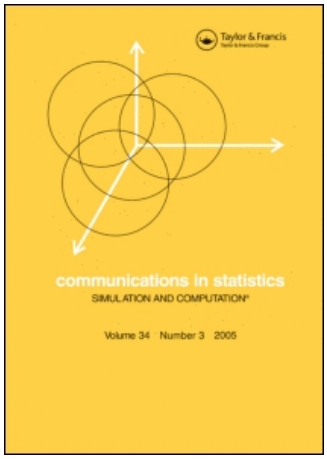

Communications in Statistics - Simulation and Computation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597237

Hierarchical Spatial Modeling and Prediction of Multiple Soil Nutrients and Carbon Concentrations

Anandamayee Majumdar [a]; Jason Kaye [b]; Corinna Gries [c]; Diane Hope [c]; Nancy Grimm [c]
[a] Department of Mathematics and Statistics, Arizona State University, Tempe, Arizona, USA
[b] Department of Crop and Soil Sciences, Pennsylvania State University, University Park, Pennsylvania, USA
[c] Global Institute of Sustainability, Arizona State University, Tempe, Arizona, USA

Online Publication Date: 01 February 2008
To cite this Article: Majumdar, Anandamayee, Kaye, Jason, Gries, Corinna, Hope, Diane and Grimm, Nancy (2008) 'Hierarchical Spatial Modeling and Prediction of Multiple Soil Nutrients and Carbon Concentrations', Communications in Statistics - Simulation and Computation, 37:2, 434 - 453
To link to this article: DOI: 10.1080/03610910701792588
URL: http://dx.doi.org/10.1080/03610910701792588

Communications in Statistics—Simulation and Computation®, 37: 434–453, 2008

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610910701792588

Statistics in Environment

Hierarchical Spatial Modeling and Prediction of Multiple Soil Nutrients and Carbon Concentrations

ANANDAMAYEE MAJUMDAR[1], JASON KAYE[2], CORINNA GRIES[3], DIANE HOPE[3], AND NANCY GRIMM[3]

[1]Department of Mathematics and Statistics, Arizona State University, Tempe, Arizona, USA
[2]Department of Crop and Soil Sciences, Pennsylvania State University, University Park, Pennsylvania, USA
[3]Global Institute of Sustainability, Arizona State University, Tempe, Arizona, USA

Modeling the multivariate spatial distribution of soil carbon and nutrients has been a challenge for ecosystem ecologists. There is a need for explanatory models, which give insight into socio-economic and biophysical controls on soil spatial variability. We propose a hierarchical Bayesian modeling specification, an approach that takes into account the spatial covariates as well as the inter-dependent nature of soil nutrients and carbon pools. We develop the model to explain variability in soil nutrient and carbon pools in the Central Arizona Phoenix Metropolitan region where soil-composition has changed considerably over the years due to socio-economic factors. A fully Bayesian statistical analysis of how these changes have affected soil nutrients provides insight as to how socio-economics influence changes in ecology. Our model included five geomorphic, ecological, and socio-economic independent variables that were used to predict soil total N, organic C, inorganic C, and extractable PO_4^{3-}*. Using six levels of hierarchy, we fit a suitable spatial hierarhical model. Using a Bayesian co-kriging strategy, we generate appropriate values used for predictions at new locations where covariate information is unavailable. We compare prediction results from standard models and show that our model is richer and so is the interpretation. To the best of our knowledge, this is the first work that applies hierarchical Bayesian modeling techniques and kriging strategies to study multivarate soil nutrient and carbon concentrations. We conclude a discussion of our findings and the broader ecological applicability of our modeling style.*

Received October 2006; Accepted June 2007
Address correspondence to Anandamayee Majumdar, Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287-1804, USA; E-mail: ananda@math.asu.edu

Keywords Bayesian framework; Coregionalization; Gibbs sampling; Hierarchical modeling; Markov chain Monte Carlo; Multivariate spatial processes.

Mathematics Subject Classification 62F15; 91B72.

1. Introduction

In this article, we handle the problem of modeling and analyzing the spatial distribution of soil variables that are correlated with each other, spatially autocorrelated due to unmeasured factors, using a variety of independent variables that reflect both traditional ecological and less conventional socioeconomic drivers, and then interpolate data from sparse measurement points to the rest of the landscape where independent variables have not been measured. The main purpose of this article is to describe a new approach to the problem of modeling soil chemical distributions using hierarchical Bayesian modeling approach.

Two of the most basic metrics of ecosystem structure are the quantities of carbon (C) and nutrients stored in soil. Organic carbon is of interest because it constrains the quantity of energy available for soil food webs and because soils store a large fraction of global surface C (Amundson, 2001). Inorganic C pools reflect long-term ecosystem water and salt balances. Nutrient pool sizes affect plant productivity and soils are a sink for anthropogenic nutrient pollution (Aber et al., 1998). Given this universal importance, it is not surprising that ecologists have spent decades trying to understand variability in soil carbon and nutrient storage both within a given site and across landscapes and regions (Burke et al., 1989; Kaye et al., 2002, 2005; Parton et al., 1987; Robertson et al., 1997).

In previous studies, a common approach to modeling variability has been to assemble soil carbon and nutrient data from many sites and use traditional multiple regression techniques to determine which independent variables are significant predictors of nutrient and carbon pools (Burke et al., 1989; Kaye et al., 2002). While this approach provides estimates of variability in regional soil carbon storage, it has several weaknesses. First, large scale studies oversimplify the region by restricting analyses to a few landscape components (e.g., just natural or agricultural portions) and these components are generally described by separate models. By avoiding urban ecosystems, these models also simplify the potential drivers of soil chemical variability by excluding socioeconomic measures from the list of independent variables. Second, large scale studies are generally based on very sparse soil sampling. Many rely on soil surveys (e.g., USDA NRCS data) data that characterize the chemistry of one soil pit to map soil properties over hundreds of square kilometers. Sparse data make it difficult to scale from the soil pit to the region because simple linear interpolation between points is unrealistic and more sophisticated scaling assumes that independent variables are well known between sample points. Sparse data also are problematic because many unknown drivers of soil chemistry are not included in the model and can not (because sample locations are often not explicitly known) be incorporated as spatial autocorrelation. At smaller scales, spatial autocorrelation has been taken into account, explaining large portions of within site variability (Ettema and Wardle, 2002; Robertson et al., 1997).

A final shortcoming of traditional regression approaches is that multiple soil processes are not modeled simultaneously, even when we know they are correlated.

For example, in trying to understand how C and N vary across a landscape separate models are constructed for each element, despite the fact that C:N ratios are known to be quite predictable. It would be advantageous, and more realistic, to include C and N (and other elements) in the same model, so that variability in one soil pool could be used to inform the model prediction of other soil pools of interest.

Here we describe a hierarchical Bayesian solution to the problem of modeling the spatial distribution of multiple soil chemistry across a landscape of heterogeneous landuse where prediction at newer spatial locations are restricted by unavailability of some of the covariates. We now discuss the modeling issues in more detail.

2. Modeling Background and Issues

Various modeling lineages have been developed and applied to explain soil-nutrient and carbon concentrations (Bossatta and Agren, 1991; Burke et al., 1989; Kaye et al., 2002; Parton et al., 1987; Robertson et al., 1997 and many others) with little attempt at synthesis. These include deterministic and stochastic models, phenomenological and mechanistic models, spatial and non spatial models.

The development of satisfactory stochastic models does raise a number of methodological issues. Here, we offer an overview of a critical subset of these issues.

First, there is a matter of modeling style. Do we employ a phenemenological model that attempts to relate different data layers, each measuring different variables or mechanistic models that attempt to simulate ecosystem properties via a fundamental understanding of processes? If the process was temporal, then should it be extended to a dynamic version or is a static view of the processes adequate? There is no simple right or wrong answer to these choices. In our case, a static-phenomenological model was the clear choice because (1) we did not have a sound mechanistic understanding of controls on variability in soils across the region; and (2) our dataset included only a one-time snapshot of soil properties without information on temporal dynamics. Here we must note that no proposed explanatory model is correct, but some are more useful than the others. Also, no proposed explanatory model establishes causality, merely relationship. In our case, we hypothesized that typical geomorphic and ecological variables might be important correlates of soil properties in desert ecosystems, but that socio-economic variables would be more important determinants of soil properties in urban and agricultural ecosystems.

