Statistica Sinica **25** (2015), 205-223
doi:http://dx.doi.org/10.5705/ss.2013.245w

BAYESIAN ANALYSIS OF SPATIALLY-DEPENDENT FUNCTIONAL RESPONSES WITH SPATIALLY-DEPENDENT MULTI-DIMENSIONAL FUNCTIONAL PREDICTORS

Wen-Hsi Yang[1], Christopher K. Wikle[2], Scott H. Holan[2],
D. Brenton Myers[2] and Kenneth A. Sudduth[3]

[1]*CSIRO,* [2]*University of Missouri*
and [3]*USDA-ARS-Cropping Systems and Water Quality Unit*

Abstract: Modeling high-dimensional functional responses utilizing multi-dimensional functional covariates is complicated by spatial and/or temporal dependence in the observations in addition to high-dimensional predictors. To utilize such rich sources of information we develop multi-dimensional spatial functional models that employ low-rank basis function expansions to facilitate model implementation. These models are developed within a hierarchical Bayesian framework that accounts for several sources of uncertainty, including the error that arises from truncating the infinite-dimensional basis function expansions, error in the observations, and uncertainty in the parameters. We illustrate the predictive ability of such a model through a simulation study and an application that considers spatial models of soil electrical conductivity depth profiles using spatially dependent near-infrared spectral images of electrical conductivity covariates.

Key words and phrases: Basis functions, diffuse reflectance spectroscopy, Karhunen-Loève, matrix normal, penetrometer, principal components, soil electrical conductivity.

1. Introduction

With advances in instrumentation such as satellites, sensor networks, data storage tags, and spectroscopes, scientists are often faced with the challenging problem of incorporating extremely high-dimensional covariates into statistical models. The efficient use of such "big data" is the subject of much active research in the statistics and computer science communities. Such problems are compounded when these data are collected over space and/or time, thereby introducing dependence. In cases where the response is also spatial and/or temporal, this dependence may be accounted for through covariates, but may also require temporal and/or spatially-explicit error structures. Another complication arises when the responses themselves are inherently functional (e.g., "curves" in time and/or space). This paper presents methodology that can accommodate

multi-dimensional covariates that vary in space, as well as purely spatial scalar covariates, in the context of modeling spatially-dependent functional responses, with the goal being spatial prediction. The methodology is presented using a hierarchical Bayesian approach to account for uncertainties that arise in the observations, process, and parameters.

The use of two-dimensional functional predictors (i.e., "image predictors") has seen increased utility in statistical models over the last few years (e.g., see Reiss and Ogden (2010); Morris et al. (2011); Holan et al. (2010, 2012); Martinez et al. (2013); Yang et al. (2013)). Holan et al. (2010) showed in the context of insect communication that one could treat a time-frequency representation of a high-frequency nonstationary time signal as a two-dimensional "image" and, with suitable functional dimension reduction and stochastic search variable selection (SSVS), easily incorporate such big data covariates into traditional generalized linear mixed models. This type of modeling was subsequently considered in the context of business cycle estimation (Holan et al. (2012)) and in characterizing spawning success of shovelnose sturgeon by incorporating nonlinear interactions into the model (Yang et al. (2013)). Recently, Martinez et al. (2013) considered a functional mixed model approach to modeling acoustic signals associated with bats. To our knowledge, modeling spatially-correlated functional data with spatially-dependent image predictors has not been considered to date.

The use of spatially-dependent image predictors is compounded when one has responses that are spatially-dependent functions as well. Although functional responses have been considered in the context of image predictors (e.g., Morris et al. (2011)), the spatially-dependent functional case has not been considered. Functional data analysis is fairly mature in statistics (e.g., Bosq (2000); Ramsay and Silverman (2005), among others), yet spatial functional data analysis has just recently become an active sub-field of spatial statistics and functional data analysis. Excellent reviews of recent work in the area can be found in Delicado et al. (2010), Ruiz-Medina (2012a), and Kokoszka (2012). In general, geostatistical predictive and clustering approaches have focused on co-kriging ideas (e.g., Goulard and Voltz (1993); Monestiez and Nerini (2008); Giraldo, Delicado, and Mateu (2010, 2012)) and the general theory of spatial autoregressive and moving average Hilbertian processes (Ruiz-Medina (2011, 2012b); Ruiz-Medina and Espejo (2013)). In addition, a more traditional functional principal components approach to spatially-dependent functions, where interest is on estimation of mean functions, is given in Gromenko et al. (2012) and Gromenko and Kokoszka (2013). These approaches have been from the classical perspective, with relatively few Bayesian implementations. Notable exceptions include Baladandayuthapani et al. (2008) who consider a Bayesian hierarchical model with relatively simple

spatial dependence on the functions at low levels of the hierarchy. This was extended by Zhou et al. (2010) to have a more flexible covariance structure, and was implemented using an expectation-maximization (EM) algorithm.

The contribution of this paper is then the development of methodology for modeling spatially-dependent functional responses in terms of spatially-dependent functional-image predictors, along with spatially-dependent covariates, within a Bayesian paradigm that can account for the uncertainty associated with data, spatial processes, and parameters. Section 2 describes the methodology. A motivating example of estimating electrical conductivity in soils using visible and near-infrared (VNIR) spectral images is given in Section 3, followed by a discussion and concluding remarks in Section 4. For convenience of exposition, a comprehensive description of model choices, simulation study, sensitivity analyses corresponding to the application, full conditional distributions, and details of the sampling algorithm and implementation are provided in an online supplement.

2. Methodology

In this section, we introduce a class of spatially-explicit functional models. Different from traditional spatial models, the proposed models allow for functional responses and functional covariates. In general, these functional covariates can be curves of one dimension, images of two dimensions, or objects of higher dimensions. For simplicity, we focus on two-dimensional image covariates that exhibit spatial dependence, although the method is general with regards to the use of functional objects of higher dimension, and can be easily generalized to account for interactions of functionals (e.g., see Yang et al. (2013)). We note that the use of multiple truncated basis expansions effectively reduces this complicated spatial model to a multivariate multiple mixed-effects regression model, greatly facilitating its implementation. Inference is performed in the Bayesian hierarchical framework, which allows one to directly account for uncertainty associated with observations, functional truncations, and parameters.

2.1. Spatially-dependent functional-image process model

We denote a response functional to be a continuous spatial process $\{Y(\mathbf{s}, d) : \mathbf{s} \in D \subset \Re^2, d \in \mathfrak{D} \subset \Re\}$, where D is a continuous spatial domain and $\mathfrak{D}$ represents a continuous one-dimensional domain such as time or depth. For convenience of exposition, we refer to this dimension as "depth" to coincide with the application presented in Section 3, but note that there are many applications in which this index would correspond to time. Also, denote $X_j(\mathbf{s}, \mathbf{u}_j)$ the jth observed two-dimensional functional covariate, $j = 1, \ldots, J$, at spatial location $\mathbf{s}$ with $\{\mathbf{u}_j = (d, \omega_j)' : d \in \mathfrak{D}, \omega_j \in \Omega\}$ corresponding to the index of the two image dimensions of interest (e.g., depth and wavelength). The j subscript on

$\mathbf{u}_j$ serves to indicate that the functional coordinates may be different for the different covariates. However, to simplify notation, we subsequently assume that the functional coordinates are the same for all "image" covariates and drop the subscript. To give this notation some perspective relative to our application in Section 3, we let Ω correspond to the continuous frequency (or wavelength) domain.

