_{\beta_p}$ and $\boldsymbol{K}_{\alpha}$ need to be specified to capture a smooth temporal behavior. In this context, Random Walk (RW) priors (i.e., IGMRFs on the line) of order 1 is a popular choice; we set common temporal structure matrices $\boldsymbol{K}_{\beta_p} = \boldsymbol{K}_{\alpha} = \boldsymbol{K}_{RW}$. It follows that $rank\left(\boldsymbol{K}_{RW}\right) = T - 1$. Because our data are irregularly spaced over time, the structure matrix typically used for this process must be adapted. To this aim, Rue and Held (2005) propose a solution by considering the integration of the Wiener process (Wahba, 1978).

Structure matrix $\boldsymbol{K}_{\theta}$ is specified as the $S \times S$-dimensional Laplacian matrix of the graph induced by a first-order neighboring structure on the regular lattice formed by the plots. The *i*-th diagonal entry is equal to the number of neighbors of plot *i*, whereas the *ij*-th entry equals -1 *iff* plots *i* and *j* share a common boundary and zero otherwise. With this specification, $rank\left(\boldsymbol{K}_{\theta}\right) = S - 1$.

Regarding the spatio-temporal interaction term, four types of interactions are envisioned in Knorr-Held (2000), leading to different structure matrices $\boldsymbol{K}_{\delta}$ obtained as a Kronecker product of temporal and spatial structure matrices $\boldsymbol{K}_{\delta} = \boldsymbol{K}_{time} \otimes \boldsymbol{K}_{space}$. Specifying $\boldsymbol{K}_{\delta} = \boldsymbol{I}_T \otimes \boldsymbol{I}_S$ corresponds to a Type I interaction, in which independence among $\boldsymbol{\delta}$s is postulated; a Type II interaction, obtained by specifying the structure matrix as $\boldsymbol{K}_{\delta} = \boldsymbol{K}_{RW} \otimes \boldsymbol{I}_S$, is suitable if plots show different temporal trends with no structure in space; a Type III interaction, $\boldsymbol{K}_{\delta} = \boldsymbol{I}_T \otimes \boldsymbol{K}_{\theta}$, assumes different spatial trends with no structure in time. Finally, a Type IV interaction, $\boldsymbol{K}_{\delta} = \boldsymbol{K}_{RW} \otimes \boldsymbol{K}_{\theta}$, involves both a spatial and a temporal structure.

To ensure model identifiability, appropriate linear constraints need to be imposed on the spatial ($\boldsymbol{\theta}$) and spatio-temporal ($\boldsymbol{\delta}$) random effects, whereas the temporal random

effects ($\boldsymbol{\alpha}$) are unconstrained and can be considered time-dependent random intercepts. The number of linear constraints needed to ensure identifiability corresponds to the rank deficiency of the structure matrix of an IGMRF. As pointed out in Schrödle and Held (2011), identifiability can be ensured by computing the null space of the structure matrix and using the obtained eigenvectors as linear constraints. A simple sum-to-zero constraint is required for the identifiability of the spatial effects because $rank\left(\boldsymbol{K}_{\theta}\right) = S - 1$ (i.e., the null space of $\boldsymbol{K}_{\theta}$ is spanned by a constant vector). Because of the properties of the Kronecker product, the rank deficiency of $\boldsymbol{K}_{\delta}$ depends on the specified type of interaction: $rank\left(\boldsymbol{K}_{\delta}\right) = T \times S - rank\left(\boldsymbol{K}_{time}\right) \times rank\left(\boldsymbol{K}_{space}\right)$; when a Type IV interaction is considered, the number of required linear constraints becomes extremely high when S and T are large.

Model hierarchy is completed by prior specification for precision parameters of the IGMRFs: we assume independent, small-parameter Gamma distributions both to preserve non-informativeness and to take advantage of the conjugacy between the Gaussian and the Gamma distribution in building the MCMC sampler.

3.1 Managing compositional covariates

As discussed in section 2, the main aim of this study is to investigate the effect of substrate typology on vegetation cover. Because at each plot substrate typology is expressed as a proportion, these explanatory variables need to be managed coherently with compositional algebra. For the sake of simplicity, in this section, spatio-temporal subscripts are dropped, and the focus is on a simple linear regression model; nevertheless, the developed theory can be readily applied to model (1).

Consider a D-dimensional vector of compositional covariates, $\boldsymbol{z}=(z_1,...,z_d,...,z_D)$, that can be represented with unit-sum constraint in the simplex $S^D=\left\{\boldsymbol{z}=(z_1,...,z_d,...,z_D)',z_d>0,\sum_{d=1}^{D}z_d=1\right\}$. If untransformed compositional covariates are used, the design matrix of the proportional representation of compositions is singular due to the sum-to-one constraint. This issue can be addressed by a transformation that allows for a transition from the D-dimensional simplex to unconstrained $(D-1)$-dimensional real space, $\Re^{D-1}$. In this paper, we adopt the *ilr* transformation (Egozcue et al., 2003), which is an isometry between S^D and $\Re^{D-1}$, leading to *ilr*-transformed covariates $\boldsymbol{x}=ilr(\boldsymbol{z})=(x_1,...,x_i,...,x_{D-1})'$. Following Egoczue et al. (2003), an orthonormal basis matrix $\boldsymbol{B}$ associated with the coordinate system generated by $ilr(\boldsymbol{z})$ can be obtained by sequential binary partitioning. The i-th row of this matrix is

$$\boldsymbol{b}_i=\sqrt{\frac{D-i}{D-i+1}}\left[\underbrace{0,...,0}_{i-1\text{ times}},+1,\underbrace{-(D-i)^{-1},...,-(D-i)^{-1}}_{D-i\text{ times}}\right];\ \boldsymbol{b}_i\in\Re^D;i=1,...,D-1. \quad (0)$$