When data layers are to be inter-related, they can be introduced at various layers of the hierarchical specification, which becomes an explanatory model for the process (Wu and David, 2002). In this specification, the lowest is one response variable. This modeling strategy is referred to as hierarchical or multilevel modeling. Here, for example, the hierarchical levels for the full model include the four soil-nutrient concentrations, percent of area covered with lawn, and percent of impervious area in the plot. This approach allows for considerable flexibilty in modeling and the possibility of incorporating mechanistic components in linking certain levels. Hierarchical modeling is achieving increased utilization in analyzing data collected from complex processes (see e.g., Clark, 2005; Gelfand et al., 2003, 2004; Wikle, 2003).

In the context of explaining soil nutrient characteristics, the data layers are inherently spatial. Data that are obviously spatial should be modeled in an

explicitly spatial fashion. Only rarely has this been done in modeling soil-nutrient concentrations, and where spatial models have been developed, the analysis does not interconnect the data layers. While there may be issues of model choice in providing spatial structure, it seems clear that measurements which are closer to each other in space should be more strongly associated than those farther from each other.

Another challenge of using a phenomenological model is that extrapolation or interpolation from measured data points requires complete data layers for explanatory variables outside the measurement points. For at least some data layers there will be missing data. Usually, ecologists and soil scientists assign a value based on a "best guess" or mean from known samples (e.g., all soils within the same soil series have the same soil C value). Models of spatial association for gridded data assume there are no missing observations. To accommodate such *holes* in the dataset, kriging (i.e., statistical prediction of missing values) provides an alternative to subjective assignments of values for explanatory variable that are not measured at points where soil characteristics are being predicted by the model. That is to say, one kriges observations at the missing locations according to the likelihood of the different possible values. Such kriging must be done randomly. Moreover, in the presence of missing values, multiple kriging are necessary in order to capture the effect of the uncertainty associated with kriging. Christensen and Ameniya (2002) have dealt with latent variable models in the multivariate context, where as our approach can be seen as a way to handle variables that are observed at some locations, but unobserved in others. Our approach may be viewed as a theoretically equivalent version to Wackernagel's (1998) model of *corregionalization*, but here the parametrization is different, and hence the computational aspect to the problem is easier to tackle. We observe that hierarchical spatial approach for multivariate data has also been applied in a non Gaussian setting in Liu et al. (2005). Commonly used constructive approaches also include a moving average (kernel convolution approach), as suggested in Ver Hoef and Barry (1998).

In our analysis, two important explanatory variables were not available outside of our sample collection points so we designed an Bayesian kriging strategy that accounts for both stochasticity and spatial structure in the independent variables.

3. The Dataset

The research was conducted within and around the Phoenix Metropolitan area of 3.5 million people (U.S. Census Bureau, 2000). Natural vegetation in the region is Sonoran desert, but Native American irrigation agriculture was prevalent up to the 1400's and Anglo American agriculture and urbanization have occupied large portions of the region since the 1950s (Fig. 1). Most of the agricultural land is flood irrigated and urban land includes both xeric (desert-like) and mesic landscaping with various modes of irrigation. Annual daily (1948–2003) maximum and minimum temperatures are 30° and 15°C, respectively, and annual average rainfall is 193 mm.

The study region was a roughly rectangular area of 6,400 km^2 that included the city and surrounding agricultural lands and desert (Fig. 1). We use probability-based sampling to acquire a spatially dispersed, unbiased group of sample points from the region (Peterson et al., 1999; Stevens, 1997). A randomized, tessellation-stratified design was achieved by superimposing a 4 km × 4 km grid on the study area, giving 462 potential sampling units. We expected that landscape heterogeneity would be greater in the urban core (Luck and Wu, 2002), so a random sample point was

Figure 1. Land-use in Phoenix (Stefanov et al., 2001).

assigned within every square inside the urban core and in every third square outside that area, giving a total sample size of 206. No a priori stratification according land cover, land use type, or other characteristics was used.

At each of the sample points, a 30 m × 30 m plot (the size of a Landsat Thematic Mapper pixel) centered on the randomly assigned sites located with the aid of a Global Positioning System (Trimble Pro XRS real-time satellite corrected mapping grade unit) regardless of land cover or ownership was selected. Access was granted to all but 8 sites, 6 of these were relocated to the nearest (within 100 m) similar accessible site, but for 2 sites access was denied and no suitable surrogate could be found, giving a total of 204 sample sites. Soil cores were taken using a hand-impacting corer (2.5 cm diameter to a depth of 30 cm) at four points in each plot. Core samples were separated into 0–10 cm (top) and 10–30 cm (bottom) depth intervals. Soils from the four cores for each depth were composited and refrigerated. At a small number of survey plots where the entire surface was covered by impervious urban surface, soil samples were collected from the nearest accessible site within 100 m of the plot boundary. The composite samples were sieved (2 mm) at field moisture, air dried, and homogenized by hand. To determine available P pool sizes, a 10 g sample was extracted with $NaHCO_3$ and ortho-phosphate concentrations were determined colorimetrically (Clesceri et al., 1998). For other analyses, a subsample was ground to a fine powder and analyzed for total N and C by dry combustion elemental analysis and for inorganic C by pressure calcimeter (Sherrod et al., 2002). An additional soil core (5 cm diameter by 15 cm deep) was

collected at the center of each plot. The main land use categories used in our analysis were: (1) urban residential with xeric vegetation (N = 22); (2) urban residential with mesic vegetation (N = 23); (3) urban residential with a mixture of xeric and mesic vegetation (N = 8); (4) urban non residential (includes industrial, commercial, transportation, parks, and golf courses; N = 41); (5) water (N = 4); (6) desert (N = 73); (7) agriculture (N = 23); and (8) a mixture of multiple land-use types (N = 11).

4. Model Details

We model the joint distribution of the natural logarithms of Total Nitrogen (Y_1), Organic Carbon (Y_2), Inorganic Carbon (Y_3), and Total Phosphorous (Y_4) concentrations (gm/m^2). Let Y_{1j}, Y_{2j}, Y_{3j}, and Y_{4j} denote the corresponding concentrations in the jth spatial location. We also observe at each plot percent of lawn cover, L_j, slope (meters), S_j, elevation (meters), E_j, percent of impervious area (P_j), and a categorical variable δ_j (1–0) measuring whether or not the plot was ever used in agriculture. In order to achieve model parsimony these 5 explanatory variables were chosen from a suite of 13 explanatory variables by observing that each of the coefficients for the remaining eight variables yielded 95% credible intervals that contained zero. Lastly, let $W_{\cdot j}$ denote the spatial random effects in the model $Y_{\cdot j} = T(\cdot j) + W_{\cdot j} + \epsilon_{\cdot j}$, where $T(\cdot j)$ is the part that corresponds to the spatial regression component, and $\epsilon_{\cdot j}$ is random noise. The notations of the variables used in the model are given in Table 1.

In specifying the distribution, our concern is not only to explain the four kinds of soil-nutrient concentrations with the the remaining five explanatory variables, but also to predict at new locations (for the purpose of kriging). In this regard, it is important to note that among the five explanatory variables, percent of impervious area and percent of lawn cover are not available at new locations. The remaining three explanatory variables are available for non sampled locations. (Land-use categories are significant covariates for the percent of impervious area and percent of lawn cover and so a sixth variable, namely the land-use is incorporated in the model. Note that this covariate information is available for any location in the area

Table 1
Notations corresponding to the variables used in the model

Y_1	log Total Nitrogen
Y_2	log Organic Carbon
Y_3	log Inorganic Carbon
Y_4	log Total Phosphorous
S	slope (meters)
E	elevation (meters)
δ	= 0 if never in agriculture, = 1 if ever used in agriculture
P	percent of impervious area
L	percent of lawn cover
LU	Land use category

of our interest.) Hence, to address this issue, we treat percent of impervious area and percent of lawn cover as stochastic processes.