The primary model can be written at location $\mathbf{s}$ and depth d as the relation between the response functional, $Y(\mathbf{s}, d)$, and functional covariates as follows

$$Y(\mathbf{s}, d) = \sum_{j=1}^{J} \int X_j(\mathbf{s}, \mathbf{u}) \beta_j(\mathbf{u}, d) d\mathbf{u} + \boldsymbol{z}'(\mathbf{s}) \boldsymbol{\delta}(d) + \eta(\mathbf{s}, d), \tag{2.1}$$

where $\beta_j(\mathbf{u}, d)$ is a square integrable functional coefficient corresponding to $X_j(\mathbf{s}, \mathbf{u})$, $\boldsymbol{\delta}(d)$ is a n_p-vector of depth-specific regression coefficients associated with the $n_p \times 1$ spatially indexed covariate vector $\boldsymbol{z}(\mathbf{s})$, and $\eta(\mathbf{s}, d)$ is a mean zero random process capturing spatial and depth dependence. Typically, we specify $\eta(\mathbf{s}, d)$ to be a Gaussian process with covariance function $C(\mathbf{s}, \mathbf{s}'; d, d') \equiv \text{cov}(\eta(\mathbf{s}, d), \eta(\mathbf{s}', d'))$. In a traditional spatial analysis, one might consider a three-dimensional spatial covariance function if d corresponds to the vertical dimension, or a spatio-temporal covariance function if d corresponds to time (e.g., Cressie and Wikle (2011)). Here, we deliberately keep these indices separate given that the response and covariate vary functionally in the dimension d, and, motivated by the application in Section 3, this depth dimension is assumed to operate on a different scale of variability relative to the two-dimensional horizontal spatial component of the process. That is, in the application that motivates this work, it is reasonable to assume separability between the horizontal and vertical spatial dimensions. This assumption can be relaxed, however, as discussed in Section 2.1.1. Different from traditional geostatistics concerning scalar or vector variables at location $\mathbf{s}$, our methodology considers that the response variable of interest is a functional (curve) in depth (or time) and some of the covariates are also functionals. In this sense, the function $\beta_j(\mathbf{u}, d)$ can be thought of as a kernel that "distributes" the covariate $X_j(\mathbf{s}, \mathbf{u})$ to the depth of the response variable at location $\mathbf{s}$.

We now derive a basis function expansion representation for the model in (2.1). First, assume that $\{\phi_{jk}(\mathbf{u}) : k = 1, 2, \ldots\}$ form a complete orthonormal basis corresponding to the jth functional covariate. Then, we have the unique representation of the functional covariate $X_j(\mathbf{s}, \mathbf{u}) = \sum_{k=1}^{\infty} \xi_{jk}(\mathbf{s}) \phi_{jk}(\mathbf{u})$, where $\xi_{jk}(\mathbf{s})$ are expansion coefficient functions (for a given location $\mathbf{s}$) associated with the jth functional covariate. By considering the same basis, we also have $\beta_j(\mathbf{u}, d) = \sum_{k=1}^{\infty} b_{jk}(d) \phi_{jk}(\mathbf{u})$, where $b_{jk}(d)$ are expansion coefficient functions

(for a given depth d) associated with the jth square integrable function. Substituting these expressions into (2.1) and making use of the orthogonality, we obtain

$$Y(\mathbf{s},d) = \sum_{j=1}^{J}\sum_{k=1}^{\infty} \xi_{jk}(\mathbf{s}) b_{jk}(d) + \boldsymbol{z}'(\mathbf{s})\boldsymbol{\delta}(d) + \eta(\mathbf{s},d). \tag{2.2}$$

In addition, assume that $\{\psi_i(d) : i = 1, 2, \ldots\}$ form a complete orthonormal basis corresponding to depth. Then, we have the unique representations $Y(\mathbf{s},d) = \sum_{i=1}^{\infty} \alpha_i(\mathbf{s})\psi_i(d)$, $b_{jk}(d) = \sum_{i=1}^{\infty} b_{jki}\psi_i(d)$, and $\eta(\mathbf{s},d) = \sum_{i=1}^{\infty} \theta_i(\mathbf{s})\psi_i(d)$, where $\alpha_i(\mathbf{s})$ and $\theta_i(\mathbf{s})$ are expansion coefficient functions (of $\mathbf{s}$) corresponding to $Y(\mathbf{s},d)$ and $\eta(\mathbf{s},d)$, respectively, and b_{jki} are expansion coefficients associated with $b_{jk}(d)$. Replacing $Y(\mathbf{s},d)$, $b_{jk}(d)$, and $\eta(\mathbf{s},d)$ in (2.2) with these expansions, we can rewrite (2.2) as

$$\sum_{i=1}^{\infty} \alpha_i(\mathbf{s})\psi_i(d) = \sum_{j=1}^{J}\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \xi_{jk}(\mathbf{s}) b_{jki}\psi_i(d) + \boldsymbol{z}'(\mathbf{s})\boldsymbol{\delta}(d) + \sum_{i=1}^{\infty} \theta_i(\mathbf{s})\psi_i(d), \tag{2.3}$$

where we could additionally write the pth element of $\boldsymbol{\delta}(d)$ as $\delta_p(d) = \sum_{i=1}^{\infty} h_{pi}\psi_i(d)$, depending on whether it was determined appropriate to view the depth response to the spatial covariates, $\boldsymbol{z}(\mathbf{s})$, as functions.

Finally, assume that $\{w_\ell(\mathbf{s}) : \ell = 1, 2, \ldots\}$ form a complete orthonormal basis corresponding to $\mathbf{s}$, which then gives the unique representations $\alpha_i(\mathbf{s}) = \sum_{\ell=1}^{\infty} w_\ell(\mathbf{s}) a_{i\ell}$, $\xi_{jk}(\mathbf{s}) = \sum_{\ell=1}^{\infty} w_\ell(\mathbf{s}) f_{jk\ell}$, and $\theta_i(\mathbf{s}) = \sum_{\ell=1}^{\infty} w_\ell(\mathbf{s}) g_{i\ell}$, where $a_{i\ell}$, $f_{jk\ell}$, and $g_{i\ell}$ are the expansion coefficients associated with $\alpha_i(\mathbf{s})$, $\xi_{jk}(\mathbf{s})$, and $\theta_i(\mathbf{s})$, respectively. As with the $\boldsymbol{\delta}(d)$ coefficients, if one is interested in considering the spatial predictors as functionals, then it would be appropriate to expand the pth element of $\boldsymbol{z}(\mathbf{s})$ as $z_p(\mathbf{s}) = \sum_{\ell=1}^{\infty} w_\ell(\mathbf{s}) q_{p\ell}$. Substituting the above expansions into (2.3), we obtain the representation

$$\begin{aligned}\sum_{i=1}^{\infty}\sum_{\ell=1}^{\infty} w_\ell(\mathbf{s})\psi_i(d) a_{i\ell} = &\sum_{j=1}^{J}\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\sum_{\ell=1}^{\infty} w_\ell(\mathbf{s})\psi_i(d) f_{jk\ell} b_{jki} + \boldsymbol{z}'(\mathbf{s})\boldsymbol{\delta}(d) \\ &+ \sum_{i=1}^{\infty}\sum_{\ell=1}^{\infty} w_\ell(\mathbf{s})\psi_i(d) g_{i\ell}. \end{aligned} \tag{2.4}$$

From a functional data analysis perspective, using the spatial basis expansions to accommodate spatial structure is quite reasonable. However, one could alternatively consider a more traditional spatial co-kriging approach (see the references in Section 1). We choose the basis expansion formulation here because our application in Section 3 is concerned with fairly smooth functional spatial surfaces, which, along with the potential for future "big data" applications, is facilitated by the use of rank-reduced spatial models (e.g., see the review in Wikle (2010)).