The i-th component of the *ilr* coordinates vector $\boldsymbol{x}$ is defined as

$$x_i=ilr(\boldsymbol{z})_i=\sqrt{\frac{D-i}{D-i+1}}\,ln\frac{z_i}{\sqrt[D-i]{\prod_{j=i+1}^{D}z_j}}\ \text{ for }\ i=1,...,D-1. \quad (0)$$

In the context of compositional data analysis, other widely used transformations include the centered log-ratio (*clr*) and the additive log-ratio (*alr*) transformations. The *alr*

299 transformation raises several issues in terms of interpretability because it is not an
300 isometric transformation. According to Tolosana-Delgado and Van Den Boogaart
301 (2013), the *clr* representation of a composition is often convenient in a regression
302 framework because each coefficient can be related to an original component.
303 Nonetheless, the *clr* is not expressed in terms of an orthogonal basis, neither of the
304 simplex nor of the real space and it generates a singular design matrix that requires a
305 sum-to-zero constraint for estimation of the regression coefficients. In a simple linear
306 regression framework, this does not raise particular issues, since estimation of the
307 regression coefficients is obtained by computing the Moore-Penrose inverse. This is not
308 practical in models where regression coefficients are modelled along time, as in our
309 application, since time-specific constraints would result in considerable computational
310 burden. Hence, our strategy for regression on compositional covariates consists in
311 estimating the model with *ilr* coordinates and in obtaining the *clr* coefficients by
312 exploiting the relationship betwee *ilr* and *clr* coordinates.
313 The *ilr* transformation is directly associated with orthogonal coordinates in the simplex
314 (Egozcue et al. 2003). The first component of the *ilr*-transformed vector is defined as a
315 function of the ratio between z_1 and the geometric mean of *all* of the remaining
316 components of $\boldsymbol{z}$. Thus, when *ilr*-transformed covariates, $\boldsymbol{X}$, are introduced into the
317 linear model $\boldsymbol{y} = \alpha + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, the first regression coefficient, β_1, is directly interpreted
318 as the *relative* effect of part 1; namely, it is the expected increment of the response
319 variable when the first part increases its relative weight with respect to an average of all
320 the other parts. The other regression coefficients cannot be interpreted analogously,
321 nonetheless, as will be discussed in the following, meaningful regression coefficients
322 for all parts can be obtained by switching from *ilr* to *clr* coefficients by a simple linear
323 transformation exploiting the relationship between *ilr* and *clr* coordinates.

The orthonormal basis B introduced in equation (XX) allows to switch from *ilr* to *clr* coordinates by the linear transformation $\tilde{\boldsymbol{x}} = clr(\boldsymbol{z}) = ilr(\boldsymbol{z})' \boldsymbol{B} = \boldsymbol{x}' \boldsymbol{B}$.

The model fit is not sensitive to the chosen transformation, i.e., the models $\boldsymbol{y} = \alpha + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ and $\boldsymbol{y} = \alpha + \tilde{\boldsymbol{X}}\tilde{\boldsymbol{\beta}} + \boldsymbol{\varepsilon}$, based respectively on *ilr* and *clr* transformed covariates, deliver the same fit and the same estimate of the intercept, in fact the two models have the same hat matrix:

$$\tilde{\boldsymbol{X}}\left(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}}\right)^{-}\tilde{\boldsymbol{X}}' = \boldsymbol{X}\boldsymbol{B}\left(\boldsymbol{B}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{B}\right)^{-}\boldsymbol{B}'\boldsymbol{X}' = \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'$$

since $\left(\boldsymbol{B}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{B}\right)^{-} = \boldsymbol{B}'\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{B}$ and $\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{I}$, where $\boldsymbol{A}^{-}$ denotes the Moore-penrose generalized inverse of $\boldsymbol{A}$. Eventually, *clr* coefficients can be obtained by means of the transformation $\tilde{\boldsymbol{\beta}} = \boldsymbol{B}\boldsymbol{\beta}$.

Vector $\tilde{\boldsymbol{\beta}}$ should be considered as the key parameter in the context of regression on compositional covariates. Indeed, the vector contains all the relevant information concerning the effect of each part on the response variable. A remarkable feature of $\tilde{\boldsymbol{\beta}}$ is that

$$\boldsymbol{1}_D'\tilde{\boldsymbol{\beta}} = \sum_{d=1}^{D}\tilde{\beta}_d = 0 \qquad (0)$$

holds by construction. Relationship (0) demonstrates an intuitive dependence structure characterizing $\tilde{\boldsymbol{\beta}}$ that is readily interpretable and particularly appealing in the context of regression on compositional covariates. For further illustration, consider an obvious

feature of compositional data: if the proportion of part *d* increases, then the proportion of at least one of the other parts must decrease because of the sum-to-one constraint. Thus, if increasing the share of the *d*-th part does have a positive effect on the response variable, then increasing the share of some other part will necessarily have a negative effect. For this reason, we interpret $\tilde{\boldsymbol{\beta}}$ coefficients as "relative effects" of a part w.r.t. all of the others.

Referring to model (1), the set of compositional covariates $\boldsymbol{z}_{st}$ at plot *s* and time *t* is included in the linear predictor after *ilr* transformation as $\boldsymbol{x}_{st} = ilr\left(\boldsymbol{z}_{st}\right)$, i.e., $\Phi^{-1}\left(\pi_{st}\right) = \boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}$; hence, model (1) contains $P = D - 1$ regressors.