As a result, there are now six stochastic processes in the model and the joint distribution form becomes $f(\mathbf{L}, \mathbf{P}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3, \mathbf{Y}_4 \mid \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU})$. We factor this joint distribution as

$$f(\mathbf{P} \mid \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}) f(\mathbf{L} \mid \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}) f(\mathbf{Y}_1 \mid \mathbf{L}, \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}) \ldots$$
$$f(\mathbf{Y}_4 \mathbf{Y}_3, \mathbf{Y}_2, \mathbf{Y}_1, \mathbf{L}, \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}). \tag{1}$$

Let us consider the first distribution in (1). This is the model for percent of impervious area. We assume that the P_j are conditionally jointly Gaussian given the S_j, E_j, δ_j, and LU_j (we take the log transformation to symmetrize the very skewed data). And thus, there is also a spatial random effect in this Gaussian model, W_{Pj}, so that the P_j are conditionally independent and identically distributed given $\mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}$, and $\mathbf{W}_P$. A similar interpretation goes for the conditional distribution of the percent of lawn-cover given P_j, S_j, E_j, δ_j, and LU_j.

We have:

$$P_j \mid \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{W}_P \overset{iid}{\sim} N\left(\beta_{P0} + \beta_{P|S} S_j \cdots + \beta_{P|LU} LU_j + W_{Pj}, \sigma_P^2\right)$$
$$L_j \mid \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{W}_L \overset{iid}{\sim} N\left(\beta_{L0} + \beta_{L|P} P_j + \cdots + \beta_{L|LU} LU_j + W_{Lj}, \sigma_L^2\right); \tag{2}$$
$$j = 1, \ldots, N.$$

We now consider the third distribution in (1). This is the Total Nitrogen model. We assume that the Y_{1j} are conditionally jointly Gaussian given the P_j, L_j, S_j, E_j, δ_j, and LU_j (we take the log transformation to symmetrize the very skewed data). And thus there is also a spatial random effect in this Gaussian model, W_{1j}, so that the Y_{1j} are conditionally independent and identically distributed given $\mathbf{P}, \mathbf{L}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}$, and $\mathbf{W}_1$. We thus have

$$Y_{1j} \mid \mathbf{P}, \mathbf{L}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{W}_P \overset{iid}{\sim} N\left(\beta_{10} + \beta_{1|P} P_j + \beta_{1|L} L_j + \cdots + W_{1j}, \beta_{1|LU} LU_j \sigma_1^2\right);$$
$$j = 1, \ldots, N.$$

In the case of the second distribution in (1), the log transformation of the organic carbon, we again assume conditionally independent Gaussian distribution, given a spatial random effect $\mathbf{W}_2$ and have

$$Y_2 \mid \mathbf{P}, \mathbf{L}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}, \mathbf{W}_2 \overset{iid}{\sim} N\left(\beta_{20} + \beta_{2|P} P_j + \beta_{2|L} L_j + \beta_{2|S} S_j + \beta_{2|E} E_j + \beta_{2|\delta} \delta_j \right.$$
$$\left. + \beta_{2|LU} LU_j + W_{2j}, \sigma_2^2\right).$$

For the third and fourth distributions in (1), we have similar assumptions. In general, the conditional distribution of Y_{ij}, given $Y_{i-1j}, \ldots, Y_{1j}$, $j = 1, \ldots, N$, the explanatory variables and a spatial component W_{ij} are independent Gaussian over i and j. Specifically, we have:

$$Y_{1j} \mid \mathbf{L}, \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}, \mathbf{W}_1$$
$$\overset{iid}{\sim} N\left(\beta_{10} + \beta_{1|P} P_j + \beta_{1|S} S_j + \beta_{1|E} E_j + \beta_{1|\delta} \delta_j + \beta_{1|LU} LU_j + W_{1j}, \sigma_1^2\right)$$

$$Y_{ij} \mid \mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1}, \mathbf{L}, \mathbf{P}, \mathbf{S}, \mathbf{E}, \boldsymbol{\delta}, \mathbf{LU}, \mathbf{W}_1, \ldots, \mathbf{W}_i \tag{3}$$
$$\overset{iid}{\sim} N\big(\beta_{i0} + \beta_{i|1} Y_{1j} + \cdots + \beta_{i|i-1} Y_{i-1j} + \beta_{i|P} P_j \beta_{i|S} S_j + \beta_{i|E} E_j + \beta_{i|\delta} \delta_j + \beta_{i|LU} LU_j + W_{ij}, \sigma_i^2\big), \quad i = 2, 3, 4.$$

In all of the distribution specifications (2) and (3), the $\beta_{\cdot|j}$'s are regressian coefficients while the $W_{\cdot j}$'s are spatial random effects, and are anticipated to capture the spatial dependence of the corresponding variables. Unlike usual random effects which are assumed to be independent and identically distributed normal variables, the distributions of the W's are specified by a joint covariance matrix where

$$\operatorname{cov}(W_{\cdot j}, W_{\cdot j'}) = \tau_{\cdot}^2 \exp(-\phi_{\cdot} \Delta_{jj'}).$$

Here, $\Delta_{jj'}$ is the distance between the jth and the j'th plot. The τ's capture the spatial variance of the spatial random effects and the ϕ's capture the rate of decay of the spatial random effects. The spatial random effects across each of the variables are assumed as independent.

The order of the hierarchical steps assumed in this model in (2) and in (3) have nothing special about the hierarchy. We could have as well started with organic carbon at the upper level, nitrogen at the lower level, and so on. Since in each of these cases the joint distribution is fully specified, we note that no matter what the permutations among the hierarchical steps are, as long as the responses are the four soil variables, the models are essentially equivalent and hence similar in performance.

The hierarchical conditional model as in (2) and (3) was introduced by Royle and Berliner (1999) in a spatial context. Schmidt and Gelfand (2003) showed that this model is equivalent to the corregionalization model (Wackernagel, 1998) and performs better than models that are weaker in assumption. "Weaker in assumption" might mean, for example, a model such as (1) which does not have some or all of the hierarchical steps. Including all the response variables in the same model via the hierarchy can improve model performance as opposed to non inclusion of some of the response variables (Schmidt and Gelfand, 2003).

5. Fitting and Inference Details

The model defined in (1) as described by (2) and (3) is referred to as a multilevel or hierarchical specification. At the lowest level, we have percent impervious area, then the percent lawn-cover, and after that the total nitrogen and so on, until finally the total phosphorous. This model is of high dimension with unknowns, $\{\beta_{P|\cdot}\}$, $\{W_{P\cdot}\}$, $\{\sigma_P\}$, $\{\phi_P\}$, $\{\tau_P\}$, $\{\beta_{L|\cdot}\}$, $\{W_{L\cdot}\}$, $\{\sigma_L\}$, $\{\phi_L\}$, $\{\tau_L\}$, $\{\beta_{i|\cdot}\}$, $\{W_{i\cdot}\}$, $\{\sigma_i\}$, $\{\phi_i\}$, $\{\tau_i\}$, $i = 1, 2, 3, 4$. Fitting of this model is feasible only in the Bayesian framework. That is, we have already provided the joint distribution for $\mathbf{P}, \mathbf{L}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3, \mathbf{Y}_4$ given the model parameters. If we view these latter variables as random and add a so-called prior distribution for them, we have a complete model specification. This means we have specified the joint distribution of all the variables in the hierarchical model. With this joint distribution, we can provide inference regarding any aspect of the model. The prior distribution of the $W_{\cdot j}$ was described above. For $\{\beta_{\cdot|\cdot}\}, \{\sigma_{\cdot}\}, \{\phi_{\cdot}\}$, and $\{\tau_{\cdot}\}$, the general approach is to use whatever prior or partially data-based information we may have regarding these unknowns to obtain