Similar to applications in traditional functional data analysis, one can consider finite approximations to the infinite summations in (2.4), e.g.,

$$\sum_{i=1}^{n_i}\sum_{\ell=1}^{n_\ell} w_\ell(\mathbf{s})\psi_i(d)a_{i\ell} = \sum_{j=1}^{J}\sum_{k=1}^{n_{k_j}}\sum_{i=1}^{n_i}\sum_{\ell=1}^{n_\ell} w_\ell(\mathbf{s})\psi_i(d)f_{jk\ell}b_{jki} + \boldsymbol{z}'(\mathbf{s})\boldsymbol{\delta}(d) + \sum_{i=1}^{n_i}\sum_{\ell=1}^{n_\ell} w_\ell(\mathbf{s})\psi_i(d)g_{i\ell}. \tag{2.5}$$

In practice, the truncations n_{k_j}, n_i, and n_ℓ are typically problem specific and can be chosen based on percent of variance explained, cross-validation, and/or sensitivity analysis. Critically, with a hierarchical Bayesian implementation (see Section 2.2), one can account for potential truncation and observation error correspond to the response and covariate functionals. Given these truncations, we denote the basis vectors associated with depth and spatial location by $\boldsymbol{\psi}(d) \equiv [\psi_1(d), \ldots, \psi_{n_i}(d)]'$ and $\boldsymbol{w}(\mathbf{s}) \equiv [w_1(\mathbf{s}), \ldots, w_{n_\ell}(\mathbf{s})]'$, respectively. Then, we can rewrite (2.5) in matrix form as

$$\boldsymbol{w}'(\mathbf{s})\boldsymbol{A}\,\boldsymbol{\psi}(d) = \sum_{j=1}^{J}\boldsymbol{w}'(\mathbf{s})\boldsymbol{F}_j\mathbf{B}_j\boldsymbol{\psi}(d) + \boldsymbol{z}'(\mathbf{s})\boldsymbol{\delta}(d) + \boldsymbol{w}'(\mathbf{s})\boldsymbol{G}\,\boldsymbol{\psi}(d), \tag{2.6}$$

where $\boldsymbol{A}$ and $\boldsymbol{G}$ are $n_\ell \times n_i$ random matrices (see below) with elements $\{a_{\ell i}\}$ and $\{g_{\ell i}\}$, respectively, $\boldsymbol{F}_j$ is an $n_\ell \times n_{k_j}$ matrix with elements $\{f_{j\ell k}\}$, and $\mathbf{B}_j$ is an $n_{k_j} \times n_i$ matrix with elements $\{b_{jki}\}$, for $i = 1, \ldots, n_i$, $k = 1, \ldots, n_{k_j}$, and $\ell = 1, \ldots, n_\ell$. Thus, $\mathbf{A}$ is related to the functional response, $\mathbf{G}$ the spatial error process, $\mathbf{F}_j$ the image covariates, and $\mathbf{B}_j$ the "regression" coefficients associated with the image covariates.

In practice one is typically interested in prediction at a specific set of spatial locations and depths, say $\{\mathbf{s}_1, \ldots, \mathbf{s}_{n_s}\}$ and $\{d_1, \ldots, d_{n_d}\}$, respectively. In this case, we denote the $n_i \times n_d$ depth basis matrix as $\boldsymbol{\Psi} \equiv [\boldsymbol{\psi}(d_1), \ldots, \boldsymbol{\psi}(d_{n_d})]$ and the $n_s \times n_l$ spatial basis matrix as $\mathbf{W} \equiv [\boldsymbol{w}(\mathbf{s}_1), \ldots, \boldsymbol{w}(\mathbf{s}_{n_s})]'$. Consequently, we can write (2.6) for these specific locations and depths as

$$\mathbf{W}\mathbf{A}\boldsymbol{\Psi} = \sum_{j=1}^{J}\mathbf{W}\mathbf{F}_j\mathbf{B}_j\boldsymbol{\Psi} + \mathbf{Z}\Delta + \mathbf{W}\mathbf{G}\boldsymbol{\Psi}, \tag{2.7}$$

with $n_s \times n_p$ matrix $\mathbf{Z} \equiv [\boldsymbol{z}(\mathbf{s}_1), \ldots, \boldsymbol{z}(\mathbf{s}_{n_s})]'$ and $n_p \times n_d$ matrix $\Delta \equiv [\boldsymbol{\delta}(d_1), \ldots, \boldsymbol{\delta}(d_{n_d})]$. Assuming that $\mathbf{W}$ and $\boldsymbol{\Psi}$ are known orthogonal matrices of rank n_ℓ and n_i, respectively (which is true if $n_d > n_i$ and $n_s > n_\ell$), it follows that $\mathbf{W}'\mathbf{W} = \mathbf{I}_{n_\ell}$ and $\boldsymbol{\Psi}\boldsymbol{\Psi}' = \mathbf{I}_{n_i}$ and we can simplify (2.7) to

$$\mathbf{A} = \sum_{j=1}^{J}\mathbf{F}_j\mathbf{B}_j + \mathbf{W}'\mathbf{Z}\Delta\boldsymbol{\Psi}' + \mathbf{G}.$$

In the case mentioned above where we treat $\boldsymbol{z}(\mathbf{s})$ and $\boldsymbol{\delta}(d)$ as functionals expanded in terms of spatial and depth basis functions, this further reduces to

$$\mathbf{A} = \sum_{j=1}^{J} \mathbf{F}_j \mathbf{B}_j + \mathbf{QH} + \mathbf{G}, \tag{2.8}$$

where $\mathbf{Z} = \mathbf{WQ}$, $\mathbf{Q}$ is an $n_l \times n_p$ matrix of expansion coefficients associated with the spatial covariates, and $\mathbf{H} \equiv \Delta\boldsymbol{\Psi}'$ is an $n_p \times n_i$ matrix of associated expansion coefficients associated with the depth-specific regression coefficients. Consequently, one advantage of our approach is that by using multiple reduced-rank basis expansions, our model effectively reduces to a multivariate multiple mixed-effects regression model with response matrix $\mathbf{A}$, covariates $\mathbf{F}_j$ and $\mathbf{Q}$, and parameters $\mathbf{B}_j$, where $\mathbf{G}$ is a random error matrix.

2.1.1. Error distributions

As previously described, $\eta(\mathbf{s}, d)$ is a mean-zero spatial Gaussian random process that, in the application that motivates our work (Section 3), is reasonably assumed to be separable in the horizontal and vertical dimensions. A convenient way to represent a matrix variate separable Gaussian process is via a matrix normal distribution, as introduced by Dawid (1981). That is, generally, if an $m \times t$ matrix $\mathbf{U}$ follows a matrix normal distribution, it can be expressed as $\mathbf{U} \sim N_{m,t}(\mathbf{L}, \boldsymbol{\Sigma}_m, \boldsymbol{\Sigma}_t)$, where $\mathbf{L}$ is an $m \times t$ matrix of mean values, $\boldsymbol{\Sigma}_m$ is an $m \times m$ covariance matrix between rows, and $\boldsymbol{\Sigma}_t$ is an $t \times t$ covariance matrix between columns. This can also be written as a multivariate normal distribution with $\text{vec}(\mathbf{U}) \sim MVN(\text{vec}(\mathbf{L}), \boldsymbol{\Sigma}_t \otimes \boldsymbol{\Sigma}_m)$, where $\text{vec}(\mathbf{U})$ is the vectorization of a matrix $\mathbf{U}$, and one can see by the Kronecker product that the standard matrix normal representation implies a separable covariance structure between the matrix row and column variables. In addition, matrix normal distributions have the property that linear transformations of $\mathbf{U}$ still follow a matrix normal distribution; i.e., $\mathbf{MUT}' \sim N_{m,t}(\mathbf{MLT}', \mathbf{M}\boldsymbol{\Sigma}_m\mathbf{M}', \mathbf{T}\boldsymbol{\Sigma}_t\mathbf{T}')$.

Therefore, given that $\eta(\mathbf{s}, d)$ is assumed to be a separable and mean-zero Gaussian random process, $\mathbf{G}$ in (2.8) has a matrix normal distribution after discretization via basis functions. In our model, we then note that $\mathbf{G} \sim N_{n_\ell, n_i}(\mathbf{0}, \boldsymbol{\Sigma}_{n_\ell}, \boldsymbol{\Sigma}_{n_i})$, where $\boldsymbol{\Sigma}_{n_\ell} = \mathbf{W}'\boldsymbol{C}_\mathbf{s}\mathbf{W}$ and $\boldsymbol{\Sigma}_{n_i} = \boldsymbol{\Psi}\boldsymbol{C}_d\boldsymbol{\Psi}'$. In this case, $\boldsymbol{C}_d$ and $\boldsymbol{C}_\mathbf{s}$ are $n_d \times n_d$ and $n_s \times n_s$ depth and spatial covariance matrices, respectively. These can be specified, as is typical in geostatistics, according to some valid spatial covariance function (e.g., a Matérn model) or empirically, as in functional principal components analysis. Alternatively, in the Bayesian paradigm $\boldsymbol{\Sigma}_{n_\ell}$ and $\boldsymbol{\Sigma}_{n_i}$ can be assigned prior distributions (e.g., inverse-Wishart distributions) to increase model flexibility. Note that a non-separable covariance structure could be specified between depth and space if desired, but given that the complicated

joint dependence is likely accommodated by the functional covariates, the separable structure on the errors is reasonable for most applications. Certainly, the added simplicity of the separable structure makes it appealing in this complex modeling framework, especially for "big data" applications.

