

Coregionalized Single- and Multi-Resolution Spatially-Varying Growth Curve Modelling with Application to Weed Growth

Sudipto Banerjee,^{1*} and Gregg A. Johnson²

^{1*} Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota 55455.

² Department of Agronomy and Plant Genetics, University of Minnesota, Minneapolis, Minnesota 55455.

* *email*: sudiptob@biostat.umn.edu

SUMMARY. Modelling of longitudinal data from agricultural experiments using growth curves help understand conditions conducive or unconducive to crop growth. Recent advances in Geographical Information Systems (GIS) now allow geocoding of agricultural data that help understand spatial patterns. A particularly common problem is capturing spatial variation in *growth patterns* over the entire experimental domain. Statistical modelling in these settings can be challenging because agricultural designs are often *spatially replicated*, with arrays of sub-plots, and interest lies in capturing spatial variation at possibly different *resolutions*. In this article we develop a framework for modelling spatially-varying growth curves as Gaussian processes that capture associations at single and multiple resolutions. We provide Bayesian hierarchical models for this setting, where flexible parametrization enables spatial estimation and prediction of growth curves. We illustrate using data from weed-growth experiments conducted in Waseca, Minnesota, that recorded growth of the weed *Setaria* spp in a spatially replicated design.

KEY WORDS: Bayesian inference; Coregionalization; Growth curves; Kronecker products; Markov Chain Monte Carlo; Multi-resolution models; Multivariate Gaussian processes; Non-stationary models; Non-separable models.

1 Introduction

Weed growth in agricultural fields adversely impacts productivity and profitability to farmers and society. Investigating the sources and conditions conducive to weed growth facilitate critical decision-making regarding the location and type of control measures to implement. Traditionally, weed control strategies have been applied homogeneously across the field with scant regard for the spatial nature of weed growth (Colbach et al., 2000; Johnson et al., 1995; Johnson, et al. 1996). Today, new technologies are available that allow farmers and land managers to apply weed control strategies site-specifically across a field. Yet, scientific understanding of spatially variable weed growth is still fledgling hampering full utilization of site-specific weed management (Johnson and Huggins, 1998). Recent attempts to understand the spatial arrangement of weeds use multivariate statistical methods such as canonical correlation analysis (Dieleman et al., 2000), without considering weed *growth* in the analysis, focusing only upon a single time period. Alternatively, Deen et al. (2001), Batchelor et al. (2002) employ crop growth models for investigating the temporal aspects of plant growth over homogeneous soil units, but these mechanistic models are based upon data with relatively large time intervals (e.g., between years).

Unlike weed growth modelling with *point-processes* (see, e.g., Møller and Waagepetersen, 2004), where the locations (weed-patches) are considered arising from a random process, we consider a *point-referenced* setting with the locations *fixed* by design and data (e.g. weed density over a patch) are monitored progressively through time. Hitherto, such modelling in agronomy has focused upon multi-environment trial (MET) contexts for plant variety improvement programs. They have been recently discussed in multiplicative frameworks (e.g., Smith et al. 2001a; 2001b) using factor-analytic methods. Our current work focuses upon point-referenced spatial modelling of short-term weed-growth (in days) geared towards understanding spatial *growth patterns* using spatially varying growth curves.

An added wrinkle here is that spatial variation might exist at different *resolutions* due to replicated designs that involve sites arranged in small rectangular arrays. Since soil and environmental factors are expected to vary across the terrain, *macro* level variation is expected between these arrays, while some

micro level variation is likely within an array as not all spatial variation is neutralized (possibly due to a non-exhaustive set of covariates). Indeed, the underlying *spatial range* (see, e.g., Banerjee, Carlin and Gelfand, 2004, Chapter 2) will likely be different for the macro and micro levels, given that they represent domains of differing sizes.

Multiple range problems also arise in multivariate spatial modelling with different response variables exhibiting different spatial correlation structures. *Coregionalization* techniques (Wackernagel, 2003; Gelfand et al., 2004; Banerjee et al., 2004, Chapter 7) that generalize *separable* models (e.g., Smith, Cullis and Thompson 2001a) help account for this problem. However, applying such multivariate methods for nested resolution settings is not straightforward raising concerns about valid cross-covariance functions. Our primary contribution lies in addressing these issues with a proposed framework for spatial multi-resolution problems that allow flexible non-stationarity through space-varying linear transformations of simpler processes. Our framework will encompass a wide variety of spatial models, including most of the correlation structures popularly used today.

A Bayesian approach is attractive here as we deal with inference for unobserved functional processes. More precisely, we seek ensemble studies of growth patterns over the entire spatial domain including unobserved locations, which are likely to offer additional insight into topographic relationships for weed growth. Hierarchical modelling of the growth curves as spatial processes offers a computationally feasible template that is easy to interpret, yet accommodates subsequent model comparisons. Furthermore, agronomy experiments typically have limited numbers of arrays and subplots that cause concern for model identifiability and over-parametrization for moderately rich spatial structures. We offer a formulation using nested processes that ensures parameter identifiability and stable estimation algorithms.

The remainder of the paper evolves as follows. Section 2 develops different models for the data, starting with non-spatial longitudinal growth models and moving towards the more complex coregionalization models. Section 3 discusses model implementation issues. Section 4 focuses upon the analysis of the weed growth data and Section 5 concludes the paper with a discussion.

2 Models for spatially replicated growth data

A typical spatially-replicated experimental setup for weed-growth involves N_s *main plots*, each associated with a rectangular array comprising N_r *subplots* (replicates) on which weed density is monitored on a daily basis (DAP: days after planting). Figure 1, to which we return in Section 4, displays a typical such spatial arrangement consisting of 10 arrays (with the corresponding main plots (centroids) shown as asterisks) of $3 \times 2 = 6$ subplots (shown as dots). Note that all physical measurements are obtained from repeated measurements over time in the subplots. The main plots, meanwhile, act as surrogates for these arrays and permit modelling of macro level spatial variation between the arrays.

Scientific hypotheses for weed growth involve the strength and nature of spatial association in weed growth patterns. Agronomists are particularly interested in ascertaining presence of spatial association in growth patterns and, when present, their strength at the micro and macro levels. These require formulation and assessment of statistical models with different association structures – from simple non-spatial models to more complex multi-resolution models. Subsequently, a formal model comparison framework will be used to compare these models to help confirm spatial hypotheses.

2.1 *Non-spatial growth curve models*

With insignificant spatial variation in the growth patterns, simple non-spatial growth curve models (see, e.g. Diggle et al., 2002) might suffice for data analysis. In the absence of spatial variation, we do not distinguish between the micro and macro level growth patterns, so we treat the subplots as individual locations observing the respective growth curve. Therefore, letting $N = N_s N_r$ simply be the total number of locations, if Y_{it} is the response (weed density on the log scale recorded for location i in time t) then our basic growth curve model is

$$Y_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + f_i(t) + \epsilon_{it}, \quad i = 1, \dots, N \quad (1)$$

where \mathbf{x}_{it} is a vector of covariates specific to the i^{th} location and time t , $f_i(t)$ is a function capturing the growth through time and ϵ_{it} is measurement error following $N(0, \tau^2)$. Note that we will assume $f_i(t)$ to be

linear in t , based upon the empirical evidence from the raw data.

Equation (1) includes several non-spatial growth curve models by specifying $f_i(t)$. For example, if no variation in growth patterns is expected between locations, a *uniform* growth curve model, where $f_i(t)$ does not depend upon i will suffice. Taking α_0 and α_1 as independent normal coefficients or as correlated with covariance matrix Λ and letting $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$, we have the following two models:

$$\text{Model 1a: } f_i(t) = \alpha_0 + \alpha_1 t; \boldsymbol{\alpha} \sim N(\mathbf{0}, \text{Diag}(\sigma_0^2, \sigma_1^2))$$

$$\text{Model 1b: } f_i(t) = \alpha_0 + \alpha_1 t; \boldsymbol{\alpha} \sim N(\mathbf{0}, \Lambda),$$

A richer model is obtained by letting the intercept and slope vary by location as Gaussian variables:

$$\text{Model 1c: } f_i(t) = \alpha_{0i} + \alpha_{1i} t; \boldsymbol{\alpha}_i \sim N(\mathbf{0}, \Lambda)$$

where $\boldsymbol{\alpha}_i = (\alpha_{0i}, \alpha_{1i})$ are i.i.d. bivariate normal distributions with covariance matrix Λ . Model 1c allows independent variation across sites and correlation between the intercept and slope *within* each site, but does not account for correlation *between* the sites.

2.2 *Single resolution spatial growth curve models*

It is reasonable to expect similar growth patterns in proximate locations as they share similar topographic and environmental conditions. In other words, the growth curves are likely to be *spatially associated* and should be modelled as a function of distance between locations: as this distance increases the association diminishes. Scientific interest resides in capturing this phenomenon, hypothesized through spatial models, which we classify as single-resolution and multi-resolution. The former ignores the nested structure of the spatial domain, but accounts for spatial association between the different growth curves based upon all the subplot locations, just as a simple univariate geostatistical model would (e.g., Banerjee et al., 2004, Chapter 5). Thus, as in Section 2.1, we let $N = N_s N_r$ be the total number of locations considered for modelling and $Y_t(s)$ be the weed density in subplot $s \in \mathcal{S}$ at time t , where \mathcal{S} is our spatial domain. Also,

$\mathbf{x}_t(s)$ is the associated vector of covariates, and we now write our model as

$$Y_t(s) = \mathbf{x}_t(s)^T \boldsymbol{\beta} + f(s, t) + \epsilon_t(s), \quad (2)$$

where $f(s, t) = \alpha_0(s) + \alpha_1(s)t$ and $\epsilon_t(s)$ are i.i.d. $N(0, \tau^2)$. Unlike the preceding section where $\boldsymbol{\alpha}_i$ simply indexed the site i , here $\alpha_0(s)$ and $\alpha_1(s)$ are function of s allowing spatial association with other sites.