3.2 Computational details

As is typical in complex hierarchical Bayesian models, the joint posterior distribution is not available in closed form and needs to be approximated. The Markov property of IGMRF models implies sparseness of the precision structure matrix, which allows fast computations, making spatio-temporal modeling with a large amount of lattice data feasible. Two alternative strategies are currently very popular for approximating the joint posterior distribution: MCMC sampling and Integrated Nested Laplace Approximations (INLA, Rue and Martino 2009). The latter is particularly suited for latent GMRF models and provides very accurate approximations of the posterior distribution. In addition, INLA definitely outperforms MCMC approaches in terms of computational time and for more precise approximations to be obtained. Thus, such a method appears to be preferable and should be used whenever possible. INLA has been made easily implementable by the R package INLA (Rue et al. 2013). The problem with

applying INLA to our case study is that, when a Type IV interaction is considered, the large number of linear constraints required for model identifiability makes the approach computationally unfeasible, causing computer crash.

For this reason, MCMC sampling is a necessary tool for the purpose of this paper. The MCMC algorithm has been implemented by adopting the popular auxiliary variable approach introduced in Albert and Chib (1993) and developed successively in several papers. This approach preserves full conditionals in GMRF form, which are lost in the case of non-normal likelihood: this is crucial for fast sampling. The approach based on data augmentation introduces, for each *Bernoulli* trial denoted by subscript *i*, in plot *s* at time *t*, the following auxiliary variables:

$$w_{ist} \sim N\left(\boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1\right) \qquad i = 1,...n;\ s = 1,...,S;\ t = 1,...,T$$

The Binomial likelihood is retrieved by specifying

$$\tilde{y}_{ist} = \begin{cases} 1 \text{ if } w_{ist} > 0 \\ 0 \text{ if } w_{ist} \le 0 \end{cases} \qquad \sum_i \tilde{y}_{ist} = y_{st}$$

where $\tilde{y}_{ist} = 1$ and $\tilde{y}_{ist} = 0$ appear y_{st} and $n - y_{st}$ times, respectively. It follows that the full conditional distributions of the auxiliary variables are truncated Normal:

$$w_{ist} \left| ... \sim \begin{cases} N\left(\boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1\right) I\left(w_{ist} > 0\right) \text{ if } \tilde{y}_{ist} = 1 \\ N\left(\boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1\right) I\left(w_{ist} \le 0\right) \text{ if } \tilde{y}_{ist} = 0 \end{cases} \right. \qquad (0)$$

Summing over the pixels, we obtain $w_{\bullet st} = \sum_{i=1}^{n} w_{ist}$. Conditionally on $w_{\bullet st}$, MCMC sampling can proceed as in a standard Bayesian linear model, where $w_{\bullet st}$ serves as pseudo-data. The linear predictor can be written in compact form as follows:

$$\boldsymbol{w}_{\bullet} = \boldsymbol{X}^{'}\boldsymbol{\beta} + \boldsymbol{\alpha} \otimes \boldsymbol{1}_S + \boldsymbol{1}_T \otimes \boldsymbol{\theta} + \boldsymbol{\delta}$$

where $\boldsymbol{1}_k$ denotes a k-dimensional unit vector, $\boldsymbol{X} = diag\left(\boldsymbol{X}_1,...,\boldsymbol{X}_t,...,\boldsymbol{X}_T\right)$ is the $ST \times T\left(D-1\right)$-dimensional block-diagonal matrix with $S \times \left(D-1\right)$-dimensional blocks $\boldsymbol{X}_t$ containing *ilr*-transformed covariates at time t and $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t,...,\boldsymbol{\beta}_T\right)$ is the $\left(D-1\right)T$-dimensional vector of time-varying coefficients.

In the context of the Bayesian linear regression model, implementing a Gibbs sampler involves standard MCMC tools; it is worth noting that sampling under linear constraints for the interaction term, even for a Type IV interaction model, is not an hard task. In fact, it can be shown that the null space spanned by the eigenvectors associated with null eigenvalues of the structure matrix $\boldsymbol{K}_{\delta}$ is equivalent to the null space spanned by the eigenvectors associated with null eigenvalues of the Kronecker product $\boldsymbol{C}_S \otimes \boldsymbol{C}_T$, where $\boldsymbol{C}_k$ denotes the k-dimensional centering matrix. Therefore, samples from the full conditional distribution of $\boldsymbol{\delta}$ can be obtained by centering samples over space and time on the fly at each MCMC iteration. Full conditional distributions for all of the parameters are reported in Appendix A.

Several further refinements to the Gibbs sampler built for our case study can be adopted to improve the efficiency of the MCMC algorithm. Regarding the auxiliary variable

approach, Frühwirth-Schnatter et al. (2007) propose an augmented model with lower dimensionality, where the number of required latent variables per observation becomes independent of the number of trials. Furthermore, block-sampling has been shown to considerably improve the mixing of Markov chains when some parameters show high correlation. Nonetheless, we adopt a Gibbs sampler in our case study because it allows for accurate results to be obtained within a fairly reasonable computational time.

4 APPLICATION

4.1 Model selection

Starting from model (1), several specifications can be proposed, according to different assumptions regarding the effect of substrate typology and the structured spatio-temporal effects. Namely, we estimate and compare a total of 10 competing models: all of them include a pure temporal effect (i.e., α_t) and a pure spatial effect (i.e., θ_s). The models differ with respect to the spatio-temporal interaction type (I,II,III,IV); moreover substrate effect is specified either as constant or time-varying. All temporal effects are modeled as RW1 opportunely adapted for considering unequally spaced times.

