



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE  
DELLA RICERCA

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

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

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Spatio-temporal regression on compositional covariates: modeling vegetation in a gypsum outcrop / Bruno, F; Greco, F; Ventrucci, M. - In: ENVIRONMENTAL AND ECOLOGICAL STATISTICS. - ISSN 1352-8505. - STAMPA. - 22:3(2015), pp. 445-463. [10.1007/s10651-014-0305-4]

*Availability:*

This version is available at: <https://hdl.handle.net/11585/372118> since: 2016-11-30

*Published:*

DOI: <http://doi.org/10.1007/s10651-014-0305-4>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

**Bruno, F., Greco, F. & Ventrucci, M. Spatio-temporal regression on compositional covariates: modeling vegetation in a gypsum outcrop. Environ Ecol Stat 22, 445–463 (2015). <https://doi.org/10.1007/s10651-014-0305-4>**

The final published version is available online at: <https://doi.org/10.1007/s10651-014-0305-4>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

1 SPATIO-TEMPORAL REGRESSION ON COMPOSITIONAL COVARIATES:

2 MODELING VEGETATION IN A GYPSUM OUTCROP

3  
4 **Francesca Bruno, Fedele Greco, Massimo Ventrucci**

5 Dipartimento di Scienze Statistiche “P. Fortunati”,

6 University of Bologna, Italy

7  
8 SUMMARY

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

22  
23 **KEYWORDS:** binomial data; compositional covariates; Intrinsic Gaussian Markov  
24 Random Fields; hierarchical Bayesian model; vegetation cover; suitability

27

28

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

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

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

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

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

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

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

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

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

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

88

89

90

## 2 MOTIVATING EXAMPLE

91

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

99 Data were collected over a study period going from April 2012 to late March 2013.

100 Sampling campaigns were adjourned during the dryness period of August and

101 September, when low vegetation cover is expected, and on January because of snow

102 cover. A total of nine sampling campaigns were conducted at unequally spaced times,

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

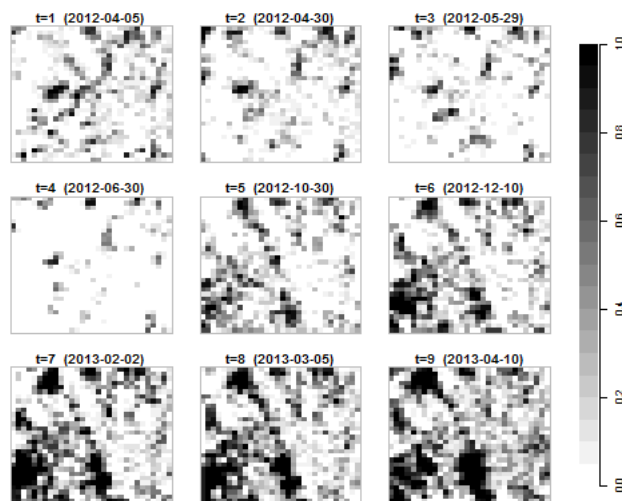
116 Some descriptive statistics concerning pixel-specific vegetation cover (expressed as  
117  $y_{st}/n$ ) are summarized in Table 1, where means and quartiles over the study area are  
118 reported for each time.

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

120 **Table 1** Descriptive statistics of vegetation cover ( $y_{st}/n$ ) over time  
121

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

124 increases in terms of both the number of pixels where vegetation is present (i.e., Q1 is  
 125 greater than zero only for  $t \geq 7$ ) and average coverage. Last row in Table 1 reports the  
 126 Moran's I statistics measuring spatial correlation. Moran's I index (Moran 1948)  
 127 measures the degree of correlation between observations at neighboring plots. This  
 128 index takes continuous values from -1 (maximum negative spatial correlation) to 1  
 129 (maximum positive spatial correlation) with 0 indicating absence of spatial correlation.  
 130 A spatio-temporal representation of the vegetation cover is reported by maps in Figure  
 131 1. The spatial correlation characterizing the vegetation cover measured at the pixel level  
 132 is considerable throughout the entire study period and becomes more evident from the  
 133 fifth to the last time point, as evidenced by the Moran's *I* values in Table 1. Overall, the  
 134 data show both spatial and temporal correlation.  
 135