Modelling the intercept and slope as spatial processes permit full inference in a Bayesian hierarchical setting, including prediction of growth patterns at arbitrary locations. We model spatial association for the intercept and slope processes and the correlation *between* them using a *bivariate* Gaussian process (see, e.g., Banerjee et al., 2004, Chapter 7) for $\boldsymbol{\alpha}(s) = [\alpha_0(s), \alpha_1(s)]$, denoted by $GP[\boldsymbol{\mu}(s), \Gamma_\alpha(\cdot)]$. Here $\boldsymbol{\mu}(s) = [\mu_0(s), \mu_1(s)]$ are the process mean functions, and $\Gamma(\cdot)$ is a 2×2 matrix function ($p \times p$ if $f(s, t)$ had p coefficient processes) defining the cross-covariance functions of the process. Equivalently, we may define $\tilde{\alpha}_0(s) = \alpha_0(s) - \mu_0(s)$, and $\tilde{\alpha}_1(s) = \alpha_1(s) - \mu_1(s)$. These yield a corresponding zero-centered bivariate Gaussian process, $[\tilde{\alpha}_0(s), \tilde{\alpha}_1(s)] \sim GP[\mathbf{0}, \Gamma(\cdot)]$. In particular, when the process means are constant across sites, say $\mu_0(s) = \mu_0$ and $\mu_1(s) = \mu_1$, the growth function can be looked upon as the sum of a non-spatial component and a spatially-varying component, $f(s, t) = g(t) + \tilde{f}(s, t)$, where $\tilde{f}(s, t) = \tilde{\alpha}_0(s) + \tilde{\alpha}_1(s)t$ and $g(t) = \mu_0 + \mu_1 t$.

The cross-covariance matrix $\Gamma_\alpha(\cdot)$ plays a central role in the construction of valid multivariate processes. It is a function on $\mathcal{S} \times \mathcal{S}$ into a subset of 2×2 matrices such that, for any two sites s and s' in \mathcal{S} :

$$\Gamma_\alpha(s, s') = \begin{pmatrix} \text{Cov}[\alpha_0(s), \alpha_0(s')] & \text{Cov}[\alpha_0(s), \alpha_1(s')] \\ \text{Cov}[\alpha_1(s), \alpha_0(s')] & \text{Cov}[\alpha_1(s), \alpha_1(s')] \end{pmatrix}.$$

In particular, note that with $s = s'$, $\Gamma_\alpha(s, s)$ is a covariance matrix for the elements of $\boldsymbol{\alpha}(s)$ within site s . Given the process realization over a set of sites $(s_i)_{i=1}^N$, the joint dispersion matrix of $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(s_i)]_{i=1}^N$ is determined by the cross-covariance matrix as $\Sigma_\alpha = [\Gamma_\alpha(s_i, s_j)]_{i,j=1}^N$. That is, Σ_α is a $2N \times 2N$ block matrix of 2×2 matrices with the $(i, j)^{th}$ block being $\Gamma_\alpha(s_i, s_j)$. A valid $\Gamma_\alpha(\cdot)$ is one that ensures a symmetric and positive definite Σ_α . Note, however, that $\Gamma_\alpha(s, s')$ for $s \neq s'$ need not be positive definite or even

symmetric, except in the limiting sense as $s \rightarrow s'$, but must satisfy $\Gamma_\alpha(s, s') = \Gamma_\alpha^T(s', s)$ to ensure the symmetry of Σ_α . These multivariate processes are *stationary* when the cross-covariances are functions of the separation between the sites, in which case we write $\Gamma_\alpha(s, s') = \Gamma_\alpha(s - s')$.

Characterizing valid cross-covariances that ensure positive-definiteness of Σ_α is indeed more demanding than, say, for correlation functions in univariate spatial modelling. More precisely, we require that for an arbitrary number and choice of locations, the resulting Σ_α be positive definite. For ensuring this in a computationally feasible manner, we adopt a constructive approach that models Γ_α by treating $\boldsymbol{\alpha}(s)$ as a nonsingular linear transformation (centered at $\boldsymbol{\mu}(s)$) of *latent* independent, zero-centered stationary processes $v_0(s)$ and $v_1(s)$ with unit variances. Letting $\mathbf{v}(s) = [v_0(s), v_1(s)]$ and A be a 2×2 non-singular matrix, we can write this relation down as $\boldsymbol{\alpha}(s) = \boldsymbol{\mu}(s) + A\mathbf{v}(s)$. Specifying $\mathbf{v}(s)$ as a bivariate $GP[\mathbf{0}, \Gamma_{\mathbf{v}}(\cdot)]$, we obtain $\boldsymbol{\alpha}(s)$ as a bivariate $GP[\boldsymbol{\mu}(s), \Gamma_\alpha(\cdot)]$ with $\Gamma_\alpha(s - s') = A\Gamma_{\mathbf{v}}(s - s')A^T$. Note that $\Gamma_{\mathbf{v}}(s - s') = \begin{pmatrix} \rho_0(s - s') & 0 \\ 0 & \rho_1(s - s') \end{pmatrix}$, where $\rho_0(\cdot)$ and $\rho_1(\cdot)$ are the correlation functions for $v_0(s)$ and $v_1(s)$ respectively. The independence of $v_0(s)$ and $v_1(s)$ is not necessary and any positive-definite specification of $\Gamma_{\mathbf{v}}(s - s')$ could be used. However, we emphasize that specification of such positive definite matrices can be an arduous task, especially for more general $p \times p$ settings, so we intentionally avoid this route.

Instead, we choose to incorporate richness in spatial covariances using the coefficient matrix A . Irrespective of whether a physical interpretation is attached to it, $\mathbf{v}(s)$ provides us with this constructive framework: using simple (stationary) specifications for $\mathbf{v}(s)$, we build richer and more versatile correlation structures for $\boldsymbol{\alpha}(s)$ by appropriately choosing the linear transformation A . Any valid process must have $\Gamma_\alpha(0)$ as positive definite and, since $\Gamma_\alpha(0) = AA^T$, A can be unambiguously chosen as the lower or upper triangular Cholesky square root of $\Gamma_\alpha(0)$. The one-one correspondence between the elements of A and $\Gamma_\alpha(0)$ is well-known (see, e.g., Harville 1997, p229), so for statistical estimation the elements of A are identifiable only up to a lower or upper triangular structure; we opt for the lower-triangular structure. Again, $\Gamma_{\mathbf{v}}(s - s')$ yields the dispersion matrix $\Sigma_{\mathbf{v}}$, which in turn provides the dispersion for $\boldsymbol{\alpha}$ as $\Sigma_\alpha = A\Sigma_{\mathbf{v}}A^T$,

where $\mathcal{A} = (I_N \otimes A)$ and \otimes denotes the Kronecker product (e.g., Harville, 1997). This is positive definite whenever $\Sigma_{\mathbf{v}}$ is, ensuring a valid multivariate process $\boldsymbol{\alpha}(s)$ that is *induced* from a latent process.

We illustrate our constructive approach to build three models of increasing complexity in covariance structures. Model 2a assumes that $\alpha_0(s)$ and $\alpha_1(s)$ are independent processes with possibly different correlation structures. This is simple enough for direct construction, but is obtained in the above framework by letting $A = \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_1 \end{pmatrix}$, so that

$$\text{Model 2a: } \Gamma_{\alpha}(s - s') = A\Gamma_{\mathbf{v}}A^T = \begin{pmatrix} \sigma_0^2\rho_0(s - s') & 0 \\ 0 & \sigma_1^2\rho_1(s - s') \end{pmatrix}$$

Clearly, the resulting $\Sigma_{\mathbf{v}}$ is p.d., hence so is Σ_{α} .

Our next model, Model 2b, associates $\alpha_0(s)$ and $\alpha_1(s)$ through an “intrinsic” specification so that

$$\text{Model 2b: } \Gamma_{\alpha}(s - s') = \Lambda\rho(s - s'),$$

where $\Lambda = \Gamma_{\alpha}(0)$ is a 2×2 variance-covariance matrix between the processes (within each site), and $\rho(s - s')$ is a *common* spatial correlation function capturing spatial association. Now $v_0(s)$ and $v_1(s)$ are also *identical* Gaussian processes so that $\Gamma_{\mathbf{v}}(\cdot)$ reduces to $\rho(s - s')I_2$ with $AA^T = \Lambda$. This is precisely the *separable* modelling (e.g. Smith et al., 2001a), which factors out the cross-covariance function into a spatial and non-spatial component. Recall that AA^T can be interpreted as the covariance matrix between the intercept and slope processes. In particular, its diagonal elements play the roles of spatial variance parameters (similar to σ_0^2 and σ_1^2 in Model 2a). Separability has nicer interpretations and offers significant computational benefits in likelihood evaluations, but imposes a common spatial correlation function (hence a common spatial range) for both the intercept and slope processes. Since there is no reason, a priori, to make such an assumption, this may be undesirable.

In order to incorporate association between the intercept and slope processes (absent in Model 2a) *and* let them have different spatial correlation functions (unlike in Model 2b), we resort to a more general

coregionalization model by taking $v_0(s)$ and $v_1(s)$ as in Model 2a (independent, but not identical) while letting A remain an unspecified lower-triangular matrix as in Model 2b. This yields,

$$\text{Model 2c: } \Gamma_\alpha(s - s') = A\Gamma_\nu(s - s')A^T; A \text{ is lower-triangular.}$$

The nicer interpretation of separability is somewhat lost, but we achieve richer modelling with process specific spatial ranges. Not only does this richness lead to better fitting models (see Section 4.2), but our framework can help assess appropriateness of separability and other hypotheses.

2.3 *Multi-resolution spatial growth curve models*

While single-resolution models are simpler to interpret and implement, they are inadequate for assessing possible multi-level variation in the intercept and slope processes. Each array of replicates occupies a unique position in the spatial domain (see Figure 1), making micro-level variation (between subplots) *site-specific*. We conceive nested micro-level spatial associations within a global macro field, distinguishing between *main plots* and *subplots* (or replicates) within the main plots.

The centroid of each array defines the locations of the main plots, say s , while the subplots, denoted by r , constitute replicates *nested* within a given array. Thus, our unit of spatial referencing, or location, here is an ordered pair (r, s) , treated as r residing within s . Perhaps a stricter notation would be $(s(r), s)$ to bring out the nesting, but we *always* consider r inside an array s and, hence, do not foresee confusion. This is different from Section 2.2, where s denoted the subplot coordinates and the spatial orientation of the main plots was never used. Now we distinguish between the main plot and subplots letting N_s be the number of main plots, N_r the number of subplots in each array and N_t the number of time points measuring the response. Letting $Y_t(r, s)$ be the response at time t from location (r, s) and $\mathbf{x}_t(r, s)$ be a $p \times 1$ vector of covariates from this location, a general multi resolution model is given by

$$Y_t(r, s) = \mathbf{x}_t^T(r, s)\boldsymbol{\beta} + f[(r, s), t] + \epsilon_t(r, s). \quad (3)$$

Here $\epsilon_t(r, s) \stackrel{iid}{\sim} N(0, \tau^2)$, denotes the measurement error, and $f[(r, s), t]$ is a function capturing the effect

of time on the response at time t in location (r, s) . We again consider linear growth functions, $f[(r, s), t] = \alpha_0(r, s) + \alpha_1(r, s)t$, but now incorporate micro level variation in the coefficient processes.