The models are compared in terms of the Deviance Information Criterion (DIC, Spiegelhalter et al. 2002) in Table 3.

	Interaction				
Substrate effect	**Type 1**	**Type 2**	**Type 3**	**Type 4**	**No**
Constant	33501	33518	33349	33510	188053
Time-varying	33496	33513	**33343**	33515	183137

Table 3 Deviance Information Criterion (DIC) corresponding to different model specifications for spatio-temporal interaction and substrate effect (time-constant or time-varying)

From the last column of this table, it can be observed that neglecting spatio-temporal interactions produces very poor performance. The selected model (whose DIC is reported in bold in the table) includes time-varying coefficients and a Type 3 interaction. The inclusion of a Type 4 interaction does not improve the model fit, regardless of the assumption made about substrate effects. Despite the small differences in terms of DIC exhibited by models with a Type 3 interaction, the results presented in the following section provide evidence of a significant change in the covariate effect over time, supporting the selection of the time-varying coefficient model.

Ultimately, the selected model is:

$$y_{st}\,|\,\pi_{st} \sim Binomial\left(\pi_{st}, n\right) \qquad s = 1,...,S;\ t = 1,...,T$$

$$\Phi^{-1}\left(\pi_{st}\right) = \boldsymbol{x}_{st}^{'}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st} \qquad \text{where} \quad \boldsymbol{\beta}_t = \left(\beta_{1t},...,\beta_{dt},...,\beta_{(D-1)t}\right)$$

$$\boldsymbol{\beta}_d = \left(\beta_{d1},...,\beta_{dt},...,\beta_{dT}\right) \sim IGMRF_T\left(\tau_{\beta_d}\boldsymbol{K}_{RW}\right);\ d = 1,...,D-1$$

$$\boldsymbol{\alpha} \sim IGMRF_T\left(\tau_\alpha \boldsymbol{K}_{RW}\right),\ \boldsymbol{\theta} \sim IGMRF_S\left(\tau_\theta \boldsymbol{K}_\theta\right);\ \boldsymbol{\delta} \sim IGMRF_{ST}\left(\tau_\delta \boldsymbol{K}_\delta\right)$$

where $\boldsymbol{K}_\delta = \boldsymbol{I}_T \otimes \boldsymbol{K}_\theta$, in accordance with a Type III interaction. Because, as mentioned in section 3, vector $\boldsymbol{\alpha}$ is unconstrained in model estimation, this purely temporal effect is devoted to capture, in the *probit* scale, the natural vegetation life cycle.

4.2 Results

Results are obtained from a post-convergence MCMC sample of size 10,000 obtained by thinning a 100,000 sample to reduce correlation while guaranteeing an adequate effective sample size.

463 In the upper panels (a1-a3) of Figure 3, posterior distributions ($\boldsymbol{\beta}_t \,|\, \boldsymbol{y}$) of parameters
464 $\boldsymbol{\beta}_t = \left(\beta_{1t}, ..., \beta_{dt}, ..., \beta_{(D-1)t}\right)$ are reported for $t = 9$ as an example. The lower panels (b1-
465 b4) report posterior distributions of *clr*-regression coefficients $\tilde{\boldsymbol{\beta}}_t$; these posterior
466 distributions are obtained by transforming $\boldsymbol{\beta}_t$ at each MCMC iteration. Panels (a1) and
467 (b1) contain the same information, up to a scaling constant relating $\beta_{1,9}$ and $\tilde{\beta}_{1,9}$. The
468 positiveness of this coefficient implies that the first substrate typology (moss)
469 encourages vegetation cover when compared with other substrates typologies, i.e., moss
470 shows positive "relative suitability". Panels (a2) and (a3) contain, respectively, the
471 posterior distributions of $\beta_{2,9}$ and $\beta_{3,9}$: these parameters do not convey readily
472 interpretable information concerning substrate suitability because, as discussed in
473 section 3.1, they do not refer to any direct comparison between one compositional part
474 and all others. As a consequence, testing for the nullity of these coefficients is pointless
475 because a null coefficient does not directly imply a non-significant effect of some
476 compositional part on the response variable. As a matter of fact, no sound
477 considerations about the relative suitability of litter, soil and bare rock can be drawn
478 from panels (a2) and (a3). On the other hand, panels (b2)-(b4) convey direct
479 information about substrate relative suitability: the posterior distribution of regression
480 coefficient associated with litter (panel b2) is centered at zero, meaning that litter shows
481 average suitability, whereas soil (panel b3) and bare rock (panel b4) have positive and
482 negative relative suitability, respectively.
483