136  
 137 **Fig. 1** Maps of vegetation cover at each collection time (indicated within brackets)  
 138  
 139



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

142

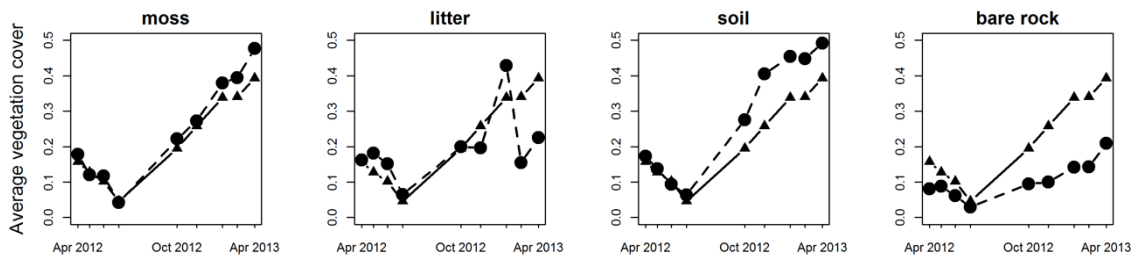
<i>T</i>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
<b>Moss</b>	0.46	0.30	0.35	0.35	0.42	0.50	0.57	0.54	0.49
<b>Litter</b>	0.29	0.15	0.14	0.19	0.06	0.01	0.03	0.01	0.01
<b>Soil</b>	0.09	0.30	0.26	0.14	0.22	0.23	0.17	0.20	0.18
<b>Bare rock</b>	0.16	0.25	0.25	0.32	0.30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

143  
 144

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

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

155



156  
 157  
 158  
 159

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

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

178

179

180

### 3 MODEL

181

182 The relationship between substrate typology and vegetation cover is modeled by the  
183 hierarchical Bayesian model described in what follows; the strategy for handling  
184 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  $\pi_{st}$ , 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

$$191 \quad y_{st} | \pi_{st} \sim \text{Binomial}(\pi_{st}, n) \quad s = 1, \dots, S; t = 1, \dots, T$$

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

$$200 \quad \Phi^{-1}(\pi_{st}) = \eta_{st} = \mathbf{x}'_{st} \boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st} \quad s = 1, \dots, S; t = 1, \dots, T \quad (0)$$

201

202 where  $\mathbf{x}_{st}$  is a  $P$ -dimensional covariate vector and  $\boldsymbol{\beta}_t = (\beta_{1t}, \dots, \beta_{pt}, \dots, \beta_{Pt})$  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:  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_t, \dots, \alpha_T)$  and

208  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s, \dots, \theta_S)$  are vectors of temporal and spatial main effects, respectively,

209 whereas vector  $\boldsymbol{\delta} = (\delta_{11}, \dots, \delta_{st}, \dots, \delta_{ST})$  contains space-time interactions.

210 Modeling of both regression coefficients and random effects is based on the theory  
211 concerning Intrinsic Gaussian Markov Random Fields (IGMRFs, see Rue and Held  
212 2005, for a comprehensive review), which are very popular tools for spatio-temporal  
213 modeling of lattice data. Because of the Markov property, which implies the sparseness  
214 of the precision matrix, IGMRFs allow for extremely fast computations. IGMRFs have  
215 also been extended to cope with irregularly spaced time locations by Lindgren and Rue  
216 (2008); this is particularly useful in our case study, in which measurements have been  
217 collected at unequally spaced times, as often occurs in ecological studies.