As a natural progression from Section 2.2, we model $\boldsymbol{\alpha}(r, s) = [\alpha_0(r, s), \alpha_1(r, s)]$ as a bivariate processes, $GP[\mathbf{0}, \Gamma_\alpha(\cdot)]$ whose 2×2 cross-covariance function is defined on the (r, s) space as $\Gamma_\alpha[(r, s), (r', s')]$. We denote the $2N_r \times 2N_r$ cross-covariance matrix for the vector process $\boldsymbol{\alpha}(s) = [\boldsymbol{\alpha}(r_j, s)]_{j=1}^{N_r}$ as $\Gamma_\alpha(s, s') = [\Gamma_\alpha\{(r_i, s), (r_j, s')\}]_{i,j=1}^{N_r}$. A fully general nested spatial hypothesis suggests association at three levels: (1) correlation between the intercept and slope for each location (r, s) , (2) coefficients arising from the subplots within each main plot (or array) s , and (3) the association for coefficients arising in different main plots. We capture these nested associations using latent processes at the different resolutions (one at the micro and another at the macro level) and “equate” them to construct consistent multi-resolution processes.

First, for each main plot s we construct a local bivariate, zero-centered, stationary process on the subplots, $\mathbf{u}_s(r) = [u_{0s}(r), u_{1s}(r)]$, with $\Gamma_{\mathbf{u}_s}(r - r') = \begin{pmatrix} \rho_{0s}^{mic}(r - r') & 0 \\ 0 & \rho_{1s}^{mic}(r - r') \end{pmatrix}$ generating the within-site dispersion matrix $\Sigma_{\mathbf{u}_s}$. We define the micro-level covariance structure on $\boldsymbol{\alpha}(r, s)$ through $\boldsymbol{\alpha}(r, s) = \boldsymbol{\mu}(r, s) + B_s(r)\mathbf{u}_s(r)$. For a fixed s , this defines a coregionalized single-resolution model over the subplots in s , where $B_s(r)$ is a space-varying linear transformation at the micro level. Forming a micro level cross-covariance matrix function of pairs of replicates r and r' within s as

$$\Gamma_{\alpha(s)}(r, r') = \begin{pmatrix} Cov[\alpha_0(r, s), \alpha_0(r', s)] & Cov[\alpha_0(r, s), \alpha_1(r', s)] \\ Cov[\alpha_1(r, s), \alpha_0(r', s)] & Cov[\alpha_1(r, s), \alpha_1(r', s)] \end{pmatrix},$$

we obtain $\Gamma_{\alpha(s)}(r, r') = B_s(r)\Gamma_{\mathbf{u}_s}(r - r')B_s(r')^T$. In particular when $r = r'$, we obtain $\Gamma_{\alpha(s)}(r) = B_s(r)B_s(r)^T$ (since $\Gamma_{\mathbf{u}_s}(0) = I_2$) as the covariance matrix between the slope and intercept coefficients within location (r, s) . As for the single resolution models, here $B_s(r)$ is identified up to a lower-triangular structure.

Collecting over replicates $\{r_j\}_{j=1}^{N_r}$ within any given s , we arrive at a main plot level equation relating the $2N_r \times 1$ vector processes $\boldsymbol{\alpha}(s) = [\boldsymbol{\alpha}(r_j, s)]_{j=1}^{N_r}$ and $\mathbf{u}_s = [\mathbf{u}_s(r_j)]_{j=1}^{N_r}$ as

$$\boldsymbol{\alpha}(s) = \boldsymbol{\mu}(s) + \mathcal{B}_s \mathbf{u}_s, \tag{4}$$

where \mathcal{B}_s is a $2N_r \times 2N_r$ block-diagonal matrix with lower-triangular $B_s(r_j)$'s as 2×2 blocks. For full inference in a generalized multi-resolution setting, the process $\boldsymbol{\alpha}(s)$ in (4) must generate a positive definite dispersion matrix Σ_α for any realization $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(s_i)]_{i=1}^{N_s}$ over a collection $\{s_i\}_{i=1}^{N_s}$. This would follow from a legal cross-covariance function of the process \mathbf{u}_s over s : once this cross-covariance has been legally defined, say $\Gamma_{\mathbf{u}}(s, s')$, then the cross-covariance for $\boldsymbol{\alpha}(s)$ over s , say $\Gamma_\alpha(s, s')$, equals $\mathcal{B}_s \Gamma_{\mathbf{u}}(s, s') \mathcal{B}_{s'}^T$.

Constructing a legal cross-covariance matrix for \mathbf{u}_s that retains the above micro-level covariance structure, however, is not immediately obvious. For this, we start with a latent process $\mathbf{v}(r, s)$ with cross-covariance function $\Gamma_{\mathbf{v}}[(r, s), (r', s')] = \begin{pmatrix} \rho_0^{mac}(s - s') & 0 \\ 0 & \rho_1^{mac}(s - s') \end{pmatrix}$, where ρ_0^{mac} and ρ_1^{mac} are correlation functions for the macro domain that do *not* depend upon r or r' . So $\mathbf{v}(r, s)$ is independent across the subplots, whence the $2N_r \times 1$ vector process $\mathbf{v}(s) = [\mathbf{v}(r_j, s)]_{j=1}^{N_r}$ has $\Gamma_{\mathbf{v}}(s - s') = I_{N_r} \otimes \begin{pmatrix} \rho_0^{mac}(s - s') & 0 \\ 0 & \rho_1^{mac}(s - s') \end{pmatrix}$ as its cross-covariance function. Note that with $s = s'$, $\Gamma_{\mathbf{v}}(0) = I_{2N_r}$. Now letting $\boldsymbol{\alpha}(s) = [\boldsymbol{\alpha}(r_j, s)]_{j=1}^{N_r}$ be a space-varying linear transformation of $\mathbf{v}(s)$, we obtain

$$\boldsymbol{\alpha}(s) = \boldsymbol{\mu}(s) + A(s)\mathbf{v}(s). \quad (5)$$

Equating (4) and (5) yields $\mathcal{B}_s \mathbf{u}_s = A(s)\mathbf{v}(s)$ and the micro-level covariance specifications

$$\Gamma_\alpha(s, s') = \mathcal{B}_s \Gamma_{\mathbf{u}}(s, s') \mathcal{B}_{s'}^T = A(s) \Gamma_{\mathbf{v}}(s - s') A(s')^T, \quad (6)$$

thereby relating $A(s)$ to $\Gamma_\alpha(s, s')$. In particular, for $s = s'$, (6) implies $A(s)A(s)^T = \mathcal{B}_s \Sigma_{\mathbf{u}_s} \mathcal{B}_s^T$ and equals $\Gamma_\alpha(s, s)$. Since $\Sigma_{\alpha(s)} = \Gamma_\alpha(s, s)$ can be interpreted as the covariance matrix of the process within main plot s , $A(s)$ identifies with its unique lower-triangular square-root and is consistent with setting $A(s) = \mathcal{B}_s \Sigma_{\mathbf{u}_s}^{1/2}$, where $\Sigma_{\mathbf{u}_s}^{1/2}$ is the lower-triangular Cholesky square-root of $\Sigma_{\mathbf{u}_s}$ and \mathcal{B}_s is also lower-triangular (from (4)). Equation (6) also shows that we obtain a legal cross-covariance matrix, $\Gamma_\alpha(s, s')$, for the macro-level process $\boldsymbol{\alpha}(s)$ for any valid $\Gamma_{\mathbf{v}}(s - s')$, yielding the valid dispersion matrix, $\Sigma_\alpha = \mathcal{A} \Sigma_{\mathbf{v}} \mathcal{A}^T$, for $\boldsymbol{\alpha}$ with \mathcal{A} being a block-diagonal matrix with $A(s)$'s as blocks and $\Sigma_{\mathbf{v}}$ the dispersion matrix from $\Gamma_{\mathbf{v}}(s - s')$.

In spite of the relatively complex scenario, (6) clarifies several important issues about estimation and identifiability of parameters in multi-resolution settings. In fact, the apparent over-identification wrought by the macro and micro level parameters is resolved through (6) by restricting estimation to the independent parameters in the micro and macro correlation functions (specifying \mathbf{u}_s and $\mathbf{v}(r, s)$) and the matrix \mathcal{B}_s . Specifically, estimation of the parameters in the micro level correlation functions delivers $\Sigma_{\mathbf{u}_s}$ which, along with estimation of \mathcal{B}_s , identifies $A(s)$. The $A(s)$'s for each main plot s and the parameters in $\Sigma_{\mathbf{v}}$ identify Σ_{α} . The flexibility of our setup is clear: *any* valid micro-level covariance matrix will result in a valid process for $\boldsymbol{\alpha}(s)$, which allows rich modelling of the $\Sigma_{\alpha(s)}$ using micro-level covariance functions and subsequent coregionalizations. Unlike in the single-resolution models, $\boldsymbol{\alpha}(s)$ is not a stationary process, but generates valid multivariate processes (and hence joint distributions) for any choice of $A(s)$ leading to spatially-varying Wishart processes (see Gelfand et al., 2004). However, these result in over-parametrized systems that are difficult to estimate, especially with a limited number of sites.

Although, in principle, we could model \mathcal{B}_s 's as in (6), their identifiability would require more replicates per main plot than are usually available (such as for us in Section 4). A further simplification assumes a common $B_s(r) = B$ for each (r, s) , i.e., the variance-covariance matrix of the coefficients does not vary spatially. Then, since $\mathcal{B}_s = I_{N_r} \otimes B$, computational benefits follow from the Kronecker structure and results in enhanced model identifiability for our most general model:

$$\text{Model 3c: } \Gamma_{\alpha}(s, s') = A(s)\Gamma_{\mathbf{v}}(s - s')A(s)^T; \quad A(s) = (I_{N_r} \otimes B)\Sigma_{\mathbf{u}_s}^{1/2}.$$

Here $\boldsymbol{\alpha}(s)$ is a non-stationary process that involves slope and intercept specific correlation functions at *both* the micro and macro levels. In addition, the micro level correlation functions, $\rho_{0s}^{mic}(\cdot)$ and $\rho_{1s}^{mic}(\cdot)$ are *site-specific* and can vary with s . This generality enables modelling correlation decay in the micro process specific to each array and assessing spatial strength in the coefficients over the terrain. We also consider two *special cases* of the above model that correspond to simpler spatial hypotheses. For instance, given that the subplots reside on a smaller domain, a single micro-level process, say $\rho^{mic}(\cdot)$, that is not site-specific and is the same for the coefficient processes might be sufficient. Writing $\Sigma^{mic} = [\rho^{mic}(r_j - r_{j'})]_{j,j'=1}^{N_r}$, we

obtain $\Sigma_{\mathbf{u}_s} = \Sigma^{mic} \otimes I_2$, which results in:

$$\text{Model 3b: } \Gamma_{\alpha}(s - s') = A\Gamma_{\mathbf{v}}(s - s')A^T; \quad A = (\Sigma^{mic})^{1/2} \otimes B.$$

Note that A no longer depends upon s and $\alpha(s)$ is now a stationary process. This model still incorporates different macro level correlation functions in $\Gamma_{\mathbf{v}}(s - s')$ for the intercept and slope. A further simplification can hypothesize a common macro-level correlation function, say $\rho^{mac}(\cdot)$, implying $\Gamma_{\mathbf{v}}(s - s') = \rho^{mac}(s - s')I_{2N_r}$ and offering a *fully separable* model:

$$\text{Model 3a: } \Gamma_{\alpha}(s - s') = \rho^{mac}(s - s')\Sigma^{mic} \otimes B^T.$$

This is a separable extension of single-resolution Model 2b, imposing common spatial ranges for the different coefficient processes at both resolutions. Writing $\Sigma^{mac} = [\rho^{mac}(s_i - s_{i'})]_{i,i'=1}^{N_s}$ and $\Lambda = BB^T$, the above model *factorizes* the joint dispersion of α into a macro dispersion, a micro dispersion and a between-coefficient dispersion: $\Sigma_{\alpha} = \Sigma^{mac} \otimes \Sigma^{mic} \otimes \Lambda$.