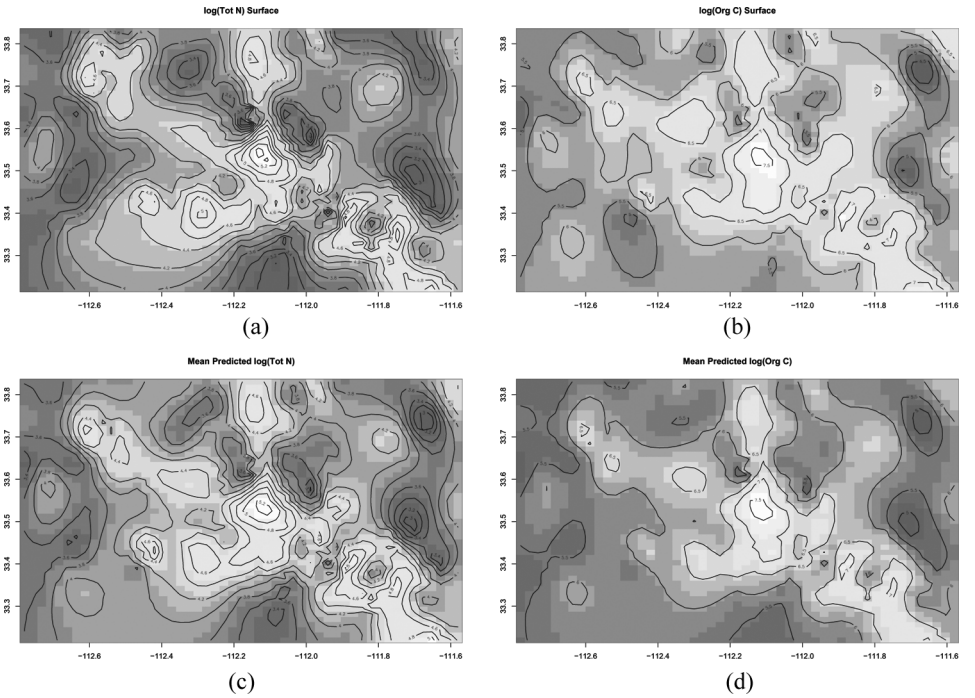

Figure 2. A: Tot N surface; B: Org C surface; C: Mean predicted Tot N surface under Model I; D: Mean predicted Org C surface under Model I.

some idea of what range they are likely to fall in. We use this range to develop proper prior distributions, which are non informative (i.e., as vague) as possible (in order to let the observed data drive the inference) while still retaining stable computation. Details of the prior specification and the posterior distribution are supplied as Appendix. Comparison between true and predicted surfaces of the soil nutrients and carbon using the model is presented in Figs. 2 and 3.

The resulting model is fitted using simulation-based methods, i.e., Gibbs sampling (Gelfand and Smith, 1990) and Markov chain Monte-Carlo methods. The output from such simulation is sampled from the joint posterior distribution of all model unknowns. Unfortunately, such model fitting requires considerable effort and time. The reward is exact inference (without relying on possibly inappropriate asymptotics) and more accurate measurement of variability by capturing the uncertainty regarding the model unknowns, which is not obtained using classical statistical approaches. In fact, we know of no other way to fit model (2) and (3). Simplified versions may be fitted using the E–M algorithm (Dempster et al., 1977) to obtain likelihood-based inference. But then we have concerns regarding associated asymptotic variances.

6. Bayesian Prediction for Missing Values of Covariates

In order to make inferences about locations outside the data, some of the explanatory spatial variables become unobtainable. These variables are percent of impervious area and missing percent of lawn cover. Rather than relying on

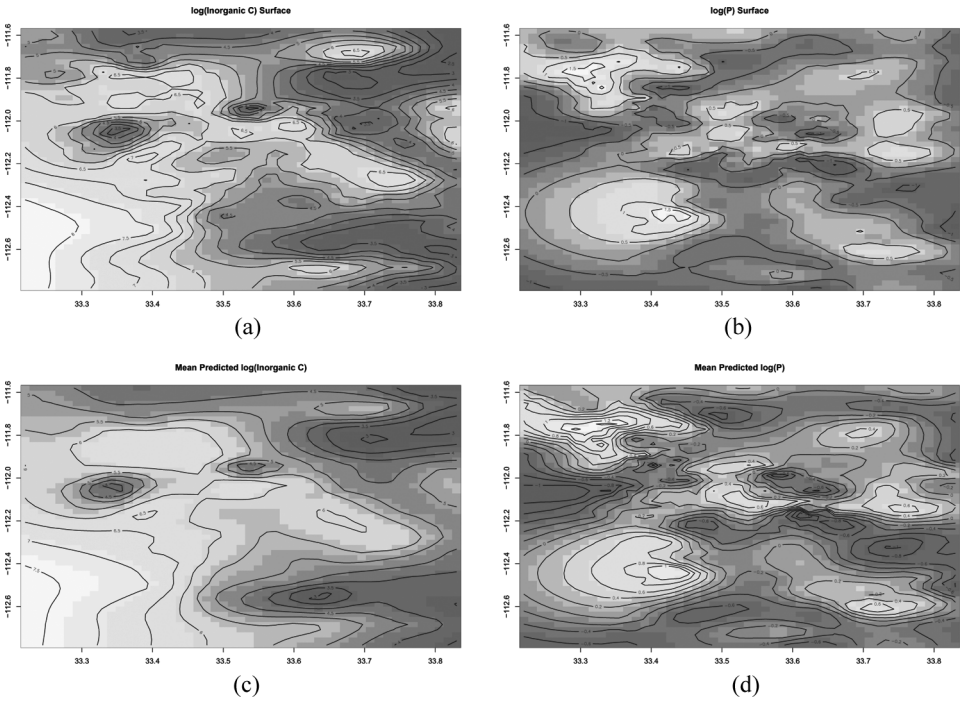

Figure 3. A: Inorg C surface; B: P surface; C: Mean predicted Inorg C surface under Model I; D: Mean predicted P surface under Model I.

subjective or ad hoc determination of these missing explanatory variables, we choose to krige missing percent of impervious area and missing percent of lawn cover in order to have a complete set for prediction using our hierarchical model. In fact, the kriging provides the entire set of missing percent of impervious area and missing percent of lawn cover using joint spatial distribution for the plots. Such a distribution should support (but not require) percent of impervious area and missing percent of lawn cover for a given plot to be similar to that of its neighbors.

7. Model Determination

The model described using (2) and (3), has four levels of hierarchy, and is therefore elaborate and challenging to fit. Below we discuss the importance of this effort by comparing our full model (called "Model I" hereafter) to two simplifications that were easier to fit and predict that result in Models II and III. Using the prediction criterion, we show that they are not as good.

A crude model, which we will refer to as Model II, would use some crude estimate for percent impervious area and percent lawn cover. For example, to predict at a certain plot known to belong to a certain land-use type, the estimate might be the sample averages obtained from that particular land-use type. Another way to fit the same model stated in (2) and (3) would be to omit the spatial random effects of each of the specified distributions the model and carry out a similar exercise of prediction. Let us refer to this non spatial hierarchical specification model as Model III.

To obtain predictions at a new location indexed by j^*, using the four-level hierarchical model in (2) and (3), first the prediction of P_{j^*} and L_{j^*} would have to be made given the posterior distribution of the unknowns. These estimates would then be substituted in the appropriate equations to get the posterior predictive distribution of Y_{ij^*}, $i = 1, 2, 3, 4$. The outline of the posterior predictive inference is given in Appendix, part B.

To compare the foregoing models we find 95% posterior predictive intervals for 15 spatial points. These 15 points were left out from the data set to obtain a cross-validation as well as model comparison. These 15 points were chosen so that they covered all the 8 different land use types.