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

223 Let  $\boldsymbol{\beta}_p = (\beta_{p1}, \dots, \beta_{pt}, \dots, \beta_{pT})$  denote the  $T$ -dimensional vector of regression coefficients  
224 referring to the  $p$ -th covariate obtained by re-arranging the  $p$ -th elements of vectors

225  $\boldsymbol{\beta}_t = (\beta_{1t}, \dots, \beta_{pt}, \dots, \beta_{Pt})$  for all  $t = 1, \dots, T$ . We introduce a temporal dependence among  
226 these coefficients, assuming a smooth variation of the covariate effect over the study

227 period, by specifying the prior  $\boldsymbol{\beta}_p \sim IGMRF_T(\tau_{\beta_p} \mathbf{K}_{\beta_p})$ ;  $p = 1, \dots, P$ . Following the  
228 modeling strategy first proposed in Knorr-Held (2000) in the context of spatio-temporal  
229 disease mapping, prior distributions for spatio-temporal random effects are specified as  
230 follows:  $\boldsymbol{\alpha} \sim IGMRF_T(\tau_{\alpha} \mathbf{K}_{\alpha})$ ,  $\boldsymbol{\theta} \sim IGMRF_S(\tau_{\theta} \mathbf{K}_{\theta})$  and  $\boldsymbol{\delta} \sim IGMRF_{ST}(\tau_{\delta} \mathbf{K}_{\delta})$ .

231 Structure matrices  $\mathbf{K}_{\beta_p}$  and  $\mathbf{K}_{\alpha}$  need to be specified to capture a smooth temporal  
 232 behavior. In this context, Random Walk (RW) priors (i.e., IGMRFs on the line) of order  
 233 1 is a popular choice; we set common temporal structure matrices  $\mathbf{K}_{\beta_p} = \mathbf{K}_{\alpha} = \mathbf{K}_{RW}$ . It  
 234 follows that  $\text{rank}(\mathbf{K}_{RW}) = T - 1$ . Because our data are irregularly spaced over time, the  
 235 structure matrix typically used for this process must be adapted. To this aim, Rue and  
 236 Held (2005) propose a solution by considering the integration of the Wiener process  
 237 (Wahba, 1978).

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

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

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

254 effects ( $\alpha$ ) are unconstrained and can be considered time-dependent random intercepts.  
255 The number of linear constraints needed to ensure identifiability corresponds to the rank  
256 deficiency of the structure matrix of an IGMRF. As pointed out in Schrödle and Held  
257 (2011), identifiability can be ensured by computing the null space of the structure  
258 matrix and using the obtained eigenvectors as linear constraints. A simple sum-to-zero  
259 constraint is required for the identifiability of the spatial effects because  
260  $rank(\mathbf{K}_\theta) = S - 1$  (i.e., the null space of  $\mathbf{K}_\theta$  is spanned by a constant vector). Because  
261 of the properties of the Kronecker product, the rank deficiency of  $\mathbf{K}_\delta$  depends on the  
262 specified type of interaction:  $rank(\mathbf{K}_\delta) = T \times S - rank(\mathbf{K}_{time}) \times rank(\mathbf{K}_{space})$ ; when a  
263 Type IV interaction is considered, the number of required linear constraints becomes  
264 extremely high when  $S$  and  $T$  are large.

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

269

### 270 *3.1 Managing compositional covariates*

271

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

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

290

$$291 \quad \mathbf{b}_i = \sqrt{\frac{D-i}{D-i+1}} \left[ \underbrace{0, \dots, 0}_{i-1 \text{ times}}, +1, \underbrace{-(D-i)^{-1}, \dots, -(D-i)^{-1}}_{D-i \text{ times}} \right]; \quad \mathbf{b}_i \in \mathfrak{R}^D; i = 1, \dots, D-1. \quad (0)$$

292

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

294

$$295 \quad x_i = \text{ilr}(\mathbf{z})_i = \sqrt{\frac{D-i}{D-i+1}} \ln \frac{z_i}{\sqrt[D-i]{\prod_{j=i+1}^D z_j}} \quad \text{for } i = 1, \dots, D-1. \quad (0)$$

296

297 In the context of compositional data analysis, other widely used transformations include  
 298 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 between *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  $z$ . Thus, when *ilr*-transformed covariates,  $X$ , are introduced into the  
317 linear model  $y = \alpha + X\beta + \varepsilon$ , the first regression coefficient,  $\beta_1$ , 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.