3 Bayesian implementation and model comparisons

We adopt a Bayesian approach specifying prior distributions on the parameters to build hierarchical models that are estimated using a Gibbs sampler, with Metropolis updates when required, for fitting our models (see, e.g., Gelman et al., 2003; Chapter 11). Although such algorithms are usually problem-specific, often requiring intensive coding, casting the problem in a general template allows several models to be fit without rewriting vast amounts of code.

Typically, for each of the models discussed in Section 2, the data equation can be cast into the following first-stage mixed model framework:

$$\mathbf{Y} = X\boldsymbol{\beta} + \tilde{Z}\boldsymbol{\alpha} + \boldsymbol{\epsilon}; \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \tau^2 I), \tag{7}$$

where \mathbf{Y} is the response vector, X is the covariate design matrix, $\boldsymbol{\beta}$ is the corresponding vector of regression coefficients, \tilde{Z} is the design matrix for the growth curve coefficients, $\boldsymbol{\alpha}$ are the multivariate realizations of

the growth curve processes and $\boldsymbol{\epsilon}$ is a vector of uncorrelated random errors. Markov Chain Monte Carlo model fitting proceeds with a Gibbs sampler with Metropolis steps (see, e.g., Gelman et al., 2003) on the marginalized scale, after integrating out $\boldsymbol{\alpha}$, to reduce the parameter space. The marginalized likelihood becomes $N(X\boldsymbol{\beta} + \tilde{Z}\boldsymbol{\mu}, \tilde{Z}\Sigma_{\alpha}\tilde{Z}^T + \tau^2I)$, where $\boldsymbol{\mu} = E[\boldsymbol{\alpha}]$ and Σ_{α} is the spatial covariance matrix for the respective model. The mean of the intercept process precludes inclusion of a separate intercept term in the covariate design matrix X for preserving model identifiability. Also, τ^2 acts as a residual variance after the spatial variation has been accounted for by the growth curve process. In principle, one may have different residual variances for each array (say, τ_s^2) which would result in a diagonal dispersion matrix for $\boldsymbol{\epsilon}$, but with most of the spatial heterogeneity being modelled through the intercept and slope process, identifiability of these additional parameters is practically impossible.

Specifically considering the single-resolution setting, where N is the number of plots and $\mathbf{t} = (t_1, \dots, t_{N_t})$ is the $N_t \times 1$ vector of observation time-points for growth, we have $\mathbf{Y}(s) = [Y_{t_k}(s)]_{k=1}^{N_t}$ as the N_t -dimensional vector of weed density (in the log scale) at site s for all the time points. Then, the site-specific growth curve equation can be written as,

$$\mathbf{Y}(s) = X(s)\boldsymbol{\beta} + Z\boldsymbol{\alpha}(s) + \boldsymbol{\epsilon}(s), \quad (8)$$

where $X(s) = [\mathbf{x}_{t_k}^T(s)]_{k=1}^{N_t}$, $\boldsymbol{\epsilon}(s) = [\epsilon_{t_k}(s)]_{k=1}^{N_t}$, $Z = [\mathbf{1}_{N_t} : \mathbf{t}]$, with $\mathbf{1}_{N_t}$ being the N_t dimensional vector of one's, and $\boldsymbol{\alpha}(s)$ as in Section 2.2. Setting $\mathbf{Y} = [\mathbf{Y}(s_i)]_{i=1}^N$, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(s_i)]_{i=1}^{N_s}$ and $\tilde{Z} = (I_N \otimes Z)$ immediately results in (7). Similarly, for the multi-resolution models in Section 2.3 we let $\mathbf{Y}(r, s) = [Y_{t_k}(r, s)]_{k=1}^{N_t}$ be the observations across time from location (r, s) and form $\mathbf{Y}(s) = [\mathbf{Y}(r_j, s)]_{j=1}^{N_r}$ and $\mathbf{Y} = [\mathbf{Y}(s_i)]_{i=1}^{N_s}$. Analogously, we have $X = ([\{\mathbf{x}_{t_k}(r_j, s_i)\}_{k,j}]_i)$ as the $N_t N_r N_s \times p$ dimensional covariate matrix, and $\tilde{Z} = (I_{N_r N_s} \otimes Z)$ with $Z = [\mathbf{1}_{N_t}, \mathbf{t}]$ and $\boldsymbol{\alpha}(s)$ as given in equation (4) for the first stage specification.

Bayesian hierarchical models are completed by assigning prior distributions on the parameters. Choice of prior distributions can play an important role in the efficiency of the algorithm. A flat prior is assigned for the covariate slope vector $\boldsymbol{\beta}$, while the error variance τ^2 is assigned an inverse-Gamma prior. Prior

distribution for $\boldsymbol{\mu}$ is again taken as flat, although typically some simplification is made. We assume that the mean levels of the processes vary by main plots, but remain constant across the subplots within a main plot. For instance, in the multi-resolution setting this implies $E[\boldsymbol{\alpha}(r, s_i)] = \boldsymbol{\mu}(s_i) = (\mu_{0i}, \mu_{1i})^T$. The variance covariance parameters in Σ_α depend upon the specific model and how the associated cross-covariance(s) is defined. When random effect (or spatial) variances are characterized by scalars, say σ_0^2 and σ_1^2 , (as in Models 1a and 2a), we assign inverse-Gamma distributions to them. More generally, when the dispersion is captured through a matrix Λ (e.g. in Models 1b, 1c and 2b and 3a), an inverse-Wishart prior is assigned to Λ . The non-separable single and multi-resolution models (Models 2c, 3b and 3c) involve estimation of lower-triangular matrices (A in 2c, B in 3b and 3c). The priors for these are derived by assigning inverted Wishart priors to their “squares”, AA^T or BB^T .

In general the spatial decay parameters are weakly identifiable and prior specifications become a more delicate issue, particularly for the multi-resolution models. Reasonably informative priors are needed for satisfactory MCMC behavior. We adopted the Matérn correlation function (e.g., Stein, 1999) with a smoothness parameter ν and a correlation decay parameter ϕ given by

$$\rho_{\phi, \nu}(d) = \frac{1}{2^{\nu-1} \mathcal{G}(\nu)} (2\sqrt{\nu}d\phi)^\nu K_\nu(2\sqrt{\nu}d\phi),$$

where K_ν is a modified Bessel function of the second kind with order ν and \mathcal{G} is the usual Gamma function. For a given (ϕ, ν) the *spatial range* solves $\rho_{\phi, \nu}(d) = 0.05$ for d (see, e.g., Banerjee et al., 2004, Chapter 2). We set prior distributions for the macro and micro decay parameters relative to the size of their domains. Distributions that set prior mean spatial ranges to approximately half the maximum inter-site distance usually result in a weak enough specification for stable MCMC behavior. For the micro-level function ρ^{mic} in models 3b and 3c, where learning from the data is much weaker and the smoothness parameter further impairs identifiability, we fix the smoothness parameter to $\nu = 1/2$ which gives an exponential correlation function $\rho_\kappa(d) = e^{-\kappa d}$ (denoting the micro-level spatial decay parameter by κ). The spatial range can now be explicitly solved as $\approx 3/\kappa$. At the macro level, the smoothness parameter is estimated using $U(0, 2)$ prior distributions. This choice is motivated by earlier findings (Banerjee et al. 2003) that it is almost

impossible for the data to distinguish between these smoothness parameters for values greater than 2. The full-conditional distributions for β and μ for flat priors are Gaussian, but the variance-covariance and range parameters have non-standard kernels that are updated using a Metropolis step.

In updating the parameters, generically denoted as Ω , using the marginal model as outlined above, we do not directly sample the spatial coefficients α . This shrinks the parameter space resulting in a more efficient MCMC algorithm. Nevertheless, the posterior distribution of α can be recovered in a predictive fashion by sampling from

$$P(\alpha | Data) \propto \int P(\alpha | \Omega, Data) P(\Omega | Data) d\Omega. \quad (9)$$

Once the posterior samples from $P(\Omega | Data)$, say $\{\Omega^{(g)}\}_{g=1}^G$, have been obtained, posterior samples from $P(\alpha | Data)$ are drawn by sampling $\alpha^{(g)}$ for each $\Omega^{(g)}$ from $P(\alpha | \Omega^{(g)}, Data)$. This composition sampling is routine because $P(\alpha | \Omega, Data)$ in (9) is Gaussian; in fact, from (7) we have this distribution as $N\left[(\Sigma_\alpha^{-1} + \tilde{Z}^T \tilde{Z} / \tau^2)^{-1} (\Sigma_\alpha^{-1} \mu + \tilde{Z}^T (\mathbf{Y} - X\beta) / \tau^2), (\Sigma_\alpha^{-1} + \tilde{Z}^T \tilde{Z} / \tau^2)^{-1}\right]$. The posterior estimates of these realizations can subsequently be mapped with contours to produce image and contour plots of the growth curve processes, revealing trends (or hot-spots) in baseline growth (intercept) or rate of growth (slope). Also note that spatial interpolation and prediction can also be carried out at any arbitrary site, even where no data monitoring has been performed, by computing

$$P[\alpha(s_0) | Data] \propto \int P[\alpha(s_0) | \alpha, \Omega, Data] P(\alpha | \Omega, Data) P(\Omega | Data) d\Omega d\alpha, \quad (10)$$

where s_0 is any location. In particular, for the multi-resolution models, we interpolate $\alpha(r_j, s_0)$. One added complication here is that we need to marginalize over the process parameters that depend upon the main plot s_0 . Rather than a posterior predictive approach, (10) is more conveniently implemented using a full Gibbs sampler that updates all the unknown parameters from their full conditional distributions (using appropriate Metropolis steps when necessary).