8. Model Results

The first nutrient in our hierarchical model was total soil nitrogen, which was explained using the percent of pervious area that was occupied by lawn and whether the site had an agricultural history (Table 2). Of the variability not explained by these two independent variables (median $\tau_1^2 + \sigma_1^2 = 0.407$), more than one third could be attributed to spatial variation (median $\tau_1^2 = 0.142$). The correlation of the median predicted values had a 0.94 correlation with the actual data. Nonsignificance of any dependent variable was determined by the 95% posterior credible of the corresponding regression coefficient in the respective hierarchy of the model.

The best fit model for organic carbon included as explanatory variables, total nitrogen and percent impervious surface area, and in this case half of the unexplained variance in the data (median $\tau_2^2 + \sigma_2^2 = 0.230$) was attributable to spatial variation (median $\tau_2^2 = 0.111$). The correlation of the median predicted values and the actual data was 0.94.

Only elevation was significant in predicting inorganic carbon values but posterior median prediction still accounted for 49% of the variability in the data

Table 2
Posterior sample summaries of the parameters in Model I

Parameter	2.5th%	Median	97.5th%
	log(Total Nitrogen)		
Regression parameters			
Intercept	3.686	3.878	4.073
Ever in agriculture(0–1)	0.091	0.326	0.559
% of lawn in pervious area	0.004	0.007	0.010
Spatial variance τ_1^2	0.085	0.142	0.238
Error variance σ_1^2	0.204	0.265	0.343
	log(Organic Carbon)		
Regression parameters			
Intercept	1.715	2.143	2.568
log(Tot N)	0.847	0.945	1.043
Percent impervious $\times 10^{-3}$	0.067	2.57	5.12
Spatial variance τ_2^2	0.070	0.111	0.175
Error variance σ_2^2	0.087	0.119	0.161

Table 3
Posterior sample summaries of the parameters in Model I

Parameter	2.5th%	Median	97.5th%
	log(Inorganic Carbon)		
Regression parameters			
Intercept	7.490	8.450	9.371
Elevation (meters)	−0.009	−0.007	−0.005
Spatial variance τ_3^2	0.251	0.510	0.954
Error variance σ_3^2	0.470	0.703	0.969
	log(Phosphorous)		
Regression parameters			
Intercept	−4.354	−3.744	−3.164
log (Total N)	0.771	0.908	1.047
Spatial variance τ_4^2	0.115	0.192	0.316
Error variance σ_4^2	0.185	0.254	0.338

(Table 3). The correlation of the median predicted values and the actual data was 0.87. Of the variability not explained by elevation (median $\tau_3^2 + \sigma_3^2 = 1.213$), more than one third could be attributed to spatial variation (median $\tau_3^2 = 0.510$). Extractable phosphorous was predicted using total nitrogen, with about half of

Table 4
Posterior sample summaries of the parameters in Model I

log(% of imp. area)	2.5%	Med.	97.5%
	Regression parameters		
Intercept	11.758	13.131	15.10
Landuse 4 (Urban Mixed)	−0.972	−0.646	−0.330
Landuse 5 (Open;Water)	−2.734	−1.948	−1.153
Landuse 6 (Open; Desert)	−2.659	−2.314	−1.967
Landuse 7 (Agricultural Land)	−2.844	−2.423	−2.002
Landuse 8 (Mixed)	−1.501	−0.987	−0.467
Spatial variance τ_P^2	0.315	0.432	0.575
Error variance σ_P^2	0.146	0.265	0.462
	log(% area lawn cover)		
Regression parameters			
Intercept	0.970	1.159	1.351
Landuse 2 (Mesic)	2.641	2.983	3.323
Landuse 3 (Residential Mixed)	1.773	2.295	2.814
Landuse 4 (Urban Mixed)	−6.430	−3.871	−1.305
Landuse 4 (Urban Mixed)*elevation	0.003	0.009	0.015
Landuse 4 (Urban Mixed)*ever in ag.	0.976	1.667	2.358
Spatial variance τ_L^2	0.333	0.419	0.529
Error variance σ_L^2	0.076	0.127	0.221

the unexplained variance (median $\tau_4^2 + \sigma_4^2 = 0.446$) coming from spatial variation (median $\tau_4^2 = 0.245$). The correlation between the median predicted values and actual data was 0.89.

For Bayesian kriging purposes, log(impervious area) was predicted by the last five land-use types (Table 4), all of which are categorical variables. Correlation between median predicted values and the data was 0.95. Spatial variation was about 1.5–2 times as the residual error in the model. The percent of pervious area covered by lawn was best predicted by three land use types (categorical variables) (Table 4), two of which were non significant in the percent impervious model, and two interaction terms: mixed urban and ever in agriculture, mixed urban and elevation. The correlation between actual data and median predicted values was 0.92. Spatial variation was about 3–4 times the residual variance in the model. For both variables (percent impervious area and percent of pervious area covered by lawn) we took the log transformation to symmetrize the very much skewed data.

In Tables 5, 6, and 7, prediction inferences at 15 hold-out points are made using the three models described in the previous section. These prediction inferences include 95% posterior predictive intervals given by the posterior predictive 2.5 and 97.5th percentiles and a point level prediction given by the posterior predictive median or 50th percentile. The first inference in Table 5 corresponds to Model I, the second inference in Table 6 to Model II, and the 3rd inference in Table 7 to Model III. The intervals which fail to include the real observation are marked in italics. Many of the predictive intervals corresponding to Model III (the non spatial model) fail to include the real observation. Model II which uses point estimates for % impervious area and % lawn cover fails to capture the real observation for two

Table 5
Model I: Posterior predictive intervals of the 15 points that were held out of the models during model development

Obs.	(2.5th%, Med., 97.5th%)	Obs.	(2.5th%, Med., 97.5th%)
	log(Total N)		log(Org C)
4.97	(3.13 4.30 5.55)	6.93	(6.29 7.15 8.05)
3.38	(3.14 4.32 5.52)	5.20	(4.66 5.53 6.39)
5.47	(3.41 4.58 5.86)	7.87	(6.58 7.43 8.35)
3.85	(3.11 4.25 5.42)	5.71	(5.19 6.03 6.83)
4.56	(2.58 3.85 5.08)	6.72	(5.56 6.46 7.41)
3.98	(2.57 3.84 5.07)	6.22	(5.11 6.04 6.95)
3.84	(2.46 3.76 5.00)	6.14	(4.91 5.89 6.76)
4.67	(3.11 4.42 5.67)	6.77	(5.74 6.67 7.58)
3.75	(2.98 4.18 5.36)	5.81	(5.04 5.91 6.76)
5.52	(3.78 4.95 6.17)	7.43	(6.81 7.68 8.55)
3.88	(3.33 4.51 5.70)	6.38	(5.15 6.02 6.90)
4.4	(3.28 4.42 5.58)	6.71	(5.60 6.42 7.26)
4.9	(3.13 4.32 5.57)	6.87	(5.96 6.84 7.74)
4.1	(2.59 3.84 5.07)	6.32	(5.22 6.11 7.06)
2.69	(2.43 3.70 4.90)	3.89	(3.83 4.79 5.67)

Table 6
Model II: Posterior predictive intervals of the 15 points that were held out of the models during model development