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

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

330

$$331 \quad \tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-}\tilde{\mathbf{X}} = \mathbf{X}\mathbf{B}(\mathbf{B}'\mathbf{X}'\mathbf{X}\mathbf{B})^{-}\mathbf{B}'\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

332

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

336

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

341

$$342 \quad \mathbf{I}'_D \tilde{\boldsymbol{\beta}} = \sum_{d=1}^D \tilde{\beta}_d = 0 \tag{0}$$

343

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

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

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

356

### 357 *3.2 Computational details*

358

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

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

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

381

$$382 \quad w_{ist} \sim N(\mathbf{x}'_{st}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1) \quad i = 1, \dots, n; \quad s = 1, \dots, S; \quad t = 1, \dots, T$$

383

384 The Binomial likelihood is retrieved by specifying

385

$$386 \quad \tilde{y}_{ist} = \begin{cases} 1 & \text{if } w_{ist} > 0 \\ 0 & \text{if } w_{ist} \leq 0 \end{cases} \quad \sum_i \tilde{y}_{ist} = y_{st}$$

387

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

390

$$391 \quad w_{ist} | \dots \sim \begin{cases} N(\mathbf{x}'_{st}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1)I(w_{ist} > 0) & \text{if } \tilde{y}_{ist} = 1 \\ N(\mathbf{x}'_{st}\boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st}, 1)I(w_{ist} \leq 0) & \text{if } \tilde{y}_{ist} = 0 \end{cases} \quad (0)$$

392

393 Summing over the pixels, we obtain  $w_{\bullet st} = \sum_{i=1}^n w_{ist}$ . Conditionally on  $w_{\bullet st}$ , MCMC

394 sampling can proceed as in a standard Bayesian linear model, where  $w_{\bullet st}$  serves as

395 pseudo-data. The linear predictor can be written in compact form as follows:

396

$$397 \quad \mathbf{w}_{\bullet} = \mathbf{X}'\boldsymbol{\beta} + \boldsymbol{\alpha} \otimes \mathbf{I}_S + \mathbf{I}_T \otimes \boldsymbol{\theta} + \boldsymbol{\delta}$$

398

399 where  $\mathbf{I}_k$  denotes a  $k$ -dimensional unit vector,  $\mathbf{X} = \text{diag}(\mathbf{X}_1, \dots, \mathbf{X}_t, \dots, \mathbf{X}_T)$  is the

400  $ST \times T(D-1)$ -dimensional block-diagonal matrix with  $S \times (D-1)$ -dimensional blocks

401  $\mathbf{X}_t$  containing *ilr*-transformed covariates at time  $t$  and  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_t, \dots, \boldsymbol{\beta}_T)$  is the

402  $(D-1)T$ -dimensional vector of time-varying coefficients.

403 In the context of the Bayesian linear regression model, implementing a Gibbs sampler

404 involves standard MCMC tools; it is worth noting that sampling under linear constraints

405 for the interaction term, even for a Type IV interaction model, is not an hard task. In

406 fact, it can be shown that the null space spanned by the eigenvectors associated with

407 null eigenvalues of the structure matrix  $\mathbf{K}_{\delta}$  is equivalent to the null space spanned by

408 the eigenvectors associated with null eigenvalues of the Kronecker product  $\mathbf{C}_S \otimes \mathbf{C}_T$ ,

409 where  $\mathbf{C}_k$  denotes the  $k$ -dimensional centering matrix. Therefore, samples from the full

410 conditional distribution of  $\boldsymbol{\delta}$  can be obtained by centering samples over space and time

411 on the fly at each MCMC iteration. Full conditional distributions for all of the

412 parameters are reported in Appendix A.