Upon estimation of the parameters, one can then perform prediction of the response at any location and depth by utilizing the appropriate basis expansions to obtain the posterior predictive distribution of the functional response at desired spatial locations. Importantly, one must have basis functions that are defined at any desired spatial and depth location. In addition, one should account for the uncertainty associated with the observations of the response and covariates, as well as the uncertainty associated with the basis truncations as they apply to the data. We prefer to approach this problem using the hierarchical Bayesian paradigm.

2.2. Hierarchical representation

As summarized in Cressie and Wikle (2011), the hierarchical Bayesian framework can facilitate the quantification of uncertainty in complicated spatial and spatio-temporal models. Here we consider it to accommodate the functional model presented in Section 2.1 in a way that uncertainty in the observed quantities, the basis functions and the parameters are accounted for.

2.2.1. Data models

We consider the possibility that the spatial locations at which we have observations are not the same as those at which we wish to predict. We start with two-dimensional spatial observations at locations $\{\mathbf{r}_1, \ldots, \mathbf{r}_{n_r}\}$ which could be a subset of the predictions prediction $\{\mathbf{s}_1, \ldots, \mathbf{s}_{n_s}\}$. We could easily make a similar assumption for the depths, but in the application in Section 3 we predict at the same depths at which we have observations and so avoid the extra cumbersome notation of allowing for different depth observation locations.

For the functional response variables, we have $j = 1, \ldots, n_r$ observations $\widetilde{\mathbf{y}}(\mathbf{r}_j) \equiv [\widetilde{y}(\mathbf{r}_j, d_1), \ldots, \widetilde{y}(\mathbf{r}_j, d_{n_d})]'$, which can be written in terms of the $n_r \times n_d$ matrix $\widetilde{\mathbf{Y}} \equiv [\widetilde{\mathbf{y}}(\mathbf{r}_1), \ldots, \widetilde{\mathbf{y}}(\mathbf{r}_{n_r})]'$. This is then related to the "true" $n_s \times n_d$ response matrix $\mathbf{Y} = [\mathbf{Y}(\mathbf{s}_1), \ldots, \mathbf{Y}(\mathbf{s}_{n_s})]'$, where $\mathbf{Y}(\mathbf{s}_i) = [Y(\mathbf{s}_i, d_1), \ldots, Y(\mathbf{s}_i, d_{n_d})]'$. We consider the matrix data model

$$\widetilde{\mathbf{Y}} = \mathbf{K}_{(y)}\mathbf{Y} + \mathbf{E}_{m,y} = \mathbf{K}_{(y)}\mathbf{W}\mathbf{A}\boldsymbol{\Psi} + \mathbf{E}_{t,y} + \mathbf{E}_{m,y} = \mathbf{K}_{(y)}\mathbf{W}\mathbf{A}\boldsymbol{\Psi} + \mathbf{E}_{\widetilde{y}},$$

where $\mathbf{K}_{(y)}$ is an $n_r \times n_s$ observation matrix (typically, an incidence matrix of 1's and 0's, but can accommodate change of spatial support as summarized in Cressie and Wikle (2011)), $\mathbf{E}_{m,y}$ is an $n_r \times n_d$ measurement error matrix, $\mathbf{E}_{t,y}$

is an $n_r \times n_d$ truncation error matrix, and $\mathbf{W}$, $\mathbf{A}$, and $\mathbf{\Psi}$ were defined previously. In this case, $\mathbf{A}$ corresponds to the process component specified at the next level of the hierarchy according to the model in (2.8). The main assumption here is that we are interested in doing prediction on the smooth functionals ($\mathbf{WA\Psi}$) and so the truncation error is accommodated in the covariance structure of $\mathbf{E}_{\widetilde{y}} = \mathbf{E}_{t,y} + \mathbf{E}_{m,y}$. In general, we consider the matrix normal error distribution, $\mathbf{E}_{\widetilde{y}} \sim N_{n_r,n_d}(\mathbf{0}, \mathbf{\Sigma}_{y,n_r}, \mathbf{\Sigma}_{y,n_d})$. In practice, typically one must have fairly reliable prior information concerning these error structures, as they are not otherwise identifiable (for example, see Section 2.2.2 and our motivating application and associated sensitivity analysis in Sections 3 and the online supplement, respectively).

The functional image covariates are also measured at locations $\{\mathbf{r}_1, \ldots, \mathbf{r}_{n_r}\}$, and we denote them by the n_u-vectors, $\widetilde{\mathbf{x}}(\mathbf{r}_i)$ for $i = 1, \ldots, n_r$, which is an n_u-dimensional vectorization of the $n_d \times n_\omega$ observed image. We then denote the $n_r \times n_u$ data matrix $\widetilde{\mathbf{X}}_j$, with each row given by $\widetilde{\mathbf{x}}(\mathbf{r}_i)$ for $i = 1, \ldots, n_r$, and let the corresponding latent true image be denoted by the $n_s \times n_u$ matrix $\mathbf{X}_j$. We then consider the data model

$$\widetilde{\mathbf{X}}_j = \mathbf{K}_{(x_j)}\mathbf{X}_j + \mathbf{E}_{m,x_j} = \mathbf{K}_{(x_j)}\mathbf{W}\mathbf{F}_j\mathbf{\Phi} + \mathbf{E}_{t,x_j} + \mathbf{E}_{m,x_j} = \mathbf{K}_{(x_j)}\mathbf{W}\mathbf{F}_j\mathbf{\Phi} + \mathbf{E}_{\tilde{x}_j},$$

where $\mathbf{K}_{(x_j)}$ is an $n_r \times n_s$ observation matrix, $\mathbf{\Phi}$ is an $n_{k_j} \times n_u$ matrix of basis functions for the image covariates, $\mathbf{E}_{m,x_j}$ is an $n_r \times n_u$ measurement error matrix, $\mathbf{E}_{t,x_j}$ is an $n_r \times n_u$ truncation error matrix, and $\mathbf{W}$ and $\mathbf{F}_j$ were defined in Section 2.1. The crucial portion of the covariates, $\mathbf{F}_j$, are given prior distributions at the next level of the hierarchy. As with the response, we consider the $\mathbf{E}_{\tilde{x}_j}$ error matrix as the sum of the truncation and measurement error matrices and assume this follows a matrix normal, $\mathbf{E}_{\tilde{x}_j} \sim N_{n_r,n_u}(\mathbf{0}, \mathbf{\Sigma}_{x_j,n_r}, \mathbf{\Sigma}_{x_j,n_u})$, with specific choices depending on the application (e.g., see Section 2.2.2, Section 3, and the online supplement).

Consider then the spatial covariates associated with $\boldsymbol{z}(\mathbf{s})$, which are also observed at spatial locations $\{\mathbf{r}_1, \ldots, \mathbf{r}_{n_r}\}$. In particular, consider observations as a $n_r \times n_p$ matrix $\widetilde{\mathbf{Z}}$, where the ith row is given by $\boldsymbol{z}'(\mathbf{r}_i)$, a n_p-vector associated with the p covariates at site $\mathbf{r}_i$. Given the $n_s \times n_p$ true covariate matrix of interest, $\mathbf{Z}$, we have the data model

$$\widetilde{\mathbf{Z}} = \mathbf{K}_{(z)}\mathbf{Z} + \mathbf{E}_{m,z} = \mathbf{K}_{(z)}\mathbf{W}\mathbf{Q} + \mathbf{E}_{t,z} + \mathbf{E}_{m,z} = \mathbf{K}_{(z)}\mathbf{W}\mathbf{Q} + \mathbf{E}_{\widetilde{z}},$$

where $\mathbf{K}_{(z)}$ is an $n_r \times n_s$ observation matrix as above, $\mathbf{Q}$ and $\mathbf{W}$ were defined previously, and $\mathbf{E}_{m,z}$ and $\mathbf{E}_{t,z}$ are the measurement and truncation error matrices, respectively. As before, $\mathbf{E}_{\widetilde{z}}$ corresponds to the sum of these two matrices and is assigned the matrix normal distribution $\mathbf{E}_{\widetilde{z}} \sim N_{n_r,n_p}(\mathbf{0}, \mathbf{\Sigma}_{z,n_r}, \mathbf{\Sigma}_{z,n_p})$, with

specific choices depending on the application (e.g., see Section 2.2.2, Section 3, and the online supplement).