Obs.	(2.5th%, Med., 97.5th%)	Obs.	(2.5th%, Med., 97.5th%)
	log(Total N)		log(Org C)
4.97	(3.11 4.29 5.54)	6.93	(4.87 6.35 7.86)
3.38	(3.09 4.26 5.47)	5.20	(4.81 6.31 7.75)
5.47	(3.59 4.79 6.05)	*7.87*	*(4.72 6.24 7.79)*
3.85	(3.12 4.26 5.43)	5.71	(4.75 6.25 7.75)
4.56	(2.58 3.85 5.08)	6.72	(4.53 6.06 7.62)
3.98	(2.57 3.84 5.07)	6.22	(4.66 6.17 7.72)
3.84	(2.46 3.76 5.00)	6.14	(4.69 6.17 7.75)
4.67	(3.06 4.37 5.61)	6.77	(4.61 6.17 7.74)
3.75	(2.99 4.19 5.37)	5.81	(4.74 6.25 7.78)
5.52	(3.65 4.81 6.04)	7.43	(4.82 6.27 7.83)
3.88	(3.26 4.45 5.64)	6.38	(4.76 6.29 7.84)
4.4	(3.43 4.57 5.73)	6.71	(4.89 6.25 7.75)
4.9	(3.13 4.32 5.57)	6.87	(4.66 6.12 7.62)
4.1	(2.59 3.84 5.07)	6.32	(4.49 6.16 7.63)
2.69	(2.58 3.86 5.05)	*3.89*	(*4.78 6.25 7.88*)

Table 7
Model III: Posterior predictive intervals of the 15 points that were held out of the models during model development

Obs.	(2.5th%, Med., 97.5th%)	Obs.	(2.5th%, Med., 97.5th%)
	log(Total N)		log(Org C)
4.97	(*3.39 4.30 4.71*)	*6.93*	(*5.51 6.17 6.69*)
3.38	(*3.76 4.32 5.01*)	*5.20*	(*5.89 6.18 6.90*)
5.47	(*4.09 4.34 5.32*)	*7.87*	(*6.08 6.19 7.14*)
3.85	(*4.10 4.35 5.35*)	*5.71*	(*6.09 6.19 7.23*)
4.56	(3.40 4.12 4.76)	*6.72*	(*5.55 6.19 6.70*)
3.98	(3.19 4.10 4.46)	6.22	(5.68 6.31 6.80)
3.84	(3.58 4.12 4.83)	6.14	(5.94 6.25 7.03)
4.67	(4.06 4.34 5.28)	6.77	(6.08 6.19 7.16)
3.75	(*4.11 4.35 5.34*)	*5.81*	(*6.09 6.19 7.23*)
5.52	(*3.64 4.31 4.89*)	*7.43*	(*5.58 6.17 6.68*)
3.88	(3.43 4.30 4.71)	*6.38*	(*5.58 6.17 6.69*)
4.4	(3.75 4.32 5.02)	6.71	(5.84 6.18 6.91)
4.9	(4.06 4.34 5.36)	6.87	(6.08 6.19 7.18)
4.1	(3.90 4.15 5.15)	6.32	(6.17 6.27 7.27)
2.69	(3.44 4.12 4.65)	*3.89*	(*5.56 6.17 6.63*)

observations while Model I (the full model mentioned via (2) and (3)) is successful in capturing all the real observations within the 95% posterior predictive intervals.

9. Conclusion and Discussion

To summarize what we have accomplished, we have formulated a phenomenological hierarchical model at the pixel level, which explains soil nutrient concentrations given elevation, percent of lawn cover, and the categorical variable that states whether the plot was ever used for agriculture or not (1–0). We have also been able to explain the nature of interdependence among these different soil nutrient concentrations. We have used a stochastic, spatial Bayesian kriging strategy to handle prediction at new locations where the percent of impervious area and percent of lawn cover are unavailable by predicting these unavailable covariates using the hierarchical model and using these predictions to make farther predictions on soil-nutrient concentrations. At this point, we have noted the importance of land-use as a driver for the afore-mentioned two independent variables. This is how socio-economic variables and land-use play an important role in spatial variation of the soil-chemical concentrations in the area of study. We have introduced spatial smoothing for each of the four response variables used in the hierarchical model and have seen predictive mean surfaces that are close to the real data surface. For all models, posterior predicted values are highly correlated with actual data. We see that total nitrogen, organic carbon, and phosphorous concentrations increase with increase in the urbanization covariates. We have argued and demonstrated that our hierarchical model captures uncertainty, which was underestimated by simpler models. We have also shown the sort of statistical inference that is possible with our hierarchical model, particularly with regard to the spatial smoothing.

As noted earlier, ours is just one example of the many types of models that could be constructed to explain the variability in this dataset. For example, another covariate used for explaining the soil-nutrient concentrations might be the depth of the soil. If there are data sets for multiple depths then depth could be used as covariate information.

Our present application has led to a model for land-use, which is static, since we lack temporal information. We could extend the model described in (2) and (3) if the data were temporal. Formally, we need only to add a subscript t to those measurements which change over time. Spatio-temporal random effects $W_{\cdot,j,t}$ could be introduced to capture association across both space and time. So far, the model described by (2) and (3) studies the assocation among different explanatory variables with the responses, and the association among the responses themselves. The model also gives a framework for prediction at new locations. In addition to (and making use of) the fully Bayesian spatial model formulated here, one could introduce directional gradients in the soil nutrient concetration surfaces that study directional changes and find the second order behavior (as opposed to the first order behavior given by the mean) of these surfaces. Such gradient behaviors have been introduced in the past by Gelfand et al. (2003) used in the context of land-prices by Majumdar et al. (2006). Specifically in this context, such directional gradients could study the effect of say, urbanization or desertification or any other possible socio-economic change at the infinitesimal level. For ecologists, hierarchical Bayesian models represent a new tool for quantifying patterns and processes in nature. The major strengths, as illustrated here and outlined by Clark (2005) are that: (1) the

approach enables models that reflect the real-world connectedness of dependent and independent variables; (2) they allow multiple regression to be combined with spatial dependence; and (3) they enable stochastic predictions for points outside the data set. It is difficult to imagine an ecological modeling problem that would not benefit from at least one of these capabilities.

Appendix

A. Prior Details

To complete the specification of the hierarchical model we require priors for the $\boldsymbol{\beta}_{i|\cdot}$'s, the $\sigma_\cdot^2$'s, the $\phi_\cdot$'s and the $\tau_\cdot^2$'s. In particular, we assume independent flat normal priors for the $\boldsymbol{\beta}_{i|\cdot}$, Inverse Gamma$(\alpha_{\sigma(\cdot)}, \beta_{\sigma(\cdot)})$ for $\sigma_\cdot^2$, Inverse Gamma$(\alpha_{\tau(\cdot)}, \beta_{\tau(\cdot)})$ for the error variance $\tau_\cdot^2$ and Gamma$(\alpha_{\phi(\cdot)}, \beta_{\phi(\cdot)})$ for the decay parameter $\phi_\cdot$. In particular, $\sigma_\cdot^2 \sim IG(\alpha_{\sigma(\cdot)}, 2)$ and $\tau_\cdot^2 \sim IG(\alpha_{\tau(\cdot)}, 2)$. This specification has infinite variance. We choose $\alpha_{\sigma(\cdot)}$, $\alpha_{\tau(\cdot)}$ so that the means of the $\sigma_\cdot^2$ and $\tau_\cdot^2$ are data-centered so as to make the convergence of the chain faster.

B. Posterior Distribution

Here we provide a detailed derivation of the posterior distributions that were used for the developments in the details of Model (2) and (3). We initially focus on the parameters in the model and then describe how posterior simulations obtained for the parameters can be used for prediction purposes.