413 Several further refinements to the Gibbs sampler built for our case study can be adopted

414 to improve the efficiency of the MCMC algorithm. Regarding the auxiliary variable

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

421  
 422

## 423 4 APPLICATION

424  
 425 *4.1 Model selection*

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

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

435

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

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

438

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

447 Ultimately, the selected model is:

448

$$449 \quad y_{st} | \pi_{st} \sim \text{Binomial}(\pi_{st}, n) \quad s = 1, \dots, S; t = 1, \dots, T$$

$$450 \quad \Phi^{-1}(\pi_{st}) = \mathbf{x}'_{st} \boldsymbol{\beta}_t + \alpha_t + \theta_s + \delta_{st} \quad \text{where } \boldsymbol{\beta}_t = (\beta_{1t}, \dots, \beta_{dt}, \dots, \beta_{(D-1)t})$$

$$451 \quad \boldsymbol{\beta}_d = (\beta_{d1}, \dots, \beta_{dt}, \dots, \beta_{dT}) \sim \text{IGMRF}_T(\tau_{\beta_d} \mathbf{K}_{RW}); d = 1, \dots, D-1$$

$$452 \quad \boldsymbol{\alpha} \sim \text{IGMRF}_T(\tau_{\alpha} \mathbf{K}_{RW}), \boldsymbol{\theta} \sim \text{IGMRF}_S(\tau_{\theta} \mathbf{K}_{\theta}); \boldsymbol{\delta} \sim \text{IGMRF}_{ST}(\tau_{\delta} \mathbf{K}_{\delta})$$

453

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

457

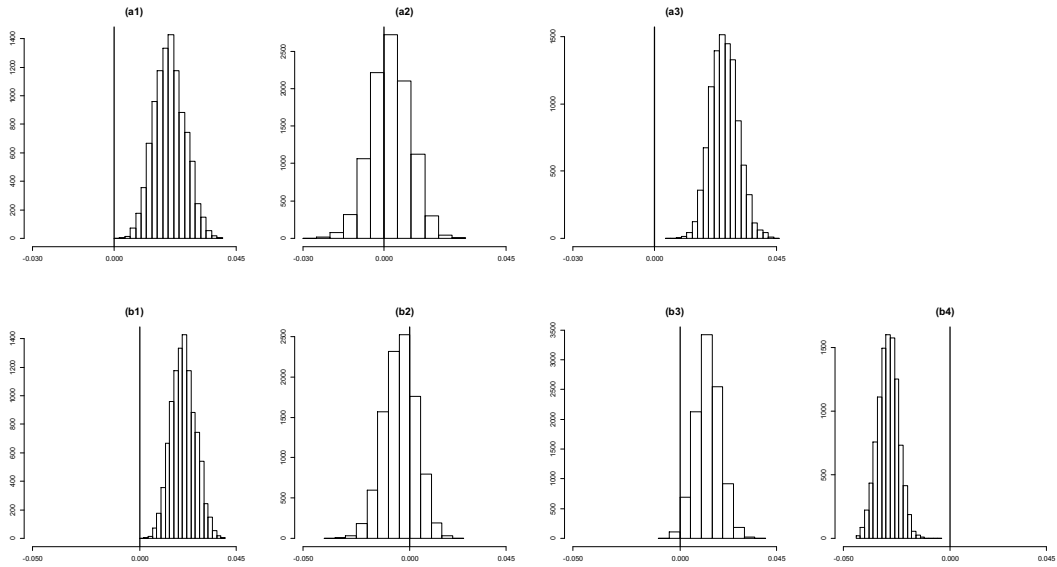
458

## 459 4.2 Results

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

463 In the upper panels (a1-a3) of Figure 3, posterior distributions ( $\beta_t | \mathbf{y}$ ) of parameters  
464  $\beta_t = (\beta_{1t}, \dots, \beta_{dt}, \dots, \beta_{(D-1)t})$  are reported for  $t = 9$  as an example. The lower panels (b1-  
465 b4) report posterior distributions of *clr*-regression coefficients  $\tilde{\beta}_t$ ; these posterior  
466 distributions are obtained by transforming  $\beta_t$  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