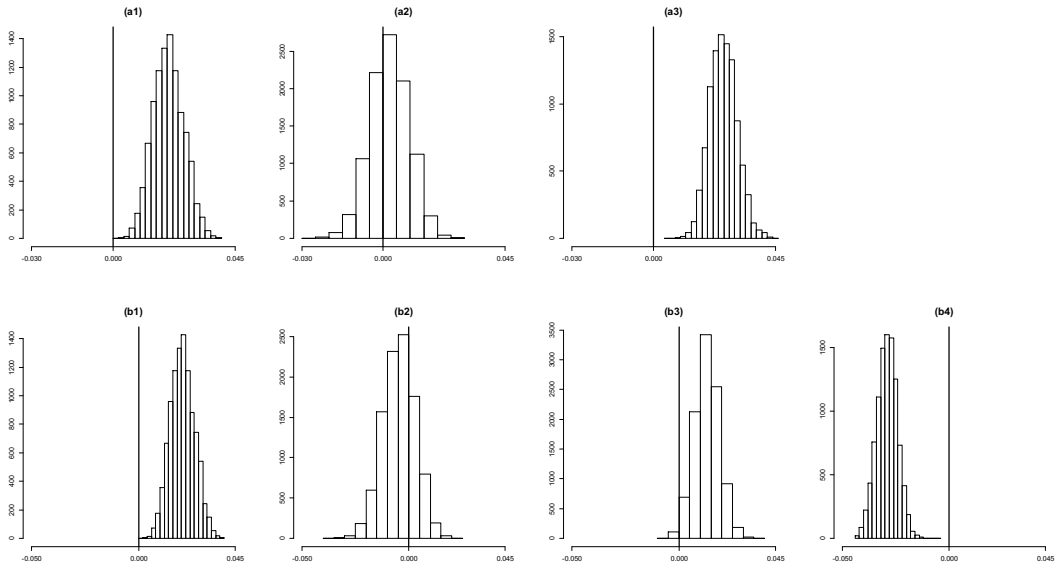
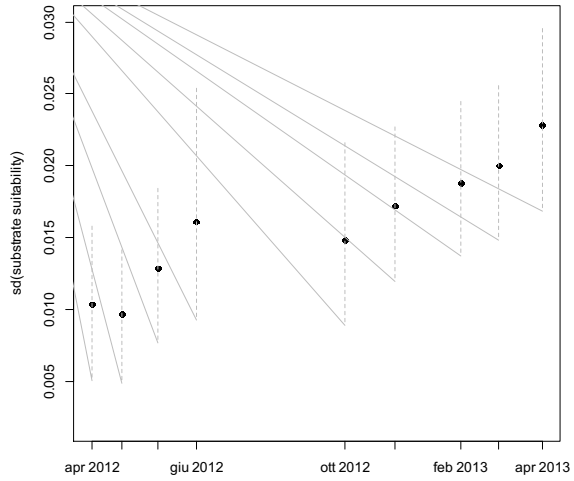


Fig. 3 Posterior distributions of the regression coefficients for the last collection time (panels a1-a3). Posterior distributions of transformed regression coefficients expressing the relative suitability for moss (b1), litter (b2), soil (b3) and bare rock (b4). Vertical lines correspond to zero

Because transformed coefficients are directly comparable, we provide some suggestions on how posterior estimates can be used to describe findings about substrate suitability. Upon first glance, at a given time, the substrates can be ordered from the most to the least suitable for vegetation cover based on the posterior means $E\left(\tilde{\beta}_{dt}\middle|\boldsymbol{y}\right)$: at time 9, moss turns out to be the most suitable substrate typology, followed by soil, litter and bare rock. Moreover, a rigorous comparison between the suitability of substrates d and d' that takes account of the uncertainty about parameter estimates can be obtained by evaluating $Pr\left(\tilde{\beta}_{dt} > \tilde{\beta}_{d't}\middle|\boldsymbol{y}\right)$, i.e., the posterior probability that the suitability of substrate d is higher than that of substrate d'. This task is easily addressed by exploiting the MCMC sample. It is observed that, at time 9, the posterior probabilities that moss is more suitable than litter, soil and bare rock are 0.98, 0.81 and 1, respectively.

Furthermore, it is interesting to determine whether the heterogeneity characterizing substrate suitability changes over time; it is worth noting that the more the estimated

503 coefficients at time *t* are dispersed around zero, the more the substrates show differences
504 in terms of suitability. Thus, the posterior distribution $sd\left(\tilde{\boldsymbol{\beta}}_t\right)\big|\boldsymbol{y}$, t=1,…,9, gives
505 information about the heterogeneity of substrates in terms of the suitability at each time.
506 Figure 4 shows that the heterogeneity increases over the study period.
507



508

509 **Fig. 4** Posterior distribution of $sd\left(\tilde{\boldsymbol{\beta}}_t\right)\big|\boldsymbol{y}$ over time

510
511 The biological motivation of this behavior can be found in the increase in vegetation
512 cover, from October to the end of the study period, caused by a higher propensity of
513 vegetation to occur in the substrates that exhibit greater availability of nutrient resources
514 (i.e., moss and soil) than do the others (litter and bare rock).
515 Finally, we describe findings about the temporal evolution of substrate relative
516 suitability that are helpful in addressing the ecological questions posed at the end of
517 section 2. In each panel of Figure 5, the posterior mean of $\tilde{\boldsymbol{\beta}}_t$ at the collection times
518 (filled circles), the predictions at unobserved times (black solid line) and credible bands

(grey dashed lines) are reported. Predictions are obtained via MCMC sampling from the posterior predictive distribution of $\tilde{\boldsymbol{\beta}}_t$.

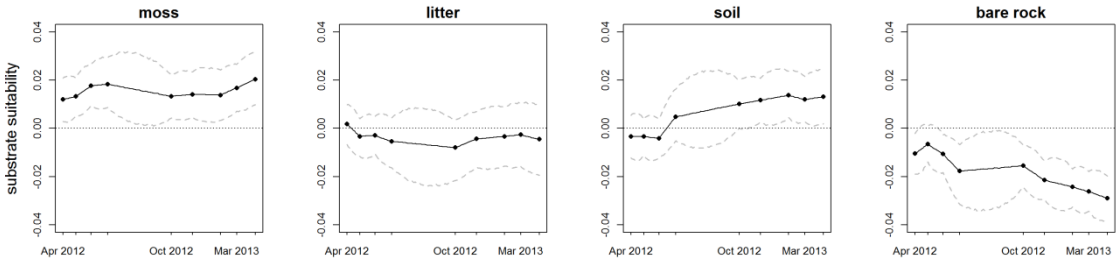


Fig. 5 Substrate relative suitability over time

Moss is revealed to be the most suitable substrate at the beginning of the study period, and its relative suitability remains nearly constant over time. On the other hand, the relative suitability of soil is approximately zero (i.e., approximately average) at the beginning and increases over time, reaching a positive relative suitability level in winter. The opposite behavior is exhibited by bare rock, which is less suitable than average and shows a decreasing pattern over time. Finally, litter presents constant average suitability over time.