Since we consider several alternative models with varying degrees of spatial richness, we use the Deviance Information Criteria (DIC) (Spiegelhalter et al., 2002) as a measure of model choice. The DIC has nice

properties for Gaussian likelihoods (as ours) and is particularly convenient to compute from posterior samples. This criteria is the sum of the Bayesian deviance (a measure of model fit) and the (effective) number of parameters (a penalty for model complexity). It rewards better fitting models through the first term and penalizes more complex models through the second term, with lower values indicating favorable models for the data. The deviance, up to an additive quantity not depending upon Ω , is simply minus twice the log-likelihood, $D(\Omega) = -2 \log L(\text{Data} | \Omega)$, where $L(\text{Data} | \Omega)$ is the first stage Gaussian likelihood from (7) for the respective models. The Bayesian deviance is the posterior mean, $\overline{D(\Omega)} = E_{\Omega | \mathbf{Y}}[D(\Omega)]$, while the effective number of parameters is given by $p_D = \overline{D(\Omega)} - D(\bar{\Omega})$. The DIC is then given by $\overline{D(\Omega)} + p_D$ and is easily computed from the posterior samples. We also use predictive fits to assess model performance using the posterior predictive distribution (e.g., Gelman et al., 2003, p8)

$$P[Y_t(r, s) | \text{Data}] = \int P[Y_t(r, s) | \Omega] P(\Omega | \text{Data}) d\Omega, \quad (11)$$

where $P[Y_t(r, s) | \Omega]$ follows from the likelihood for any given (r, s) and t . Again, this is computed by sampling $Y_t^{(g)}(r, s) \sim P[Y_t(r, s) | \Omega^{(g)}]$, where $\Omega^{(g)}$ denotes the posterior samples as before.

4 Analysis of data

4.1 Data and design

Weed growth data were collected from field research conducted in the year 2001 at the University of Minnesota's Agricultural Ecology Research Farm (AERF) located at the Southern Research and Outreach Center in Waseca, Minnesota. The field was naturally drained and chisel plowed in the fall, followed by a field cultivator to prepare the seed bed in the spring. The experimental design included six replicates (subplots) nested as a 3×2 lattice within each main plot (see Figure 1). On the basis of the local topography we label these regions as back-slopes (B's), summits (S's) and toe-slopes (T's). These make up ten research sites that were established within a sixteen hectare area of the AERF based upon a detailed terrain analysis and soil taxonomy information. The maximum distance between the main plots (array centroids) was

approximately 186 meters. Uniform geometric dimensions of the arrays yield a common inter-subplot distance matrix within each array. Hence, the maximum distance between subplot locations within an array is approximately 7 meters, which is the same for all arrays.

For modelling growth patterns within each subplot, we monitored weed density through time. The species we study in the current experiment is *Setaria* spp., a species of grasses that are particularly troublesome in highly disturbed agricultural systems. This species germinates early in the season, has a high growth rate, and is a very prolific seed producer (Wilson and Tilman, 1995). Weed density (counts per square meter area) was determined using non-destructive methods and development data were collected weekly for five consecutive weeks. The field was tilled a day ahead of planting to prepare a level seedbed. The tillage operation (along with the planting operation) essentially eliminated all existing plants. Therefore, plants emerging after planting are most competitive and are the focus of this study. Since germination and emergence from the soil takes anywhere between 7 and 14 days, we started recording density measurements after such time that there was enough vegetation above the ground to get a good measurement, i.e., 14 days after planting (or elimination of all weeds). These are given in DAP units (Days After Planting).

Several terrain attributes that define a treatment, e.g., aspect, hill-slope position, and soil characteristics, are identified. Knowledge of soil hydrology is also essential to evaluating the processes related to weed growth. Thompson and Bell (1996) introduced a soil color index, the Profile Darkness Index (PDI), which accounts for the observed morphological difference from shoulder to depressional areas in Mollisol hydrosequences. In these landscapes, PDI is well correlated with duration of saturated conditions at various soil depths. Successful weed seed germination, establishment, and growth is partly a function of soil moisture and nutrient availability. High PDI values identify areas that are saturated and high in nutrient availability. Conversely, low PDI values indicate lower organic matter, higher sand content, and relatively low nutrient availability. Therefore, it is reasonable to expect this index to be correlated to weed growth along a hill-slope and forms an important covariate.

Figure 2 plots the variation of growth patterns in three different settings. The time axis starts after

about two weeks of planting, with the checked squares denoting the days of measurement. These plots indicate linear growth patterns for (log) weed density. Part (a) of the figure corresponds to the observed growth from two different subplots within the same site, so that the variation in growth patterns would be attributed to the micro level variation. The plots, A and B, reveal a prominent difference in the intercepts and moderate difference in the slopes, justifying the need for capturing micro-level variation. Traditional analysis sometimes hypothesizes that for arrayed subplots, such as ours, “row” and “column” effects can explain the spatial variation. If so, then two subplots located within two different arrays but occupying the same relative position (therefore having same row and column effects) will not reveal significant differences. Part (b) of Figure (2) is an example that such is not the case here with Plots C and D corresponding to subplots within two different arrays, but having the same “cell number” in their respective lattice. Here the intercepts seem to be close, but the growth rates are very different. Finally, Part (c) plots E and F representing growth from different subplots (unequal “cell numbers”) from two different arrays.

4.2 *Results from spatial growth curve models*

Turning to the analysis of the data, we fit several competing models using our template. We take the time variable as centered at 14 days after planting (DAP), when the first measurement was taken, so the intercept corresponds to weed density at the beginning. We implemented the nine models listed in Table 1. All the models contain the same set of global covariates and differ only in the hierarchical specification of the growth curve coefficient distributions. The general guidelines for priors mentioned in Section 3 were followed for the respective models. Note that $N_s = 10$, $N_r = 6$ and $N_t = 5$ for our data, so $N = 300$ is the dimension of Y in the data equation (7). While simplifying model assumptions often deliver computational benefits (particularly separable structures), the most general multi-resolution updating still involve N dimensional matrix computations. The first column in Table 1 corresponds to the non-spatial growth curves, the second to the single resolution models while the third considers the multi-resolution models. For each of these models three parallel MCMC chains were run for 15000 iterations. The CODA

package in R (www.r-project.org) was used to diagnose convergence by monitoring mixing using Gelman-Rubin diagnostics and autocorrelations (e.g., Gelman et al., 2003, Section 11.6). For each of the models, 10000 iterations revealed sufficient mixing of the chains, so the remaining 15000 samples (5000×3) were retained for posterior analysis.

Returning to Table 1, the DIC scores clearly indicate the the need and room for spatial modelling. Incorporating non-spatial independently varying growth curves (Model 1c) already brings about a sharp decrease in the scores (despite the enhanced p_D due to random effects), while spatial dependencies cause further reduction. Also, coregionalization always tends to lower the deviance substantially and seem to be marginally increasing the penalty. This lowers the scores making Model 2c and 3c preferable within the single and multi-resolution settings respectively. In fact, the multi-resolution models in the third column perform considerably better, even the relatively simpler ones, compared to their single-resolution counterparts. For purposes of illustration, we next present detailed analysis of the coregionalized multi-resolution model 3c, highlighted in Table 1 as the one offering the best DIC score.

Table 2 presents the estimates for the global covariates in Model 3c. These results correspond to flat priors for β and μ , an inverse-Gamma $IG(2, 0.001)$ for τ^2 , an inverse-Wishart $IW(2, 0.0001 \times I_2)$ with low precision for $\Lambda = BB^T$ and weakly informative Gamma $G(0.0002, 0.01)$ for ϕ_0^{mac} and ϕ_1^{mac} rendering a mean of 0.02 and a prior spatial range of around 100 meters. Percentages of sand, phosphorus and potassium seem to negatively impact the growth of *Setaria* spp. Higher PDI, which is related to higher organic matter and saturated soils, seems conducive to the growth of these weeds. The macro-level spatial parameters are presented in the next segment of Table 2 and correspond to the macro-level process. These reveal moderately different spatial variation (Λ_{00} and Λ_{11}) and a significantly negative correlation between the coefficients. This is not uncommon in linear regression settings as, typically, if one decreases the other needs to increase to fit the points well. The macro process correlation decay parameters produce quite different posteriors with the spatial decay parameter for the intercept process being almost four times that of the slope process, suggesting spatial ranges of approximately 25.3 meters and 98.5 meters respectively.

The value of τ^2 shows measurement error contributing about $\tau^2/(\Lambda_{00} + \Lambda_{11} + \tau^2) \approx 0.35\%$ of the variability captured by the model. This negligible residual variation is quite expected, given that the rather rich spatial structures would account for much of the variability.

As a result of differences in topography, soils developed from similar parent material may vary greatly within a small area (Foth, 1990). A sequence of soils along a transect that differs because of topography (e.g., summit (S) to back-slope (B) to toe-slope (T)) is called a soil catena or toposequence. Some variation in growth patterns is attributed to type of site. Of our ten main plots, we had five sites in the back-slope, three in the toe-slope and two in the summit. The site-varying parameters in Model 3c are presented in Figure 3. Parts (a) and (b) show the central 95% box-plots of the posterior distribution of the site-specific means of the intercept and slope processes (elements of $\boldsymbol{\mu}$) respectively, indicated by the type of site. The site-wise variation in the intercept process means is more pronounced than for the slopes. It seems that baseline growth patterns and growth rates are more aggressive for weeds growing in the toe-slopes (T's) than for the back-slopes (B's) and summits (S's). Parts (c) and (d) of Figure 3 show the corresponding results for the site-specific micro-level correlation decay parameters for these processes. These are precisely the κ_{0s} and κ_{1s} parameters in the exponential correlation functions for the intercept and slope processes respectively. Each of these were assigned weakly informative $U(0.5, 5)$ priors, implying spatial ranges between 0.6 and 6 meters within each array, which is reasonable given the array dimensions. For the intercept process, the decay parameter seems to be higher for the toe-slopes (hence quicker correlation decay) than for the back-slopes and summits, while the slope process is less conclusive here, although the back-slopes again seem to have lower decay rates. We emphasize that this variation can be captured only with multi-resolution modelling, with modest posterior learning seen for them.