484

485 **Fig. 3** Posterior distributions of the regression coefficients for the last collection time (panels a1-a3).  
 486 Posterior distributions of transformed regression coefficients expressing the relative suitability for moss  
 487 (b1), litter (b2), soil (b3) and bare rock (b4). Vertical lines correspond to zero  
 488

489 Because transformed coefficients are directly comparable, we provide some suggestions  
 490 on how posterior estimates can be used to describe findings about substrate suitability.

491 Upon first glance, at a given time, the substrates can be ordered from the most to the  
 492 least suitable for vegetation cover based on the posterior means  $E(\tilde{\beta}_{dt} | \mathbf{y})$ : at time 9,

493 moss turns out to be the most suitable substrate typology, followed by soil, litter and  
 494 bare rock. Moreover, a rigorous comparison between the suitability of substrates  $d$  and  
 495  $d'$  that takes account of the uncertainty about parameter estimates can be obtained by

496 evaluating  $Pr(\tilde{\beta}_{dt} > \tilde{\beta}_{d't} | \mathbf{y})$ , i.e., the posterior probability that the suitability of

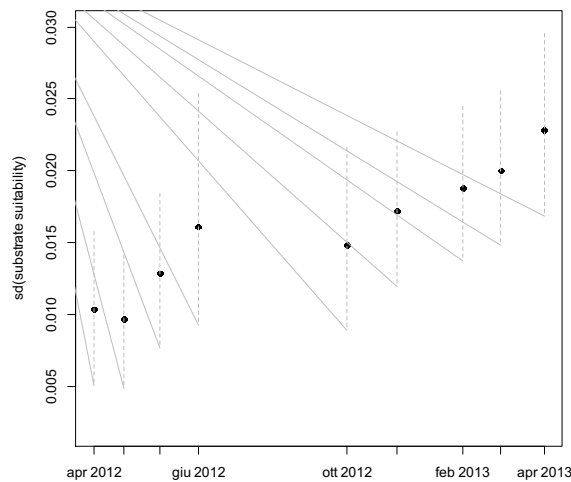
497 substrate  $d$  is higher than that of substrate  $d'$ . This task is easily addressed by

498 exploiting the MCMC sample. It is observed that, at time 9, the posterior probabilities  
 499 that moss is more suitable than litter, soil and bare rock are 0.98, 0.81 and 1,  
 500 respectively.

501 Furthermore, it is interesting to determine whether the heterogeneity characterizing  
 502 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  $sd(\tilde{\beta}_t)|\mathbf{y}$ ,  $t=1,\dots,9$ , 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  $sd(\tilde{\beta}_t)|\mathbf{y}$  over time

510

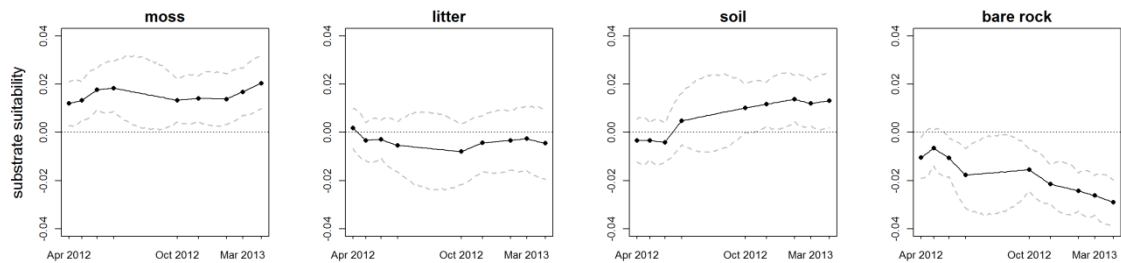
511 The biological motivation of this behavior can be found in the increase in vegetation  
 512 cover, from October to the end of the study period, caused by a higher propensity of  
 513 vegetation to occur in the substrates that exhibit greater availability of nutrient resources  
 514 (i.e., moss and soil) than do the others (litter and bare rock).