2.2.2. Prior distributions

We specify prior distributions that provide a tradeoff between computational convenience (given the "big data" nature of the applications of this methodology) and model flexibility. However, as in any complex Bayesian hierarchical model, the choice of prior distributions is subjective and can typically be modified if suggested by a particular application. Specific choices for our application and sensitivity analyses are discussed in the online supplement.

Consider the data model error covariances given in $\mathbf{E}_{\widetilde{y}} \sim N_{n_r,n_d}(\mathbf{0}, \mathbf{\Sigma}_{y,n_r}, \mathbf{\Sigma}_{y,n_d}), \mathbf{E}_{\widetilde{x_j}} \sim N_{n_r,n_u}(\mathbf{0}, \mathbf{\Sigma}_{x_j,n_r}, \mathbf{\Sigma}_{x_j,n_u})$, and $\mathbf{E}_{\widetilde{z}} \sim N_{n_r,n_p}(\mathbf{0}, \mathbf{\Sigma}_{z,n_r}, \mathbf{\Sigma}_{z,n_p})$. As mentioned previously, in hierarchical spatial and spatio-temporal models, it is typically difficult to identify such general covariance structures in data models given fairly complex dependence structures in the process model. Thus, it is common in Bayesian hierarchical spatial and spatio-temporal analyses to assume that the data model error covariance matrices are diagonal (e.g., Cressie and Wikle (2011)). This is generally a reasonable assumption and is important for computational tractability in "big data" applications. We assume $\mathbf{\Sigma}_{y,n_r} = \text{diag}(\boldsymbol{\tau}_{y,n_r})$, $\mathbf{\Sigma}_{y,n_d} = \text{diag}(\boldsymbol{\tau}_{y,n_d})$, $\mathbf{\Sigma}_{x_j,n_r} = \text{diag}(\boldsymbol{\tau}_{x_j,n_r})$, $\mathbf{\Sigma}_{x_j,n_u} = \text{diag}(\boldsymbol{\tau}_{x_j,n_u})$, $\mathbf{\Sigma}_{z,n_r} = \text{diag}(\boldsymbol{\tau}_{z,n_r})$ and $\mathbf{\Sigma}_{z,n_p} = \text{diag}(\boldsymbol{\tau}_{z,n_p})$. We could then specify conjugate (inverse gamma) priors for these variance components. However, this still generally requires fairly strong prior information in order to identify these components. As described in the online supplement, we do have prior information available and, thus, further simplify these prior distributions accordingly.

Recall that the process model was given by (2.8). In the Bayesian paradigm, the finite expansion coefficient matrices $\mathbf{F}_j$ and $\mathbf{Q}$ are random matrices and we assume that they follow matrix normal distributions given by $\mathbf{F}_j \sim N_{n_\ell,n_{k_j}}(\boldsymbol{\mu}_{F_j}, \mathbf{\Sigma}_{F_j,n_\ell}, \mathbf{\Sigma}_{F_j,n_{k_j}})$ and $\mathbf{Q} \sim N_{n_\ell,n_p}(\boldsymbol{\mu}_Q, \mathbf{\Sigma}_{Q,n_\ell}, \mathbf{\Sigma}_{Q,n_p})$, respectively. The matrix normal prior distributions are selected based on computational convenience in the sense that they facilitate conjugate sampling and the inherent separability accommodates high-dimensional data applications. Further, to facilitate conjugate sampling and to increase model flexibility, the inverse covariance (precision) matrices associated with $\mathbf{\Sigma}_{\mathbf{F}_j,n_\ell}$, $\mathbf{\Sigma}_{\mathbf{F}_j,n_{k_j}}$, $\mathbf{\Sigma}_{\mathbf{Q},n_\ell}$, and $\mathbf{\Sigma}_{\mathbf{Q},n_p}$ are assigned Wishart distributions, $\mathbf{\Sigma}^{-1}_{\mathbf{F}_j,n_\ell} \sim W_{n_\ell}(\mathbf{V}_{\mathbf{F}_j,n_\ell}, v_{\mathbf{F}_j,n_\ell})$, $\mathbf{\Sigma}^{-1}_{\mathbf{F}_j,n_{k_j}} \sim W_{n_{k_j}}(\mathbf{V}_{\mathbf{F}_j,n_{k_j}}, v_{\mathbf{F}_j,n_{k_j}})$, $\mathbf{\Sigma}^{-1}_{\mathbf{Q},n_\ell} \sim W_{n_\ell}(\mathbf{V}_{\mathbf{Q},n_\ell}, v_{\mathbf{Q},n_\ell})$, and finally $\mathbf{\Sigma}^{-1}_{\mathbf{Q},n_p} \sim W_{n_p}(\mathbf{V}_{\mathbf{Q},n_p}, v_{\mathbf{Q},n_p})$, respectively. The matrices $\mathbf{V}_{\mathbf{F}_j,n_\ell}$, $\mathbf{V}_{\mathbf{F}_j,n_{k_j}}$, $\mathbf{V}_{\mathbf{Q},n_\ell}$ and $\mathbf{V}_{\mathbf{Q},n_p}$ are specified scale matrices and $v_{\mathbf{F}_j,n_\ell}$, $v_{\mathbf{F}_j,n_{k_j}}$, $v_{\mathbf{Q},n_\ell}$, and $v_{\mathbf{Q},n_p}$ are specified degrees of freedom, with the specific choices depending on the particular application (e.g., see the online supplement).

In the process model (2.8), $\mathbf{B}_j$ is the matrix of truncated expansion coefficients associated with the jth functional (image) covariates. To facilitate conjugate computation and accommodate high-dimensional data, we specify matrix normal prior distributions $\mathbf{B}_j \sim N_{P,n_i}(\mathbf{0}, \mathbf{\Sigma_B}, \mathbf{\Sigma}_{n_i})$. The row components of $\mathbf{B}_j$ are associated with the column components of $\mathbf{F}_j$ and $\mathbf{Q}$. These covariance matrices can be specified as warranted by specific applications. It is fairly common in high-dimensional Bayesian applications to specify diagonal priors for these covariance matrices, e.g., $\mathbf{\Sigma_B} = \text{diag}(\tau_{\mathbf{B},1}, \ldots, \tau_{\mathbf{B},P})$, where $\tau_{\mathbf{B},p}$, $p = 1, \ldots, P$, are hyperparameters. Alternatively, if there is no a prior scientific preference of specific parameters, to facilitate computation we might consider the simpler form, $\mathbf{\Sigma_B} = \tau_{\mathbf{B}}\mathbf{I}_P$, with hyperparameter $\tau_{\mathbf{B}}$. In both cases, one typically specifies fairly large values (relative to the scale of the data) for the hyperparameters to make the priors less informative (e.g., see Section 3). We note also that one could fairly easily perform stochastic search variable selection (see George (2000) for detailed overview) for the elements of $\mathbf{B}_j$. Such variable selection methods have recently shown promise in other Bayesian functional/image covariate applications (e.g., see Holan et al. (2010); Wikle and Holan (2011); Holan et al. (2012)).