Returning to the selected replicate-site combinations, A through F, (seen in Figure 2). For each of these locations, we computed the posterior predictive distribution $P[Y_t(r, s)|\mathbf{Y}]$, as given in (11), for a sequence of time points t and the posterior predictive means were connected to produce the solid lines in Figures 4 (a) through (f). The dashed lines through the checked boxes are the reproductions of the raw data of

Figure 2, while the remaining two dashed lines are based upon the 2.5% and 97.5% percentiles from the posterior predictive distribution (their linear appearance is due to the small residual variation). The fits are impressive, with the raw curve lying within the credible interval in all six cases. The multi-resolution variability seems to have been adequately captured by the space-varying growth curves.

Figure 5 shows image plots, with overlaid contours, for the posterior mean surfaces of the interpolated intercept and slope process respectively, obtained from evaluating (9) and (10) within the convex hull of the monitoring locations. Lighter shades correspond to higher values of the process and darker shades to the lower ones. Generally we observe zones with high baseline weed growth (intercept process) in the south-central part, while growth rates (slope process) seem to be higher in eastern fringes. This reversal of spatial pattern in baseline growth and rate of growth is not unnatural, given the negative correlation between the coefficients. Finally, we *slice* the multivariate processes by sites, and show the variation of the intercept and slope processes by replicates within each site. These are presented in the box-plots in Figures 6 and 7. The site-types (B's or S's or T's) do not seem to drastically affect the pattern of variation in the sub-plots, although modest variation is seen across all the sites.

5 Discussion

This article has developed a Bayesian framework to model weed growth using spatially-varying growth curves. We model the coefficients as spatial processes and construct non-stationary and non-separable processes to capture different associations. We provide a computationally feasible template for fitting these models. Although our application had the same time-points measuring density in all subplots, our models are easily applied to unbalanced settings. Also, our models are easily extended to non-linear growth functions such as polynomials or exponentials.

An alternative approach will employ spatiotemporal processes instead of parametric growth curves. This can be deemed a *non-parametric* version of our current model, adding further flexibility to the

modelling. More fundamentally, spatial processes can be constructed from cross-covariance *functions* (e.g., Cressie, 1993) of which our method is a special case. However, our framework emphasizes computational feasibility and simpler interpretations of growth patterns. For example, rates of weed-growth are more easily explained (and computed) through explicit slope coefficients in the growth curves than by temporal gradients of spatiotemporal processes. A substantive part of the spatial variation is often attributed to local and global terrain attributes, although methods for testing such hypothesis have remained unaddressed. Recent work by Banerjee, Gelfand and Sirmans (2003) develop statistical theory for digital terrain models, enabling joint spatial modelling of weed-growth and digital terrain features such as slopes and aspects.

Through the spatial analysis of growth patterns that help to identify unsuitable placement zones, our current work helps in contributing towards a broader goal of characterizing the extent of spatial variability in the patterns of weed growth across an agricultural landscape. This will enhance understanding of factors contributing to differences in weed growth across field landscapes. We believe that any decision support model that will be used by agricultural professionals for targeting application of integrated weed control strategies should incorporate flexible multi-resolution spatial growth models as developed here. Such analysis can also be used as a basis for optimizing the placement of alternative crops across the landscape with the goal of increasing profit while providing ecosystem services.

REFERENCES

- Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Banerjee, S., Gelfand, A.E. and Sirmans, C.F. (2003). Directional rates of change under spatial process models. *Journal of the American Statistical Association*, **98**, 946-954
- Batchelor, W.D., Basso, B. and Paz, J.O. (2002). Examples of Strategies to Analyze Spatial and Temporal Yield Variability Using Crop models. *European Journal of Agronomy*, **18**, 141-158.

- Colbach, N., Forcella, F. and Johnson, G.A. (2000). Spatial and temporal stability of weed populations over five years. *Weed Science*, **48**, 366-377.
- Cressie, N. (1993). *Statistics for Spatial Data*, 2nd edition. New York: Wiley.
- Deen, W., Swanton, C.J. and Hunt, L.A. (2001). A Mechanistic Growth and Development Model of Common Ragweed. *Weed Science*, **49**, 723-731.
- Dieleman, J.A., Mortensen, D.A., Buhler, D.D., Cambardella, C.A., and Moorman, T.B. (2000). Identifying Association Among Site Properties and Weed Species Abundance. *Weed Science*, **48**, 567-575.
- Diggle, P.J., Heagerty, P.J., Liang, K.Y. and Zeger, S.L. (2002). *Analysis of Longitudinal Data*, 2nd Ed. Oxford, U.K.: Oxford University Press.
- Diggle, P.J., Tawn, J.A., and Moyeed, R.A. (1998). Model-based geostatistics (with discussion). *J. Roy. Statist. Soc., Ser. C (Applied Statistics)*, **47**, 299–350.
- Foth, H.D. (1990). *Fundamentals of Soil Science*. New York: John Wiley and Sons.
- Gelfand A.E., Schmidt, A., Banerjee S. and Sirmans C.F. (2004). Nonstationary multivariate process modelling through spatially varying coregionalization (with discussion). *Test*, **13**, 1–50.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004). *Bayesian Data Analysis*. Second Edition. Boca Raton, FL: Chapman and Hall/CRC Press.
- Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer.
- Johnson, G.A., Mortensen, D.A and Martin, A.R. (1995). A Simulation of Herbicide Use Based on Weed Spatial Distribution. *Weed Research*, **35**, 197-205.
- Johnson, G.A., Mortensen, D.A. and Gotway, C.A. (1996). Spatial and Temporal Analysis of Weed Seedling Populations Using Geostatistics. *Weed Science* **44**, 704-710.

- Johnson, G.A. and Huggins, D.R. (1998). Knowledge-based Decision Support Strategies: Linking Spatial and Temporal Components within Site-specific Weed Management. *Journal of Crop Production*, **2**, 225-238.
- Møller, J. and Waagepetersen, R. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Smith, A.B., Cullis, B.R. and Thompson, R. (2001a). Analysing variety by environment data using multiplicative mixed models. *Biometrics*, **57**, 1138-1147.
- Smith, A.B., Cullis, B.R., Appels, R., Campbell, A.W., Cornish, G.B., Martin, D. and Allen, H.M. (2001b). The statistical analysis of quality traits in plant improvement programs with applications to the mapping of milling yield in wheat. *Australian Journal of Agricultural Research* **52**, 1207–1219.
- Stein, M. L. (1999a), *Interpolation of Spatial Data: some theory for kriging*, New York: Springer.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society, Series B*, **64**, 583-639.
- Thompson, J.A. and Bell, J.C. (1996). Color Index for Identifying Hydric Conditions for Seasonally Saturated Mollisols in Minnesota. *Journal of the Soil Science Society of America* **60**, 1979-1988.
- Wackernagel, H. (2003). *Multivariate Geostatistics*. Third Edition. New York: Springer.
- Wilson, S.D. and Tilman, D. (1995). Competitive Responses of Eight Old-field Plant Species in Four Environments. *Ecology*, **76**, 1169-1180.

Model	pD	DIC	Model	pD	DIC	Model	pD	DIC
1a	7	979	2a	22.4	776	3a	23.2	738
1b	21.7	811	2b	23.1	771	3b	24.1	732
1c	22.3	802	2c	23.6	762	3c	26.1	726

Table 1: Model comparisons using the DIC criteria.

Parameters	Estimates: 50% (2.5%,97.5%)
Sand	-0.103 (-0.113,-0.084)
Phosphorus	-0.028 (-0.046,-0.010)
Potassium	-0.002 (-0.004,-0.001)
PDI	0.167 (0.139, 0.192)
Λ_{00}	0.380 (0.284, 0.521)
Λ_{11}	0.190 (0.141, 0.255)
$\Lambda_{01}/\sqrt{\Lambda_{00}\Lambda_{11}}$	-0.565 (-0.691,-0.390)
ϕ_0	0.080 (0.022, 0.387)
ϕ_1	0.021 (0.008, 0.018)
ν_0	1.073 (0.665, 1.594)
ν_1	0.687 (0.386, 1.197)
τ^2	0.002 (0.001, 0.004)

Table 2: Parameter estimates for the fully coregionalized Model 3c.