It is worth noting that, in contrast to the results shown in Figure 2, where the temporal pattern of substrate-specific effects is confounded with the marginal trend for vegetation, the results shown in Figure 5 provide direct information on the relative suitability level of each substrate and how these levels vary over the study period, after properly accounting for spatial and temporal correlation. These results are in agreement with the ecological literature on this type of habitat (Lundholm 2003, Otsus and Zobel 2002) as stated in Velli (2014).

5. CONCLUSIONS

In this work, a spatio-temporal model for ecological data derived from digital ground photos is presented. The primary focus is on quantifying the relation between vegetation cover and substrate typology, expressed as compositional data, by accounting for spatial and temporal correlation. The proposed models include spatial, temporal and spatio-temporal effects modeled as IGMRFs. A novel approach for estimating the relative effect of compositional covariates on the response variable is developed within a Bayesian framework, through a procedure based on appropriate transformations of the regression coefficients. This procedure exploits the properties of the *ilr* transformation, which admits an orthonormal basis representation; this transformation allows for a transition between different coordinate systems, providing meaningful regression coefficients. In our opinion, the proposed approach constitutes a basis for future developments in such fields as non-linear and nonparametric regression for compositional covariates. To the best of our knowledge, these topics have been scarcely debated in the literature concerning both compositional data and Bayesian analysis.

The lattice structure of the data make GMRFs particularly suitable tools for modeling spatial and temporal dependence by specifying a conditional dependence structure. This ability allows for feasible computations that can be implemented both by the INLA approach and MCMC sampling; fast algorithms that allow for efficient sampling by exploiting the sparseness of the precision matrices are implemented within a Gibbs sampling scheme. Both the INLA and MCMC codes are available upon request from the corresponding author.

The model is applied to a dataset concerning a habitat included in the framework of the priority defined by the European Commission, where investigating the relationship

between vegetation cover and substrate typology can be helpful in quantifying substrate suitability. The case study demonstrates that the proposed modeling strategy produces valuable outcomes for ecologists. First, in ecological studies, knowledge of substrate suitability is strategic for habitat conservation, for instance, to predict future developments and reactions to possible environmental changes. Second, a modeling approach that is able to predict substrate suitability or other ecological responses over un-monitored periods of time is extremely useful in many common situations in which surveys cannot follow a regular plan, as in the case in which fieldwork activity is impossible because of adverse environmental and weather conditions.

One interesting ecological finding of the paper is that substrate relative suitability varies over time. A future line of research will focus on extending the proposed model to allow for smooth, rather than simply linear, effects of compositional covariates on ecological responses. Non-linear effects of substrate shares on ecological responses may often be observed, especially when the response of interest is a biodiversity index, such as species richness or evenness.

APPENDIX

The Gibbs sampler is implemented as follows:

1. Sample auxiliary variables from the truncated normal full conditionals (0).

Let $\tilde{\boldsymbol{W}}$ be the nST-dimensional vector of the auxiliary variables, full conditionals for the other parameters can be obtained as in the standard linear model with Gaussian likelihood, starting from the following expression of the linear predictor:

591 $$\tilde{\boldsymbol{\eta}} = \left(\boldsymbol{X} \otimes \boldsymbol{1}_n\right)\boldsymbol{\beta} + \left(\boldsymbol{I}_T \otimes \boldsymbol{1}_S \otimes \boldsymbol{1}_n\right)\boldsymbol{\alpha} + \left(\boldsymbol{1}_T \otimes \boldsymbol{I}_S \otimes \boldsymbol{1}_n\right)\boldsymbol{\theta} + \left(\boldsymbol{I}_T \otimes \boldsymbol{I}_S \otimes \boldsymbol{1}_n\right)\boldsymbol{\delta}$$
$$= \tilde{\boldsymbol{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{A}}_1\boldsymbol{\alpha} + \tilde{\boldsymbol{A}}_2\boldsymbol{\theta} + \tilde{\boldsymbol{A}}_3\boldsymbol{\delta}$$

592

593 In what follows, the multivariate normal distribution is expressed in the canonical form

594 (Rue and Held, 2005; page 27)

595 2. Sample the regression parameters from the distribution

596 $$\boldsymbol{\beta}\left|\ldots \sim N_C\left(\boldsymbol{b}_\beta, \boldsymbol{Q}_\beta\right)\right.$$

597 $$\boldsymbol{b}_\beta = \tilde{\boldsymbol{X}}'\left(\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{A}}_1\boldsymbol{\alpha} - \tilde{\boldsymbol{A}}_2\boldsymbol{\theta} - \tilde{\boldsymbol{A}}_3\boldsymbol{\delta}\right) = n\tilde{\boldsymbol{X}}'\left(\boldsymbol{w}_\bullet/n - \left(\boldsymbol{I}_T \otimes \boldsymbol{1}_S\right)\boldsymbol{\alpha} - \left(\boldsymbol{1}_T \otimes \boldsymbol{I}_S\right)\boldsymbol{\theta} - \boldsymbol{\delta}\right);$$