Finally, to facilitate model flexibility and conjugate computation, we specify Wishart priors for $\mathbf{\Sigma}_{n_\ell}^{-1} \sim W_{n_\ell}(\mathbf{V}_{G,n_\ell}, v_{G,n_\ell})$ and $\mathbf{\Sigma}_{n_i}^{-1} \sim W_{n_i}(\mathbf{V}_{G,n_i}, v_{G,n_i})$, where $\mathbf{V}_{G,n_\ell}$ and $\mathbf{V}_{G,n_i}$ are specified scale matrices and v_{G,n_ℓ} and v_{G,n_i} are specified values of degree of freedom. Specific choices for these hyperparameters depend on the particular application (e.g., see the online supplement).

3. Soil Science Application: Spatial Prediction of Electrical Conductivity Profiles

We present our motivating problem and data, followed by Bayesian estimation, and the results. Discussion surrounding the specific model choices used for implementation and sensitivity analysis is provided in the online supplement.

3.1. Motivating problem

Scientific understanding of soil properties and processes is important for many purposes, including crop and hydrologic modeling, the improvement of sustainable agricultural production systems, and possible remediation of atmospheric carbon dioxide through carbon sequestration. Measurement of important soil properties traditionally requires significant and laborious field work to collect samples as well as significant laboratory time and often expensive methodologies to analyze the physical, chemical and biological properties of those samples. Increasingly, proximal soil sensing technologies (Viscarra Rossel et al. (2012)) are being used to obtain high-resolution soil information for applications such as precision agriculture and digital soil mapping. A key proximal soil sensing

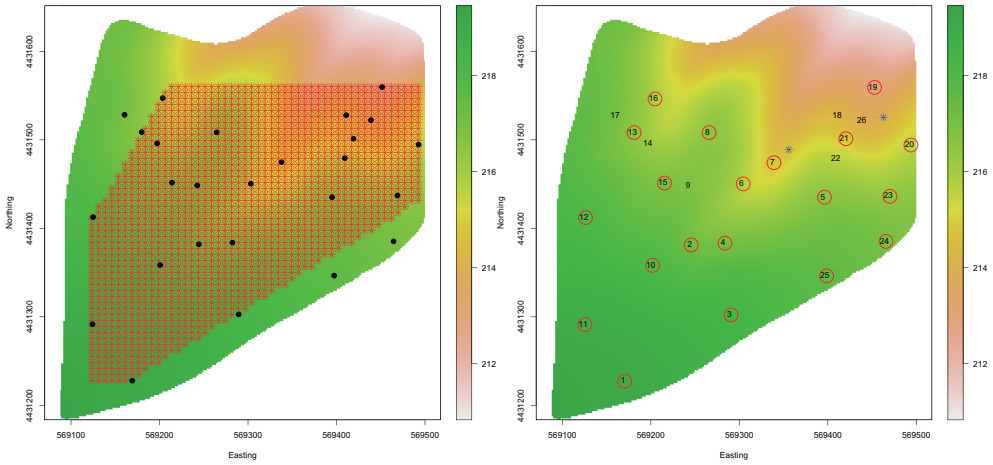

Figure 1. Left Panel: Elevation surface of the study area (measured in meters). Spatial prediction grid (stars) and observation locations (dots). Right Panel: The elevation surface (measured in meters) of the study area. Numbers correspond to the 26 observation locations considered in the analysis and stars are locations with missing image covariate information. The spatial convolution kernel knot locations are given by circles, as selected by a space-filling design.

technology is optical diffuse reflectance spectroscopy (DRS) in the visible and near-infrared (VNIR) wavelength ranges (∼400-2,500mm), which can be used to obtain a great deal of information quickly in the laboratory (Sudduth and Hummel (1996); Sudduth et al. (2010)) or in-situ with mobile sensors (Christy (2008)). In particular, VNIR-DRS can provide information about subsurface soil variation through wavelength by depth maps (or images). It is of increasing interest to be able to use such information as covariates in spatial and spatio-temporal models because it is relatively inexpensive to obtain and mitigates the need for expensive and time-consuming laboratory analyses. The purpose of the analysis presented here is to use functional/image covariates obtained from the VNIR soil spectra depth profiles to predict functional (in depth) response curves in space via the methodology described in Section 2.

3.2. Greenly research center data

We consider soil data collected at 28 sites as shown in Figure 1. In general, we are interested in the spatial prediction of profile soil electrical conductivity measured with a soil penetrometer (ECp) as a function of depth in terms of images corresponding to the VNIR soil spectra depth profiles. At each location, ECp is measured at up to 48 depth segments and the VNIR spectra are measured at wavelengths from 500 to 2,500 nanometers (nm) for each depth segment as

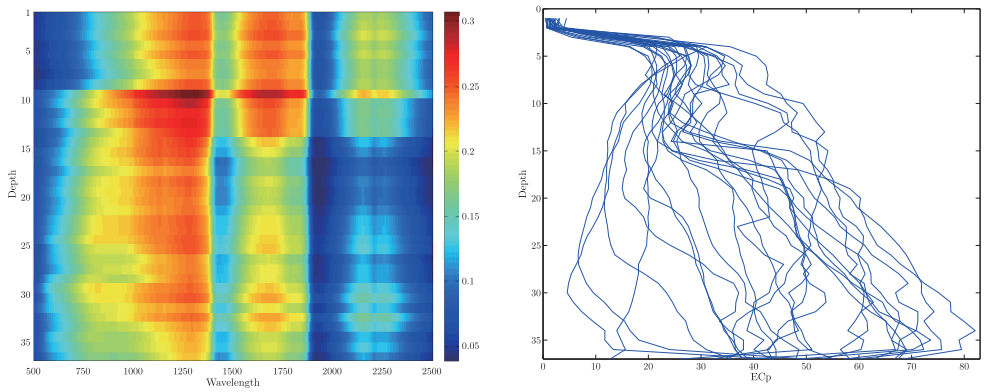

Figure 2. Left Panel: The VNIR covariate wavelength by depth image from a randomly selected data location. Wavelength is measured in nanometers, whereas each depth increment represents 2.5 cm. The color scale here represents decimal reflectance. Right Panel: ECp (measured in milliSiemens per meter) as a function of depth (where each increment represents 2.5 cm) as measured at the 26 locations shown in Figure 1.

shown in Figure 2. We note that some locations have missing information at the deeper depth segments. To retain as many locations as possible for model building, we thus only consider the first 37 depth segments. Consequently, we have 26 locations with complete covariate and response information and these are used to fit the model in our analysis. The two discarded locations are missing most of the VNIR spectra information. The elevation at each spatial location in the domain is also shown in Figure 1; this variable is used as a spatial covariate in the analysis to help predict the ECp depth responses. The ECp response data at the 26 locations considered here are shown in Figure 2.

Specifically, the soil profile measurements we consider were collected at the University of Missouri Greenly Research Center near Novelty Missouri, USA (Lat. 40.03°, Lon. -92.19°) (see Myers et al. (2011) for a comprehensive description). The exact locations sampled within the study site are shown in Figure 1. Soil cores (4.5 cm $\times$ 1.2 m) were obtained from each location for ex-situ measurements. Diffuse reflectance spectra ($500 \times 2,500$ nm, FieldSpec Pro FR, ASD Inc., Boulder, CO) were measured at 2.54 cm intervals along the core length. Soil profile electrical conductivity (ECp) was measured in-situ using a Veris Profiler 3,000 with an insulated shaft (Veris Technologies, Salina, KS, USA) and interpolated to 2.54 cm intervals. Both in-situ and ex-situ measurements were made in the late spring of 2007 (see Myers et al. (2010, 2011) for data collection details).