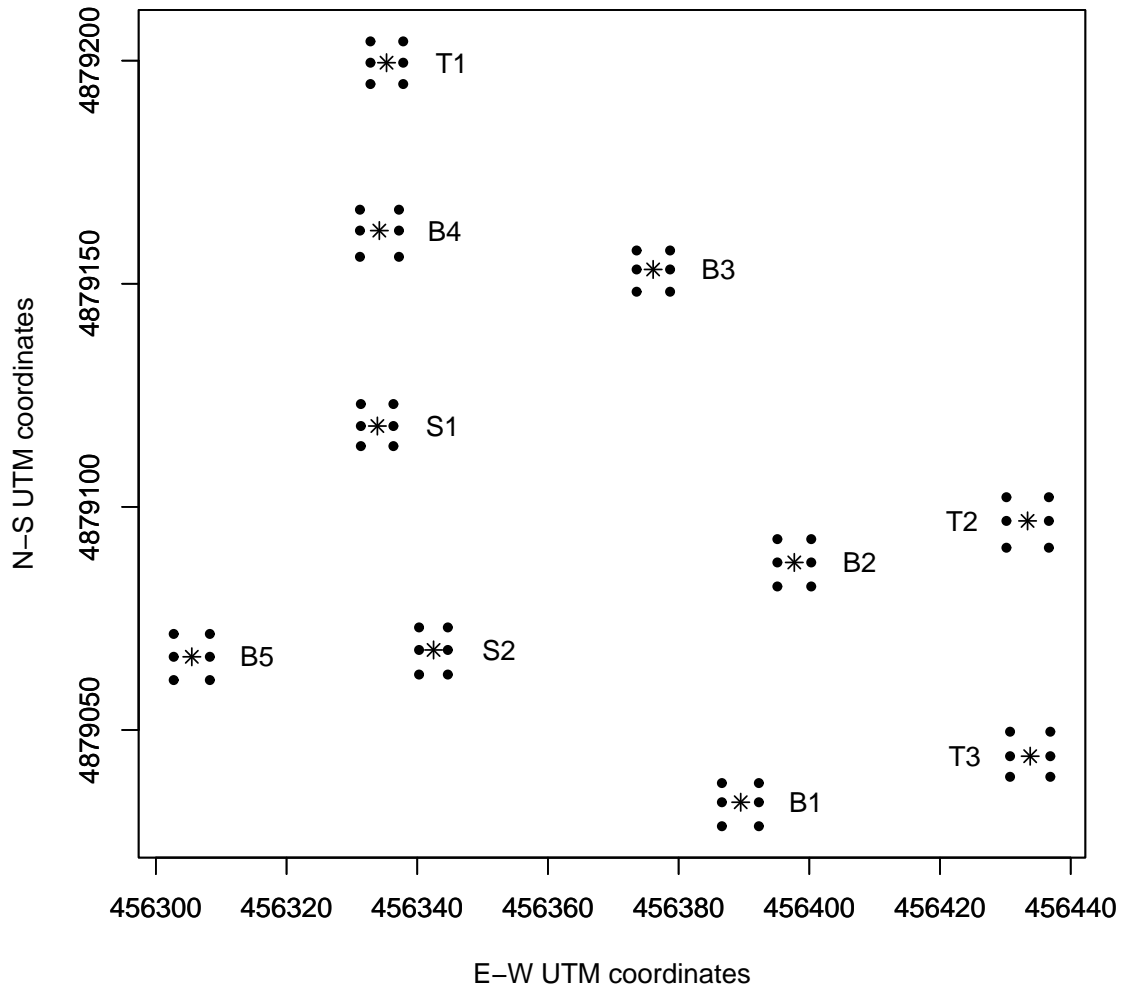
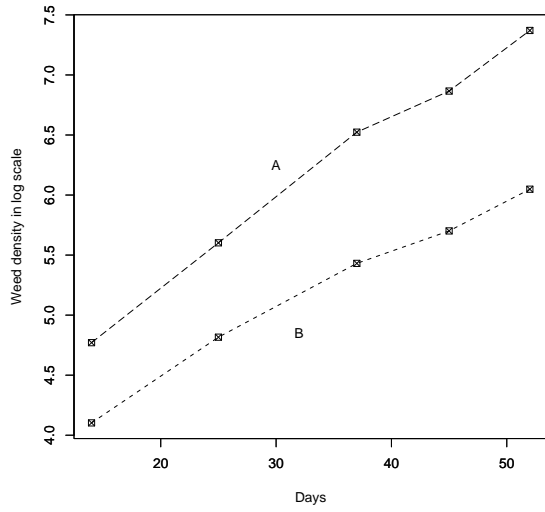
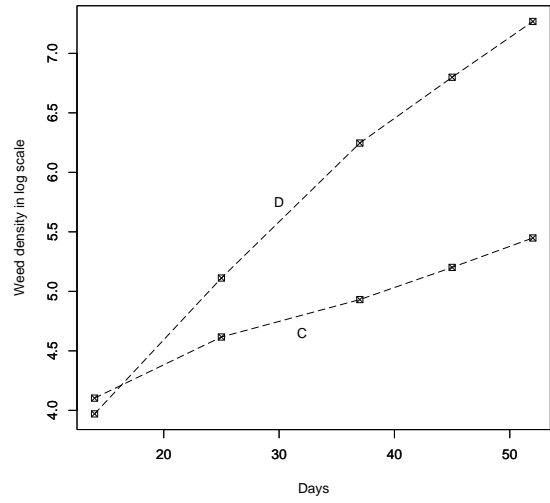


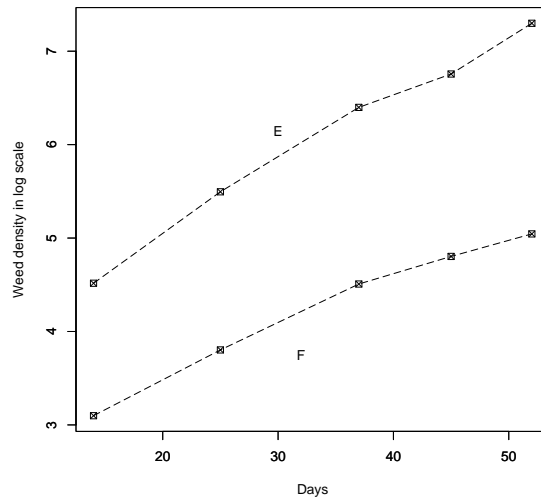
Figure 1: The location of the main plots in Waseca each of which, when zoomed in, is a rectangular 3x2 lattice of spatially arranged subplots (replicates). The centroid of the rectangular lattice is taken as the geocoding for that site. The labels indicate back-slopes (B's), toe-slopes (T's) and summits (S's).



(a)

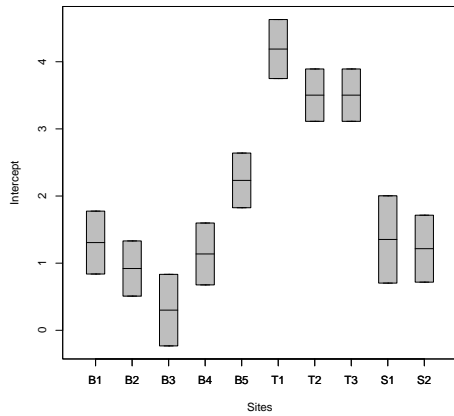


(b)

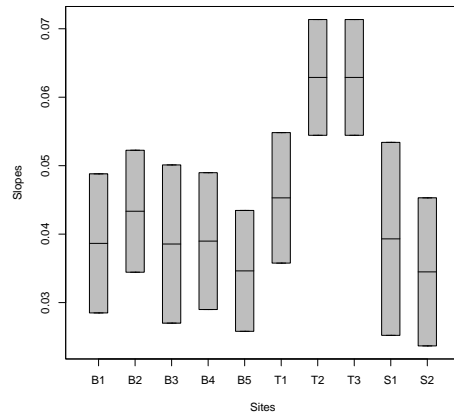


(c)

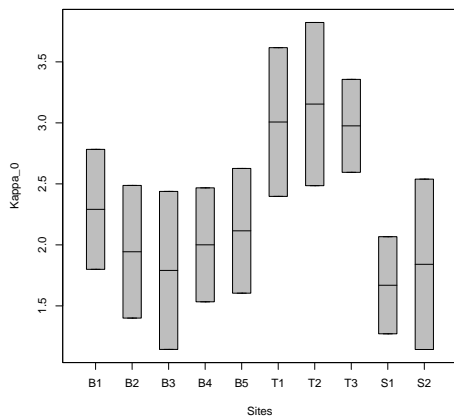
Figure 2: Figure showing variation in growth patterns for raw data. There are six curves labelled A through F. The checked boxes represent time points where the data have been collected.



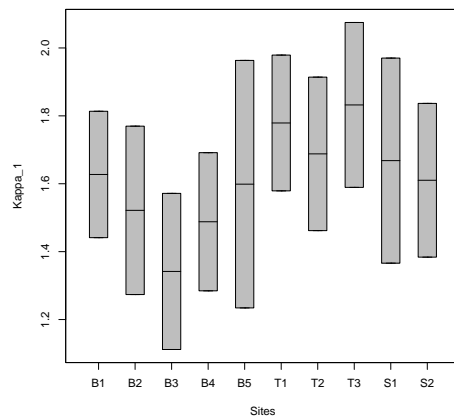
(a)



(b)

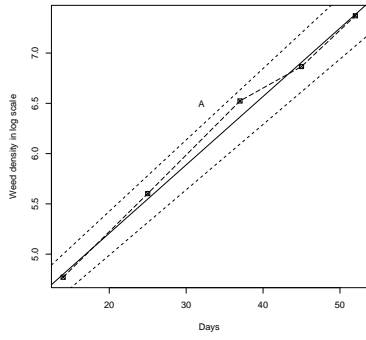


(c)

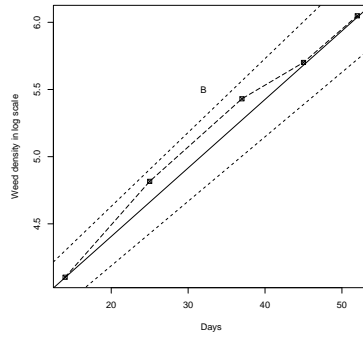


(d)

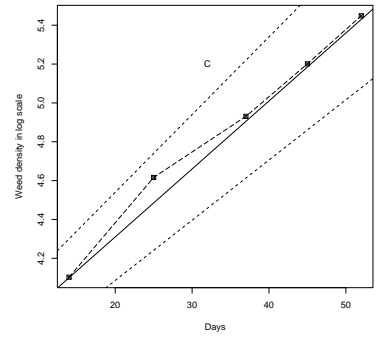
Figure 3: (a) and (b) show the central 95% posterior credible intervals for the site-wise means of intercept and slope processes respectively; (c) and (d) show the same for the micro-level spatial decay parameters.



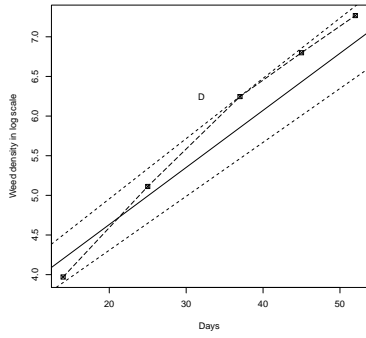
(a)



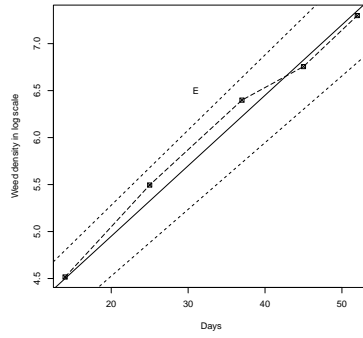
(b)



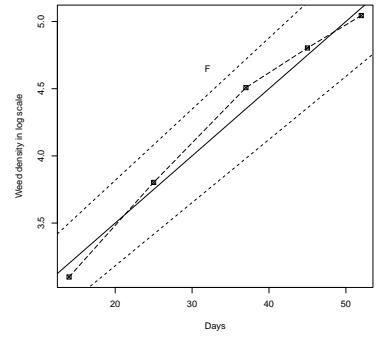
(c)



(d)



(e)



(f)

Figure 4: (a) through (f) shows the posterior predictive fit of the linear spatially-varying growth curves for the graphs labelled A through F in Figure 2.