598 $$\boldsymbol{Q}_\beta = \boldsymbol{P}_0\boldsymbol{\Gamma}_\beta\boldsymbol{P}_0' + \tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}} = \boldsymbol{P}_0\boldsymbol{\Gamma}_\beta\boldsymbol{P}_0' + n\boldsymbol{X}'\boldsymbol{X}$$

599 where $\boldsymbol{\Gamma}_\beta = diag\left(\tau_{\beta_1}, \ldots, \tau_{\beta_{D-1}}\right) \otimes \boldsymbol{K}_{RW}$ and $\boldsymbol{P}_0$ is a permutation matrix with the following

600 structure:

601

602 $$\boldsymbol{P}_0 = \begin{bmatrix} \boldsymbol{I}_{D-1} \otimes \boldsymbol{v}_1' \\ \ldots \\ \boldsymbol{I}_{D-1} \otimes \boldsymbol{v}_t' \\ \ldots \\ \boldsymbol{I}_{D-1} \otimes \boldsymbol{v}_T' \end{bmatrix}$$

603 The T-dimensional vector $\boldsymbol{v}_t$ contains 1 at position t and zero elsewhere.

604 3. Sample the temporal effects, i.e., the time-varying intercepts, from the distribution

605 $$\boldsymbol{\alpha}\left|\ldots \sim N_C\left(\boldsymbol{b}_\alpha, \boldsymbol{Q}_\alpha\right)\right.$$

606 $$\boldsymbol{b}_\alpha = \tilde{\boldsymbol{A}}_1'\left(\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta} - \tilde{\boldsymbol{A}}_2\boldsymbol{\theta} - \tilde{\boldsymbol{A}}_3\boldsymbol{\delta}\right) = n\tilde{\boldsymbol{A}}_1'\left(\boldsymbol{w}_\bullet/n - \boldsymbol{X}\boldsymbol{\beta} - \left(\boldsymbol{1}_T \otimes \boldsymbol{I}_S\right)\boldsymbol{\theta} - \boldsymbol{\delta}\right);$$

607 $$\boldsymbol{Q}_\alpha = \tau_\alpha\boldsymbol{K}_{RW} + \tilde{\boldsymbol{A}}_1'\tilde{\boldsymbol{A}}_1 = \tau_\alpha\boldsymbol{K}_{RW} + nS\boldsymbol{I}_T$$

608 4. Sample the spatial effect from the distribution

609 $$\boldsymbol{\theta}\left|\ldots \sim N_C\left(\boldsymbol{b}_\theta, \boldsymbol{Q}_\theta\right)\right.$$

610 $$\boldsymbol{b}_\theta = \tilde{\boldsymbol{A}}_2^{'}\left(\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta} - \tilde{\boldsymbol{A}}_1\boldsymbol{\alpha} - \tilde{\boldsymbol{A}}_3\boldsymbol{\delta}\right) = n\left(\boldsymbol{1}_T \otimes \boldsymbol{I}_S\right)^{'}\left(\boldsymbol{w}_\bullet/n - \boldsymbol{X}\boldsymbol{\beta} - \left(\boldsymbol{I}_T \otimes \boldsymbol{1}_S\right)\boldsymbol{\alpha} - \boldsymbol{\delta}\right);$$

611 $$\boldsymbol{Q}_\theta = \tau_\theta \boldsymbol{K}_\theta + \tilde{\boldsymbol{A}}_2^{'}\tilde{\boldsymbol{A}}_2 = \tau_\theta \boldsymbol{K}_\theta + nT\boldsymbol{I}_S$$

612 Centre the sampled values.

613 $$\tilde{\boldsymbol{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{A}}_1\boldsymbol{\alpha} + \tilde{\boldsymbol{A}}_2\boldsymbol{\theta} + \tilde{\boldsymbol{A}}_3\boldsymbol{\delta}$$

614 5. Sample the spatiotemporal random effects from the distribution

615 $$\boldsymbol{\delta}|... \sim N_C\left(\boldsymbol{b}_\delta, \boldsymbol{Q}_\delta\right)$$

616 $$\boldsymbol{b}_\delta = \left(\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta} - \tilde{\boldsymbol{A}}_1\boldsymbol{\alpha} - \tilde{\boldsymbol{A}}_2\boldsymbol{\theta}\right) = n\left(\boldsymbol{w}_\bullet/n - \boldsymbol{X}\boldsymbol{\beta} - \left(\boldsymbol{I}_T \otimes \boldsymbol{1}_S\right)\boldsymbol{\alpha} - \left(\boldsymbol{1}_T \otimes \boldsymbol{I}_S\right)\boldsymbol{\theta}\right);$$

617 $$\boldsymbol{Q}_\delta = \tau_\delta \boldsymbol{K}_\delta + \tilde{\boldsymbol{A}}_3^{'}\tilde{\boldsymbol{A}}_3 = \tau_\delta \boldsymbol{K}_\delta + n\boldsymbol{I}_{TS}$$

618 Centre the sampled values either row wise, column wise or both depending on the type

619 of interaction.