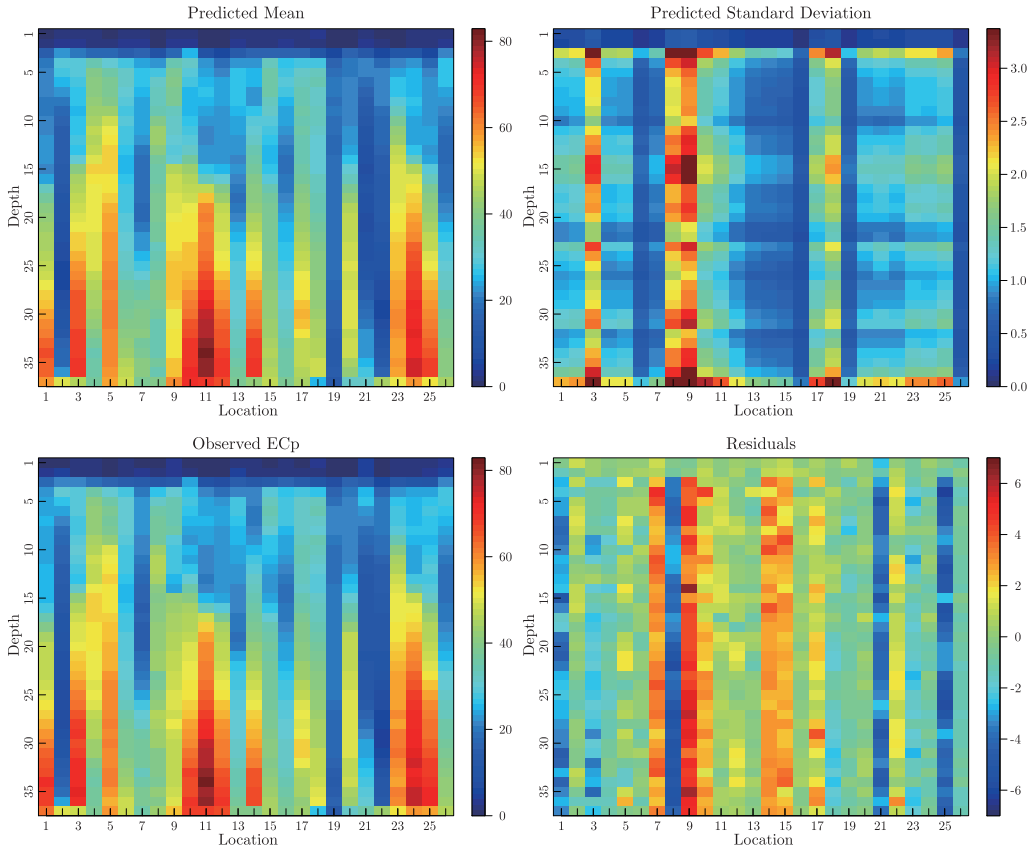

Figure 3. The leave-one-out prediction of ECp depth profile for each of the 26 observation locations. Panel (a) shows the posterior mean profile, (b) shows the posterior standard deviation, (c) shows the truth depth profile observations, and (d) shows the residuals between the posterior means and the observed depth profiles.

3.3. Results

We first evaluated the ability to predict the ECp depth profiles through a leave-one-out cross validation experiment. Figure 3 shows the out-of-sample posterior predictions of the ECp depth profiles given the VNIR image predictors. Specifically, Panel (a) shows the posterior predicted mean ECp depth profile for each of the 26 observation locations, Panel (b) shows the associated posterior standard deviation, Panel (c) gives the observed ECp depth profiles, and Panel (d) shows the associated residuals (posterior mean minus the observed). In general, from visual inspection, the functional spatial model is able to do a reasonable job in predicting the ECp depth profiles across the various locations, and the residual variation is relatively small compared to the magnitude of the

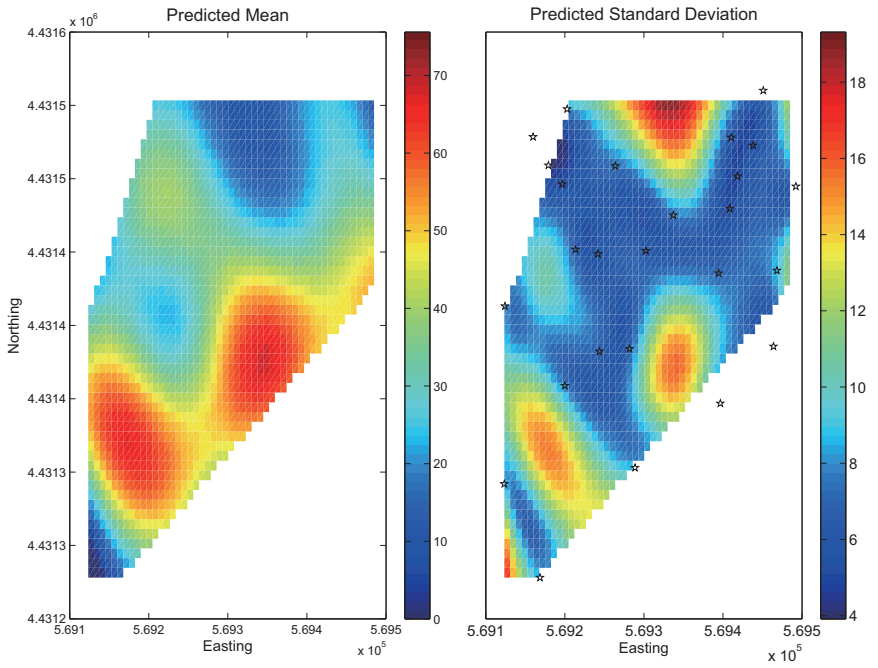

Figure 4. Predicted ECp at depth segment 16 on the prediction grid shown in Figure 1. (a) Posterior mean (negative values are truncated at 0), (b) Posterior standard deviation (note: stars correspond to observation locations).

responses. However, there are a few depth regions in some of the residual profiles that show coherent model error/bias across depth. In some cases (e.g., locations 8 and 9), the predicted standard deviation suggests this uncertainty. In other locations (e.g., 14 and 15) the model predictive standard deviation does not indicate the issues shown in the residuals. This could be due to the extreme variability in the ECp profiles (see Figure 2) at lower depths, and the possibility that the image covariates are not always helpful in predicting at those depths. These results also suggest that there should be spatially coherent prediction errors in our spatial field predictions.

We are primarily interested in predicting spatial fields of the response variable at various depths. Figure 1 shows a prediction grid over the spatial domain of interest, along with the observation locations. As an illustration, the posterior predicted mean and standard deviation on this grid for ECp at depth level 16 are shown in Figure 4. The spatial field is relatively smooth, as expected, based on the low-rank discrete kernel convolution basis representation. Spatial variation in ECp is driven mainly by spatial variation in clay content and soil density. Examination of predicted ECp at depth level 16 matches the known

trends at this study location (Myers et al. (2010, 2011)). The key trends affecting ECp are due to systematic variation in the depth to clay layers and dense glacial deposits. Each of these soil features has a relatively large ECp response. Further, the depth to these layers is related to spatial variation in elevation and geomorphology. Elevation decreases generally from southwest to northeast, with a ridge surrounding the area on the east and west edges. As expected, the northeast region of the plot indicates material with small ECp (< 30 mS/m) at depth level 16. This region corresponds to areas of coarse sediment deposition. These deposits have low electrical conductivity and bury the more conductive materials. The eastern and western edges of the study area are connected ridgelines that contain the study area. These areas have a relatively shallow depth to large clay concentrations. This is reflected in the predicted ECp values in the 40 to 50 mS/m range.

4. Discussion and Conclusion

Scientists are increasingly faced with very large data signals from new technologies and they are interested in relating these "big data" signals to various types of responses. We consider here a particular case where we have functional responses and so-called "image" predictors. In this case, we use the term image loosely, and consider it to be any two-dimensional continuous process, such as a time-frequency representation of a time signal or a depth-wavelength representation of a spectroscopic profile. We take into account here the additional complication that both the functional response and the image covariates can be spatially-dependent. This brings together several areas of research in functional data analysis, spatial statistics, and non-stationary time series analysis.

We develop a flexible, yet fairly easy to implement, methodology for the aforementioned problem by considering several layers of basis expansions. In practice, these expansions are truncated, leading effectively to a complex Bayesian mixed-effects multivariate multiple regression model. The major complication with implementation is the potentially large number of covariates (even in the basis expansion context), which is not an issue in our motivating application, but may be mitigated by the use of stochastic search variable selection priors on the parameters.

We demonstrate via an illustrative example that this methodology is useful in the context of complex soil profile data. In particular, we consider a ECp depth profile response variable as predicted by depth/wavelength images of VNIR measurements. The leave-one-out predictions of these profiles were quite promising, and spatial field prediction at a specific depth gave results that matched quite nicely with scientific insight.