515

516 Finally, we describe findings about the temporal evolution of substrate relative  
 517 suitability that are helpful in addressing the ecological questions posed at the end of  
 518 section 2. In each panel of Figure 5, the posterior mean of  $\tilde{\beta}_t$  at the collection times  
 (filled circles), the predictions at unobserved times (black solid line) and credible bands

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

521



522

523

**Fig. 5** Substrate relative suitability over time

524

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

532 It is worth noting that, in contrast to the results shown in Figure 2, where the temporal  
533 pattern of substrate-specific effects is confounded with the marginal trend for  
534 vegetation, the results shown in Figure 5 provide direct information on the relative  
535 suitability level of each substrate and how these levels vary over the study period, after  
536 properly accounting for spatial and temporal correlation. These results are in agreement  
537 with the ecological literature on this type of habitat (Lundholm 2003, Otsus and Zobel  
538 2002) as stated in Velli (2014).

539

540

## 5. CONCLUSIONS

541

542

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

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

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

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

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

581

582

583

## APPENDIX

584

585 The Gibbs sampler is implemented as follows:

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

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

590

$$\begin{aligned} \tilde{\boldsymbol{\eta}} &= (\mathbf{X} \otimes \mathbf{I}_n) \boldsymbol{\beta} + (\mathbf{I}_T \otimes \mathbf{I}_S \otimes \mathbf{I}_n) \boldsymbol{\alpha} + (\mathbf{I}_T \otimes \mathbf{I}_S \otimes \mathbf{I}_n) \boldsymbol{\theta} + (\mathbf{I}_T \otimes \mathbf{I}_S \otimes \mathbf{I}_n) \boldsymbol{\delta} \\ &= \tilde{\mathbf{X}} \boldsymbol{\beta} + \tilde{\mathbf{A}}_1 \boldsymbol{\alpha} + \tilde{\mathbf{A}}_2 \boldsymbol{\theta} + \tilde{\mathbf{A}}_3 \boldsymbol{\delta} \end{aligned}$$

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

$$596 \quad \boldsymbol{\beta} | \dots \sim N_C(\mathbf{b}_\beta, \mathbf{Q}_\beta)$$

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

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

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

600 structure:

601

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

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

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

$$605 \quad \boldsymbol{\alpha} | \dots \sim N_C(\mathbf{b}_\alpha, \mathbf{Q}_\alpha)$$

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

$$607 \quad \mathbf{Q}_\alpha = \tau_\alpha \mathbf{K}_{RW} + \tilde{\mathbf{A}}_1' \tilde{\mathbf{A}}_1 = \tau_\alpha \mathbf{K}_{RW} + n \mathbf{S} \mathbf{I}_T$$

608 4. Sample the spatial effect from the distribution

$$609 \quad \boldsymbol{\theta} | \dots \sim N_C(\mathbf{b}_\theta, \mathbf{Q}_\theta)$$

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

611 
$$\mathbf{Q}_\theta = \tau_\theta \mathbf{K}_\theta + \tilde{\mathbf{A}}_2' \tilde{\mathbf{A}}_2 = \tau_\theta \mathbf{K}_\theta + n\mathbf{I}_S$$

612 Centre the sampled values.

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

614 5. Sample the spatiotemporal random effects from the distribution

615 
$$\boldsymbol{\delta} | \dots \sim N_C(\mathbf{b}_\delta, \mathbf{Q}_\delta)$$

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

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

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  $\tau_{\beta_1}, \dots, \tau_{\beta_{D-1}}, \tau_\alpha, \tau_\theta, \tau_\delta$ . The full conditionals for these  
621 parameters are Gamma, for instance:

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

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](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 Approximation. 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.