Let $\mathbf{Y}$ denote the vector of arranged in sub-vectors corresponding to total N, organic C, inorganic C, and extractable PO_4^{3-} i.e.,

$$\mathbf{Y}^T = (\mathbf{Y}_1^T, \ldots, \mathbf{Y}_4^T)$$

with

$$\mathbf{Y}_i = (Y_{i,1}, \ldots, Y_{i,N})^T$$

the log transformation of the ith soil-concentrate for N spatial points. The vector of parameters for data from model (2) and (3)

$$\boldsymbol{\theta}^T = \big(\boldsymbol{\beta}_{P|\cdot}^T, \boldsymbol{\beta}_{L|\cdot}^T, \boldsymbol{\beta}_{1|\cdot}^T, \ldots, \boldsymbol{\beta}_{4|\cdot}^T, \sigma_P^2, \sigma_L^2, \sigma_1^2, \ldots, \sigma_4^2, \phi_P, \phi_L, \phi_1, \ldots, \phi_4, \tau_P^2, \tau_L^2, \times \tau_1^2, \ldots, \tau_4^2\big),$$

then has the associated likelihood function

$$\begin{aligned} L(\mathbf{Y}, \mathbf{L}, \mathbf{P} \mid \mathbf{W}, \boldsymbol{\theta}) \propto\ & \tau_P^{-N}\sigma_P^{-N}\exp\left\{-(2\sigma_P^2)^{-1}\sum_{j=1}^{N}\left(P_j - W_{Pj} - X_{Pj}^T\boldsymbol{\beta}_{P|\cdot}\right)^2\right\} \\ & \times \tau_L^{-N}\sigma_L^{-N}\exp\left\{-(2\sigma_L^2)^{-1}\sum_{j=1}^{N}\left(L_j - W_{Lj} - X_{Lj}^T\boldsymbol{\beta}_{L|\cdot}\right)^2\right\} \\ & \times \prod_{i=1}^{4}\tau_i^{-N}\sigma_i^{-N}\exp\left\{-(2\sigma_i^2)^{-1}\sum_{j=1}^{N}\left(Y_{ij} - W_{ij} - X_{ij}^T\boldsymbol{\beta}_{i|\cdot}\right)^2\right\} \end{aligned}$$

with (Y_{ij}, W_{ij}), $i = 1, \ldots, 4$ being the log soil concentrations and the spatial components associated with log soil-chemical concentrations for the jth spatial point and $X_{\cdot j}$ is the jth row corresponding to the matrix X_i which entails the regressor variables corresponding to the conditional distribution of $\mathbf{Y}_i$ given $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1})$. The developments in the "Model details" section now entail that the joint distribution of $\mathbf{Y}$, $\mathbf{P}$, $\mathbf{L}$, $\mathbf{W}$, and $\boldsymbol{\theta}$ is the product of $L(\mathbf{Y}, \mathbf{P}, \mathbf{L} \mid \mathbf{W}, \boldsymbol{\theta})$ with the prior density

$$\pi(\boldsymbol{\theta}, \mathbf{W}) = \pi(\mathbf{W}_P \mid \boldsymbol{\theta})\pi(\boldsymbol{\beta}_{P|\cdot})\pi(\phi_P)\pi\big(\sigma_P^2\big)\pi\big(\tau_P^2\big)\pi(\mathbf{W}_L \mid \boldsymbol{\theta})\pi(\boldsymbol{\beta}_{L|\cdot})\pi(\phi_L)\pi\big(\sigma_L^2\big)\pi\big(\tau_L^2\big) \times \prod_{i=1}^{4} \pi(\mathbf{W}_i \mid \boldsymbol{\theta})\pi(\boldsymbol{\beta}_{i|\cdot})\pi(\phi_i)\pi\big(\sigma_i^2\big)\pi\big(\tau_i^2\big).$$

The next step is to derive the posterior densities for each of $\boldsymbol{\theta}$ the components of that will allow us to carry out Bayesian inference via the multi-stage Gibbs sampler. For this development we will employ the symbol "\" to indicate removal of a particular parameter or parameters from $\boldsymbol{\theta}$.

We begin the marginalization process by noting that the posterior density for σ_i^2 is

$$\pi\big(\sigma_i^2 \mid \mathbf{Y}, \mathbf{L}, \mathbf{P}, \mathbf{W}, \boldsymbol{\theta}\backslash\sigma_i^2\big) \propto \sigma_i^{-N} \exp\left(-\sum_{j=1}^{N} \big(Y_{ij} - W_{ij} - X_{ij}^T\boldsymbol{\beta}_{i|\cdot}\big)^2/2\sigma_i^2\right) \times \sigma_i^{-2(\alpha_\sigma+1)} \exp\big(-\beta_\sigma/\sigma_i^2\big).$$

From this it follows that $\sigma_i^2 \mid \mathbf{Y}, \mathbf{L}, \mathbf{P}, \mathbf{W}, \boldsymbol{\theta}\backslash\sigma_i^2$ has an inverse gamma distribution with parameters $2^{-1}N + \alpha_\sigma$ and $2^{-1}\sum_{j=1}^{N}(Y_{ij} - W_{ij} - X_{ij}{}^T\boldsymbol{\beta}_{i|\cdot})^2 + \beta_\sigma)$. Hence, we can show that $\sigma_\cdot^2 \mid \mathbf{Y}, \mathbf{L}, \mathbf{P}, \mathbf{W}, \boldsymbol{\theta}\backslash\sigma_\cdot^2$ has an inverse gamma distribution with parameters $2^{-1}N + \alpha_\sigma$ and $2^{-1}\sum_{j=1}^{N}(\cdot_{\cdot j} - W_{\cdot j} - X_{\cdot j}^T\boldsymbol{\beta}_{\cdot|\cdot})^2 + \beta_\sigma)$.

Similarly, we can show that the posterior distribution $\tau_\cdot^2 \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}\backslash\tau_\cdot^2$ has an inverse gamma distribution with parameters $2^{-1}N + \alpha_\tau$ and $2^{-1}\mathbf{W}_\cdot^T C(\mathbf{W}_\cdot, \phi_\cdot)$ $\mathbf{W}_\cdot + \beta_\tau$. Here $C(\mathbf{W}_\cdot, \phi_\cdot)$ is a $N \times N$ variance–covariance matrix corresponding to prior of the spatial component $\mathbf{W}_\cdot$.

For the joint posterior distribution of the vectors $\boldsymbol{\beta}_{i|\cdot}$, $i = 1, \ldots, 4$, we obtain:

$$\pi(\boldsymbol{\beta}_{i|\cdot} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}\backslash\boldsymbol{\beta}_{i|\cdot}) \propto \exp\left(-\big(2\sigma_i^2\big)^{-1}\sum_{j=1}^{N}\big(Y_{ij} - W_{ij} - X_{ij}^T\boldsymbol{\beta}_{i|\cdot}\big)^2\right).$$

This leads to $\boldsymbol{\beta}_{i|\cdot} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}\backslash\boldsymbol{\beta}_{i|\cdot}$ having a multivariate normal distribution with posterior mean

$$\tilde{\boldsymbol{\mu}}_i = \widetilde{\Sigma}_i\left(\sigma_i^{-2}\sum_{j=1}^{N} X_{ij}(Y_{ij} - W_{ij})\right)$$

and posterior variance–covariance matrix

$$\widetilde{\Sigma}_i = \sigma_i^2\big(X_i^T X_i\big)^{-1}.$$

In a similar fashion, we can show that $\boldsymbol{\beta}_{\cdot|\cdot} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta} \backslash \boldsymbol{\beta}_{\cdot|\cdot}$ has a multivariate normal distribution with posterior mean

$$\tilde{\boldsymbol{\mu}}_{\cdot} = \tilde{\Sigma}_P \left(\sigma_{\cdot}^{-2} \sum_{j=1}^{N} X_{\cdot j} (\cdot_j - W_{\cdot j}) \right)$$

and posterior variance–covariance matrix

$$\tilde{\Sigma}_{\cdot} = \sigma_{\cdot}^2 (X_{\cdot}{}^T X_{\cdot})^{-1}.$$

The posterior distribution of $\phi_{\cdot}$ can be derived as follows. First, observe that:

$$\pi(\phi_{\cdot} | \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta} \backslash \phi_{\cdot}) \propto \exp\left(-(2\tau_{\cdot}^2)^{-1} \mathbf{W}_{\cdot}^T C(\mathbf{W}_{\cdot}, \phi_{\cdot}) \mathbf{W}_{\cdot}\right) \quad \exp -(\beta_{\phi} \phi_{\cdot}^{\alpha_{\phi}-1}).$$

The above expression has no closed form, and we use the Metropolis-Hastings algorithm (Hastings, 1970) to sample from this distribution within a Gibbs sampler.