Future extensions of this work include the consideration of additional covariates that are spatially dependent, including those that lack depth information. A real strength of this methodology will come from helping to suggest to practitioners where they should take additional (image) observations in order to most improve the predictions of the spatial distribution of the functional response. We will consider such optimal adaptive spatial sampling design problems as an extension to the work presented here.

Acknowledgements

We would like to acknowledge the University of Missouri Center for Agroforestry for providing access to their Novelty, MO research watershed location for the measurement activities described in this study. We would also like to thank the two anonymous reviewers and the editors for their helpful comments on an early draft. Funding for the methodological research was partially provided through National Science Foundation grant DMS-1049093 and by the U.S. National Science Foundation (NSF) and the U.S. Census Bureau under NSF grant SES-1132031, funded through the NSF-Census Research Network (NCRN) program.

References

Baladandayuthapani, V., Mallick, B. K., Hong, M. Y., Lupton, J. R., Turner, N. D. and Carroll, R. J. (2008). Bayesian hierarchical spatially correlated functional data analysis with application to colon carcinogenesis. *Biometrics* **64**, 64-73.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Springer.

Christy, C. D. (2008). Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy, *Computers and Electronics in Agriculture* **61**, 10-19.

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. John Wiley and Sons.

Dawid, A. P. (1981). Some matrix-variate distribution theory: notational considerations and a Bayesian application. *Biometrika* **68**, 265-274.

Delicado, P., Giraldo, R., Comas, C. and Mateu, J. (2010). Statistics for spatial functional data: some recent contributions. *Environmetrics* **21**, 224-239.

George, E. I. (2000). The variable selection problem. *J. Amer. Statist. Assoc.* **95**, 1304-1308.

Giraldo, R., Delicado, P. and Mateu, J. (2010). Continuous time-varying kriging for spatial prediction of functional data: An environmental application. *J. Agricultural, Biological, and Environmental Statist.* **15**, 66-82.

Giraldo, R., Delicado, P. and Mateu, J. (2012). Hierarchical clustering of spatially correlated functional data. *Statist. Neerlandica* **66**, 403-421.

Goulard, M. and Voltz, M. (1993). Geostatistical interpolation of curves: a case study in soil science. In *Geostatistics Tróia '92* (Edited by A. Soares), 805-816. Springer.

Gromenko, O. and Kokoszka, P. (2013). Nonparametric inference in small data sets of spatially indexed curves with application to ionospheric trend determination. *Comput. Statist. Data Anal.* **59**, 82-94.

Gromenko, O., Kokoszka, P., Zhu, L. and Sojka, J. (2012). Estimation and testing for spatially indexed curves with application to ionospheric and magnetic field trends. *Ann. Appl. Statist.* **6**, 669-696.

Holan, S. H., Yang, W.-H., Matteson, D. S. and Wikle, C. K. (2012). An approach for identifying and predicting economic recessions in real-time using time-frequency functional models. *Appl. Stochastic Models in Business and Industry* **28**, 485-499.

Holan, S. H., Wikle, C. K., Sullivan-Beckers, L. E. and Cocroft, R. B. (2010). Modeling complex phenotypes: generalized linear models using spectrogram predictors of animal communication signals. *Biometrics* **66**, 914-924.

Kokoszka, P. (2012). Dependent functional data. *ISRN Probability and Statistics*, 2012.

Martinez, J. G., Bohn, K. M., Carroll, R. J. and Morris, J. S. (2013). A study of Mexican free-tailed bat chirp syllables: Bayesian functional mixed models for nonstationary acoustic time series. *J. Amer. Statist. Assoc.* **108**, 514-526.

Monestiez, P. and Nerini, D. (2008). A cokriging method for spatial functional data with applications in oceanology. In *Functional and Operatorial Statistics* (Edited by S. Dabo-Niang and F. Ferraty), 237-242. Springer.

Morris, J. S., Baladandayuthapani, V., Herrick, R. C., Sanna, P. and Gutstein, H. B. (2011). Automated analysis of quantitative image data using isomorphic functional mixed models, with application to proteomics data. *Ann. Appl. Statist.* **5**, 894-923.

Myers, D. B., Kitchen, N. R., Sudduth, K. A., Grunwald, S., Miles, R. J., Sadler, E. J. and Udawatta, R. P. (2010). Combining proximal and penetrating soil electrical conductivity sensors for high-resolution digital soil mapping. In *Proximal Soil Sensing* (Edited by R. A. Viscarra Rossel, A. B. McBratney and B. Minasny), 233-243. Springer.

Myers, D. B., Kitchen, N. R., Sudduth, K. A., Miles, R. J., Sadler, E. J. and Grunwald, S. (2011). Peak functions for modeling high resolution soil profile data. *Geoderma* **166**, 74-83.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. 2nd edition. Springer-Verlag.

Reiss, P. T. and Ogden, R. T. (2010). Functional generalized linear models with images as predictors. *Biometrics* **66**, 61-69.

Ruiz-Medina, M. D. (2011). Spatial autoregressive and moving average Hilbertian processes. *J. Multivariate Anal.* **102**, 292-305.

Ruiz-Medina, M. D. (2012a). New challenges in spatial and spatiotemporal functional statistics for high-dimensional data. *Spatial Statist.* **1**, 82-91.

Ruiz-Medina, M. D. (2012b). Spatial functional prediction from spatial autoregressive Hilbertian processes. *Environmetrics* **23**, 119-128.

Ruiz-Medina, M. D. and Espejo, R. M. (2013). Integration of spatial functional interaction in the extrapolation of ocean surface temperature anomalies due to global warming. *Internat. J. Appl. Earth Observation and Geoinformation* **22**, 27-39.

Sudduth, K. A. and Hummel, J. W. (1996). Geographic operating range evaluation of a NIR soil sensor. *Trans. ASAE* **39**, 1599-1604.

Sudduth, K. A., Kitchen, N. R., Sadler, E. J., Drummond, S. T. and Myers, D. B. (2010). VNIR spectroscopy estimates of within-field variability in soil properties. In *Proximal Soil Sensing* (Edited by R. A. Viscarra Rossel, A. B. McBratney and B. Minasny), 233-243. Springer.

Viscarra Rossel, R. A., Adamchuk, V. I., Sudduth, K. A., McKenzie, N. J. and Lobsey, C. (2012). Proximal soil sensing: An effective approach for soil measurements in space and time. *Adv. Agronomy* **113**, 237-282.

Wikle, C. K. and Holan, S. H. (2011). Polynomial nonlinear spatio-temporal integro-difference equation models. *J. Time Series Anal.* **32**, 339-350.

Wikle, C. K. (2010). Low rank representations as models for spatial processes. In *Handbook of Spatial Statistics* (Edited by A. Gelfand and P. Diggle and M. Fuentes and P. Guttorp), 107-118. Chapman and Hall/CRC.

Yang, W.-H., Wikle, C. K., Holan, S. H. and Wildhaber, M. L. (2013). Ecological prediction with nonlinear multivariate time-frequency functional data models. *J. Agricultural, Biological, and Environmental Statist.* **18**, 450-474.

Zhou, L., Huang, J. Z., Martinez, J. G., Maity, A., Baladandayuthapani, V. and Carroll, R. J. (2010). Reduced rank mixed effects models for spatially correlated hierarchical functional data. *J. Amer. Statist. Assoc.* **105**, 390-400.

CSIRO Computational Informatics, QLD 4102, Australia.

E-mail: Wen-Hsi.Yang@csiro.au

Department of Statistics, University of Missouri, Columbia, MO 65211, USA.

E-mail: wiklec@missouri.edu

Department of Statistics, University of Missouri, Columbia, MO 65211, USA.

E-mail: holans@missouri.edu

Department of Plant Sciences, University of Missouri, Columbia, MO 65211, USA.

E-mail: MyersDB@missouri.edu

USDA-ARS-Cropping Systems and Water Quality Unit, Columbia, MO 65211, USA.

E-mail: SudduthK@missouri.edu

(Received August 2013; accepted February 2014)