620 6. Sample the precision parameters $\tau_{\beta_1},...,\tau_{\beta_{D-1}},\tau_\alpha,\tau_\theta,\tau_\delta$. The full conditionals for these

621 parameters are Gamma, for instance:

622 $$\tau_\theta|... \sim G\left(a + .5 \times rank\left(\boldsymbol{K}_\theta\right), b + .5 \times \boldsymbol{\theta}^{'}\boldsymbol{K}_\theta\boldsymbol{\theta}\right)$$

623

624

625 ## ACKNOWLEDGEMENTS

626 We wish to thank Carlo Ferrari, Giovanna Pezzi and Andrea Velli for introducing us to the problem,
627 providing data and performing data pre-processing.
628 The research work underlying this paper was funded by a FIRB 2012 grant (project no. RBFR12URQJ;
629 title: Statistical modeling of environmental phenomena: pollution, meteorology, health and their
630 interactions) for research projects by the Italian Ministry of Education, Universities and Research.
631
632
633 ## REFERENCES

634 Aitchison J (1986) The Statistical Analysis of Compositional Data. Chapman & Hall,
635 London
636
637 Albert JH, Chib S (1993) Bayesian Analysis of Binary and Polychotomous Response
638 Data. J Am Stat Assoc 88:669-679
639

640 Bennett LT, Judd TS, Adams MA (2000) Close-range vertical photography for
641 measuring cover changes in perennial grasslands. J Range Manage 53:634–641
642
643 Booth DT, Cox SE, Fifield C, Phillips M, Williamson N (2005) Image analysis
644 compared with other methods for measuring ground cover. Arid Land Res Manag
645 19:91–100
646
647 Booth DT, Tueller PT (2003) Rangeland monitoring using remote sensing. Arid Land
648 Res Manag 17:455–478
649
650 Council Directive (92/43/CEE) on the conservation of natural habitats and of wild fauna
651 and flora (http://ec.europa.eu/environment/nature/legislation/habitatsdirective/index_en.htm)
652
653 Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C (2003) Isometric
654 logratio transformations for compositional data analysis. Math Geol 35: 279–300
655
656 Egozcue JJ, Pawlowsky-Glahn V (2006) Simplicial geometry for compositional data. In
657 Buccianti et al. (eds) Compositional data analysis in the Geosciences: From theory to
658 practice. Geological Society, London, Special Publications, 264, pp 145-159
659
660 Fortin MJ, James PMA, MacKenzie A, Melles SJ, Rayfield B (2012) Spatial statistics,
661 spatial regression, and graph theory in ecology. Spatial Statistics 1:100-109
662
663 Frühwirth-Schnatter S, Frühwirth R, Held L, Rue H (2009) Improved auxiliary mixture
664 sampling for hierarchical models of non-Gaussian data, Stat Comput 19:479-492
665
666 Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology.
667 Ecol Model 135:147–186
668
669 Hron K, Filzmoser P, Thompson K (2012) Linear regression with compositional
670 explanatory variables. J Appl Stat 39:1115–1128
671
672 Kneib T, Muller J, Hothorn T (2008) Spatial smoothing techniques for the assessment
673 of habitat suitability. Environ Ecol Stat 15: 343-364
674
675 Knorr-Held L (2000) Bayesian modelling of inseparable space-time variation in disease
676 risk. Stat Med 19:2555–2567
677
678 Kuntz KL, Larson DW (2006) Microtopographic control of vascular plant, bryophyte
679 and lichen communities on cliff faces. Plant Ecol 185:239–253
680
681 Latimer AM, Wu S, Gelfand AE, Silander JA Jr (2006) Building Statistical Models to
682 Analyze Species Distributions. Ecol Appl 16:33-50
683
684 Lindgren F, Rue H (2008) On the Second-Order Random Walk Model for Irregular
685 Locations. Scand J Stat 35:691–700
686
687 Lundholm JT, Larson DW (2003) Relationships between spatial environmental
688 heterogeneity and plant species diversity on a limestone pavement. Ecography 6:715-
689 722

690
691 Pueyo Y, Alados CL (2007) Abiotic factors determining vegetation patterns in a semi-
692 arid Mediterranean landscape: Different responses on gypsum and non-gypsum
693 substrates. J Arid Environ 69: 490–505
694
695 Otsus M, Zobel M (2002) Small-scale turnover in a calcareous grassland, its pattern and
696 components. J Veg Sci 13:199-206
697
698 Richardson MD, Karcher DE, Purcell LC (2001) Quantifying turfgrass cover using
699 digital image analysis. Crop Sci 41:1884–1888.
700
701 Rue H, Held L (2005) Gaussian Markov random fields: theory and applications.
702 Chapman & Hall/CRC Press, London.
703
704 Rue H, Martino S (2009) Approximate Bayesian inference for latent Gaussian models
705 by using integrated nested Laplace approximations (with discussion). J Roy Stat Soc B
706 71:319-392
707
708 Rue H, Martino S, Lindgren F, Simpson D, Riebler A (2013) INLA: Functions which
709 allow to perform full Bayesian analysis of latent Gaussian models using Integrated
710 Nested Laplace Approximaxion. R package version 0.0-1383402327
711
712 Schrödle B, Held L (2011) Spatio-temporal disease mapping using INLA.
713 Environmetrics 22:725–734
714
715 Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A (2002) Bayesian Measures of
716 Model Complexity and Fit (with Discussion). J Roy Stat Soc B 64:583-616
717
718 Velli A. (2014) Relationships between plant diversity and environmental heterogeneity
719 in rupicolous grasslands on gypsum. The case study of Alysso-Sedion albi (Habitat
720 6110). Dissertation, University of Bologna
721
722 Wahba G (1978) Improper priors, spline smoothing and the problem of guarding against
723 model errors in regression. J Roy Stat Soc B 40:364–372
724
725 Wikle CK (2003) Hierarchical Bayesian models for predicting the spread of ecological
726 processes. Ecology 84:1382-1394.