In order to find the posterior predictive distribution of $Y_{ij^*} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}$, we note that $[Y_{ij^*} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}] = \left[X_{ij^*}^T \boldsymbol{\beta}_{i|\cdot} + W_{ij^*} + \epsilon_{ij^*}\right]$, where $\epsilon_{ij^*} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta} \sim N\left(0, \tau_i^2\right)$. Thus, from the S subsamples $\boldsymbol{\theta}^s$, $s = 1, \ldots, S$, generated in the MCMC step we can further generate subsamples of the posterior predictive distribution of $Y_{ij^*} | \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}, \boldsymbol{\theta}$. That is, we can compute $X_{ij^*}^{sT} \boldsymbol{\beta}_{i|\cdot}^s + W_{ij^*}^s + \epsilon_{ij^*}^s$, $s = 1, \ldots, S$ for any particular choice of j^* and i. Here $\epsilon_{ij^*}^s$ is simulated from a zero mean, normal distribution with variance equal to the τ_i^2 component of $\boldsymbol{\theta}^s$ and $X_{ij^*}^{sT}$ can be derived from the sth sample of the predictive distribution of $\left[L_{j^*}, P_{j^*} \mid \mathbf{Y}, \mathbf{P}, \mathbf{L}, \mathbf{W}\right]$. We can then find the median and the 95% posterior predictive intervals by evaluating the corresponding 2.5th and 97.5th percentiles of the the posterior predictive distribution.

Acknowledgments

We thank Steven S. Carroll for sampling design; M. Myers, A. Budet, S. Paine, M. Clary, A. Stiles, L. Stabler, and S. Holland for field and lab assistance; Salt River Project for the donation of helicopter time; Cities of Phoenix, Scottsdale, and Tempe, Maricopa County Parks, Tonto National Forest, Arizona State Lands Department, Sky Harbor Airport, and all the private property owners involved for giving us permission to access their land. This work was funded by National Science Foundation Grants # DEB-9714833 and DEB-0423704 (the Central Arizona-Phoenix Long-Term Ecological Research Project) and the NSF Biocomplexity in the Environment Program (EAR-0322065).

References

Aber, J., McDowell, W., Nadelhoffer, K., Magill, A., Berntson, G., Kamakea, M., McNulty, S., Currie, W., Rustad, L., Fernandez, I. (1998). Nitrogen saturation in temperate ecosystems. *Bioscience* 48:921–934.

Amundson, R. (2001). The carbon budgets in soils. *Annual Rev. Earth Planetory Sci.* 29: 535–562.

Banerjee, S., Carlin, B. P., Gelfand, A. E. (2003). *Hierarchical Modeling and Analysis for Spatial Data*. London: Chapman and Hall.

Bossatta, E., Agren, G. I. (1991). Dynamics of carbon and nitrogen in the organic matter of the soil: a generic theory. *Amer. Naturalist* 138:227–245.

Burke, I. C., Yonker, C. M., Parton, W. J., Cole, C. V., Flach, K., Schimel, D. S. (1989). Texture, climate, and cultivation effects on soil organic matter content in U.S. grassland soils. *Soil Sci. Soc. Amer. J.* 53:800–805.

Carlin, B. P., Louis, T. A. (2000). *Bayes and Emperical Bayes Methods for Data Analysis*. 2nd ed. Boca Raton, FL: Chapman and Hall/CRC.

Christensen, W. F., Ameniya, Y. (2002). Latent variable analysis of multivariate spatial data. *J. Amer. Statist. Assoc.* 97:457,302–317(16).

Clark, J. (2005). Why environmental scientists are becoming Bayesians. *Ecol. Lett.* 8:2–14.

Clesceri, L. S., Greenburg, A. E., Eaton, A. D. eds. (1998). *Standard Methods for the Examination of Water and Wastewater*. 20th ed. Prepared and Published jointly by APHA, AWWA and WEF. Baltimore, MD: United Book Press, Inc.

Dempster, A., Laird, N., Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. B* 39:1–38.

Ettema, C., Wardle, D. A. (2002). Trends in ecology and evolution. *Spatial Soil Ecol.* 17: 177–183.

Gelfand, A. E., Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* 5:398–409.

Gelfand, A. E., Kim, H., Sirman, C. F., Banerjee, S. K. (2003). Spatial modeling with spatially varying coefficient processes. *J. Amer. Statist. Assoc.* 98.

Gelfand, A. E., Schmidt, A. M., Banerjee, S., Sirmans, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test* 13:263–312.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B. (1995). Bayesian Data Analysis. Boca Raton, FL: Chapman and Hall/CRC.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57(1):97–109.

Kaye, J. P., Barrett, J. E., Burke I. C. (2002). Stable carbon and nitrogen pools in grassland soils of variable texture and carbon content. *Ecosystems* 5:461–471.

Kaye, J. P., McCulley, R., Burke, I. C. (2005). Carbon fluxes, nitrogen cycling and soil microorganisms in adjacent urban, native and agricultural ecosystems. *Global Change Biol.* 11:575–587.

Liu, X., Wall, M. M., Hodges, J. S. (2005). Generalized spatial structural equation models. *Biostatistics* 6:539–557.

Luck, M. A., Wu, J. (2002). A gradient analysis of the landscape pattern of urbanization in the Phoenix metropolitan area of USA. *Landscape Ecol.* 17:327–339.

MAG; Maricopa Association of Governments. (1997). *Urban Atlas, Phoenix Metropolitan Area*. Phoenix, AZ: Maricopa Association of Governments.

Majumdar, A., Munneke, H. J., Gelfand, A. E., Banerjee, S., Sirmans, C. S. (2006). Gradients in spatial response surfaces with application to land-values. *J. Bus. Econ. Statist.* 24:77–90.

Oleson, J. J., Hope, D., Gries, C., Kaye, J. (2005). A Bayesian approach to estimating regression coefficients for soil properties in land-use patches with varying degrees of spatial variation. *Environmetrics* 17:517–525.

Parton, W. J., Schimel, D. S., Cole, C. V., Ojima, D. S. (1987). Analysis of factors controlling soil organic matter levels in great plains grasslands. *Soil Sci. Soc. Amer. J.* 51:1173–1179.

Peterson, S. A., Urquhart, N. S., Welch, E. B. (1999). Sample representativeness: a must for reliable regional lake condition estimates. *Environ. Sci. Technol.* 33:1559–1565.

Robertson, G. P., Klingensmith, K., Klug, M., Paul, E., Crum, J. (1997). Soil resources, microbial activity and primary production across and agricultural ecosystem. *Ecolog. Applic.* 7:158–170.

Royle, M., Berliner, L. M. (1999). A hierarchical approach to multivariate spatial modeling and prediction. *J. Agricult. Biol. Environ. Statist.* 1:29–56.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.

Sclesinger, W. H., Rakes, J. A. Hartley, A. E., Cross, A. F. (1996). On the spatial pattern of soil nutrients in desert eco-systems. *Ecology* 77(2):364–374.

Schmidt, A. M., Gelfand, A. E. (2003). A Bayesian coregionalization approach for multivariate pollutant data. *J. Geo-phys. Res. Atmospheres* 108:D24:8783.

Sherrod, L., Dunn, G., Peterson, G. (2002). Inorganic carbon analysis by modified pressure-calcimeter method. *Soil Sci. Soc. Amer.* 66:299–305.

Stevens, D. L., Jr. (1997). Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics* 8:167–195.

U.S. Census Bureau. (2000). Phoenix-Mesa metropolitan statistical area population demographics. http://www.census.gov.

Ver Hoef, J. M., Barry, R. P. (1998). Constructing and fitting models for cokriging and multivariate spatial prediction. *J. Statist. Plann. Infer.* 69(2):275–294.

Wackernagel, H. (1998). *Multivariate Geostatistics: An Introduction with Applications*. 2nd ed. Berlin: Springer Verlag.

Wikle, C. (2003). Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* 84:1382–1394.

Wu, J., David, J. L. (2002). A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. *Ecol. Model.* 153:7–26.