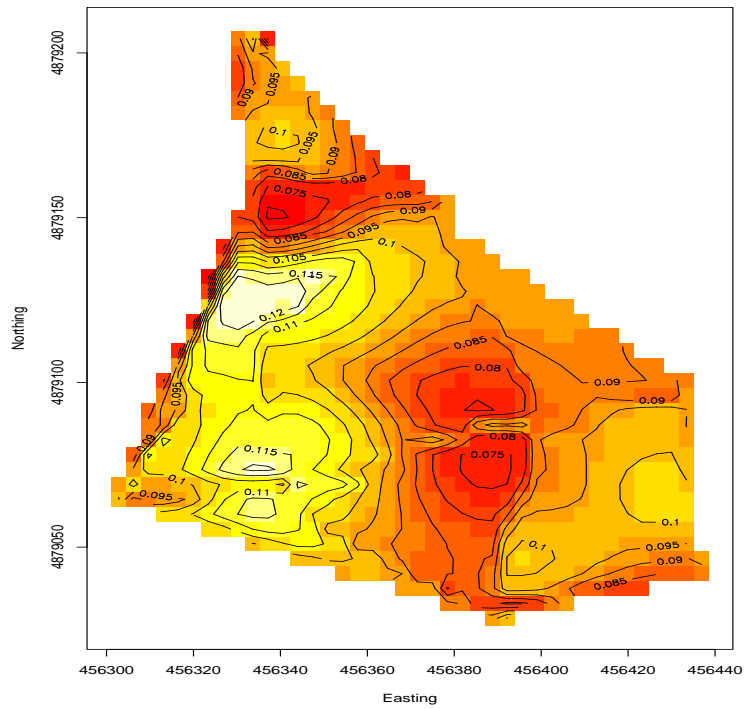
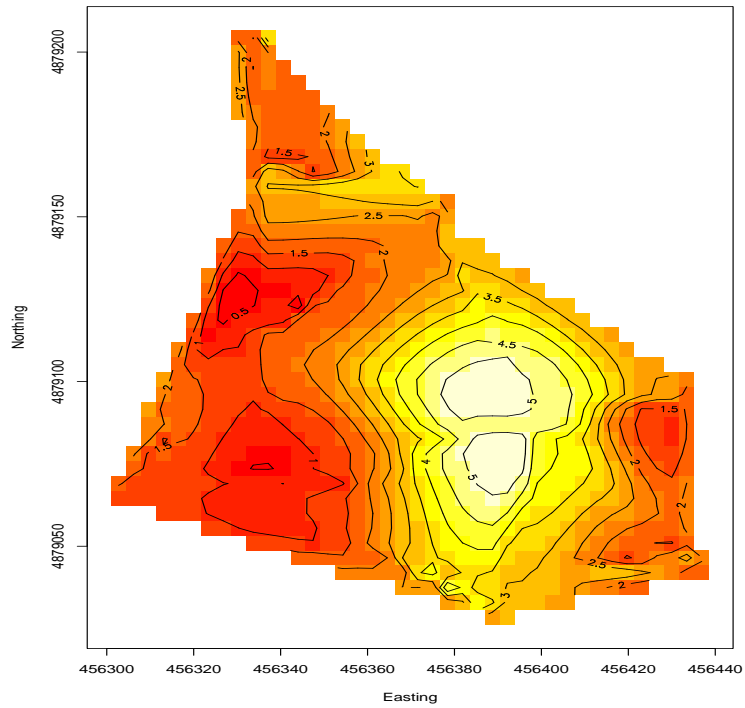


Figure 5: Image plot of the interpolated intercept process (top) and the slope process (bottom) over the field with lighter shades corresponding to higher values.

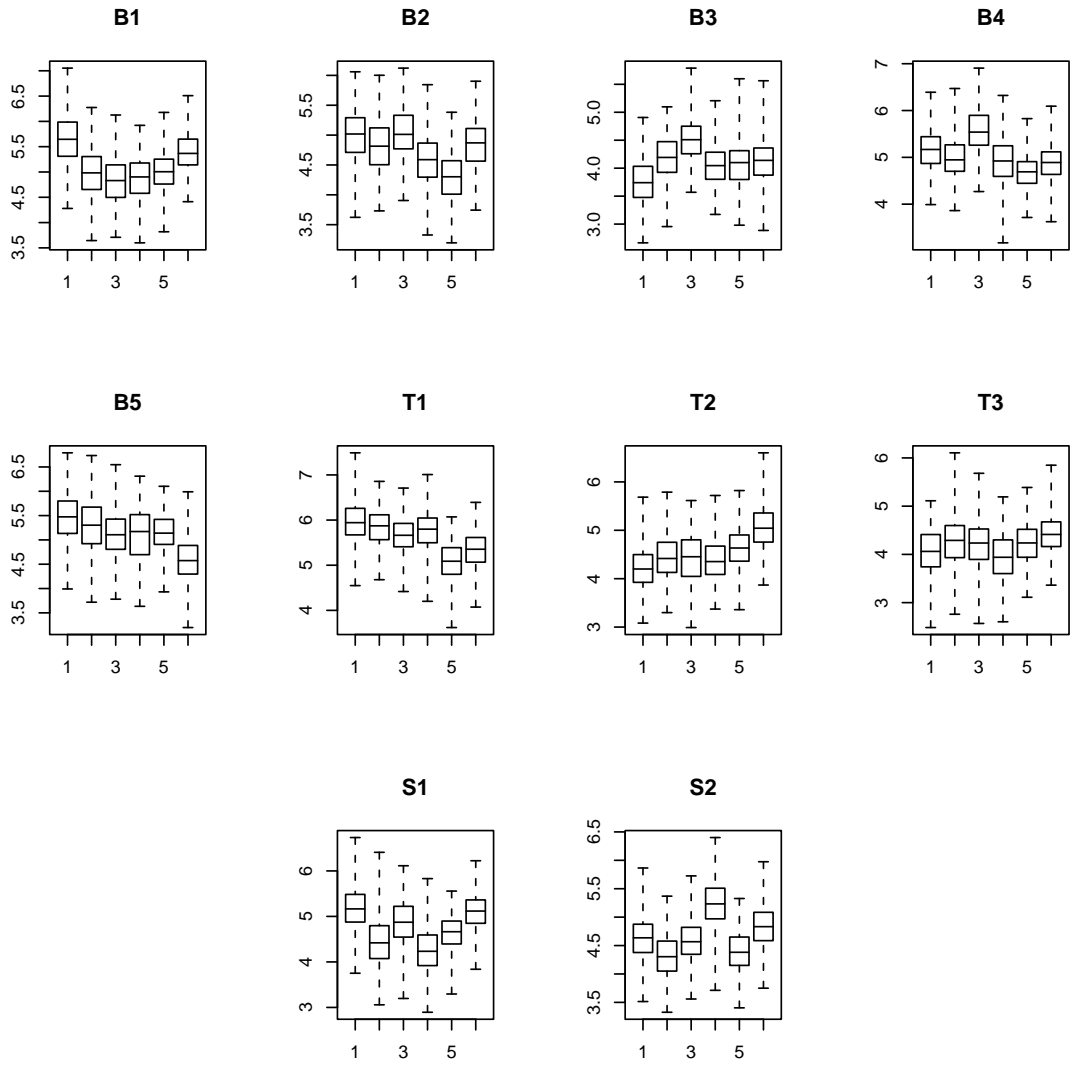


Figure 6: Site-sliced box-plots of the intercept process bringing out the variation in the process between replicates within each site.

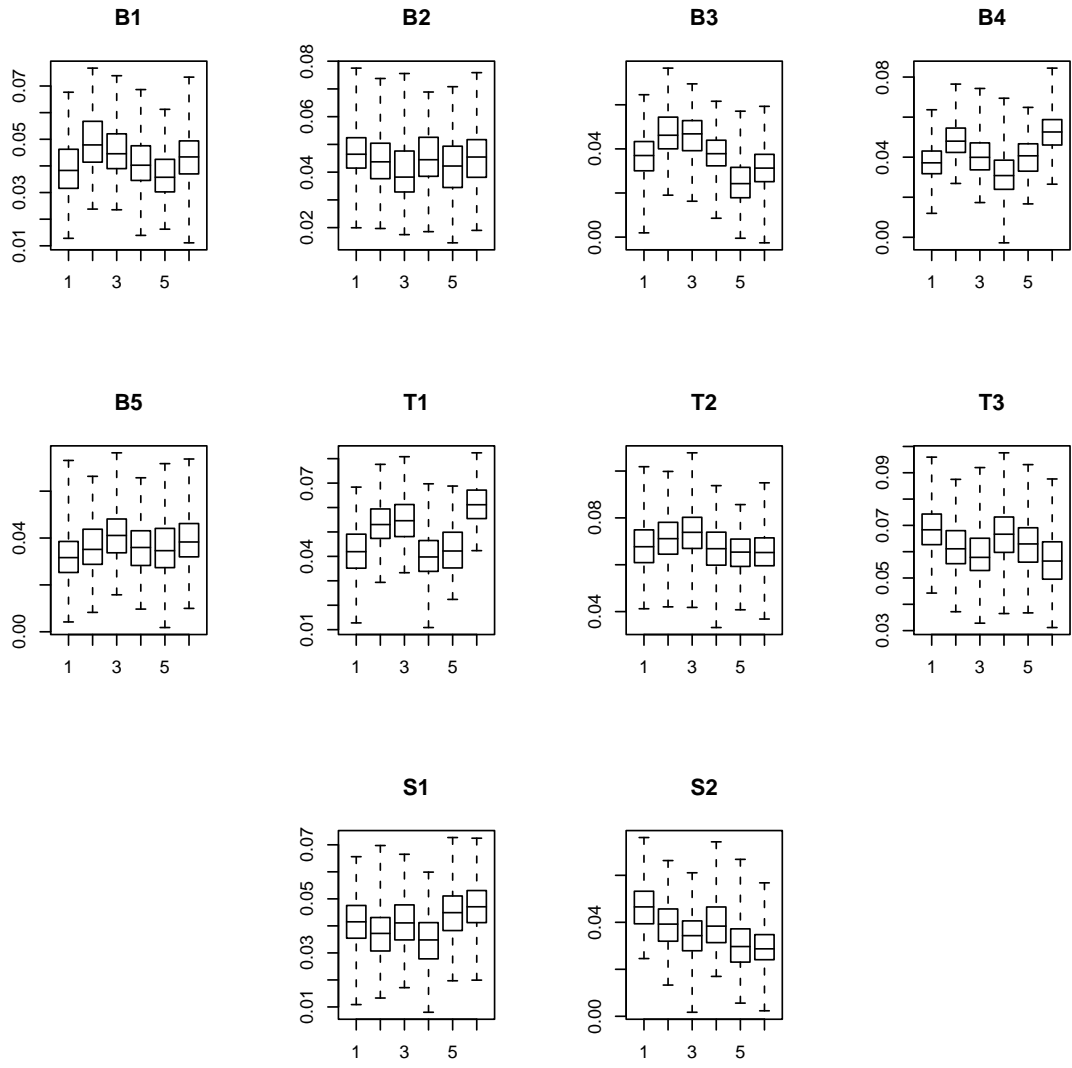


Figure 7: Site-sliced box-plots of the slope process bringing out the variation in the process between replicates within each site.