My title

Justice Akuoko-Frimpong, Jonathan Ta   Author Two

1 Introduction

Spatial heterogeneity in environmental processes presents significant challenges for traditional regression modeling approaches. Consider air pollution monitoring across an urban-rural gradient: the relationship between particulate matter (PM2.5) concentrations and predictor variables like traffic density or industrial emissions often varies substantially across space. Global regression models that assume constant coefficients fail to capture these localized relationships, potentially leading to incorrect scientific conclusions and suboptimal policy decisions. Our work develops a computationally efficient Bayesian implementation of spatially varying coefficient (SVC) models, which address this limitation by allowing regression coefficients to change smoothly over space. Let $s_i \in \mathrm{R}^2$ for each $i = 1, \dots, n$ denote spatial locations where we observe response variables $Y(\boldsymbol{s})$ and covariates $\boldsymbol{X}(\boldsymbol{s}) = (X_1(\boldsymbol{s}), \dots, X_p(\boldsymbol{s}))^\top$. The SVC model specification:

$$Y(\boldsymbol{s}) = \sum_{j=1}^{p} X_r(\boldsymbol{s}) w_r(\boldsymbol{s}) + \epsilon(\boldsymbol{s}),$$

where $w_r(\boldsymbol{s})$ are spatially varying coefficients modeled as Gaussian processes:

$$w_r \sim \mathrm{GP}(0, C(s, s'))$$

$$C(s, s') = \sigma_r^2 \exp\left(-\phi_r^{-1} \|s - s'\|^2\right)$$

with $\epsilon(\boldsymbol{s}) \sim \mathrm{MVN}(0, \tau^2 I_n)$ representing measurement error. The key innovations in our implementation include:

Computational Algorithm

We implement a Markov Chain Monte Carlo (MCMC) algorithm with three key innovations:

1. **Low rank Gaussian Process**: Using knots to reduce the $O(n^3)$ computational complexity of Gaussian process models.

2. **Adaptive Metropolis-Hastings**: For updating spatial range parameters ϕ_r with automatic tuning of proposal distributions:

$$\phi_r^{(new)} = \phi_r^{(current)} + \epsilon, \quad \epsilon \sim N(0, \sigma^2_{adapt})$$

   where σ^2_{adapt} adjusts based on acceptance rates during burn-in.

3. **Efficient linear algebra**: Utilizing Cholesky decompositions for all covariance matrix operations, with pre-computation of constant distance matrices:

2 Methods

2.1 SVC Model

Let $s_i \in \mathrm{R}^2$ for each $i = 1, \ldots, n$ denote a spatial location for which we have collected data, and $\boldsymbol{s} = (s_1, \ldots, s_n)^\top$ be the vector of all such locations. Then $Y(\boldsymbol{s})$ are univariate dependent variables and $\boldsymbol{X}(\boldsymbol{s}) = (X_1(\boldsymbol{s}), \ldots, X_p(\boldsymbol{s}))^\top$ are $p \times 1$ vectors of covariates. A linear regression model with spatially varying coefficients assumes $Y(\boldsymbol{s})$ are dependent on

$\boldsymbol{X}(\boldsymbol{s})$ as follows:

$$Y(\boldsymbol{s}) = \sum_{j=1}^{p} X_r(\boldsymbol{s}) w_r(\boldsymbol{s}) + w_0(\boldsymbol{s}) + \epsilon(\boldsymbol{s}),$$

where $w_r(\boldsymbol{s})$ are the regression coefficients corresponding to $X_r(\boldsymbol{s})$, $w_0(\boldsymbol{s})$ are spatial random effects, and $\epsilon(\boldsymbol{s})$ are independently and identically distributed multivariate normal (MVN) measurement errors, i.e.

$$\epsilon(\boldsymbol{s}) \sim \text{MVN}(0, \tau^2 I_n).$$

For convenience, we will treat $w_0(\boldsymbol{s})$ as the regression coefficients corresponding to $X_0(\boldsymbol{s}) = \mathbf{1}_n$, a vector of length n with every element equal to 1.

For each $w_r(\boldsymbol{s})$, we assign a Gaussian process (GP) prior with squared exponential covariance, i.e.

$$w_r \sim \text{GP}(0, C(s, s')), \text{ where}$$
$$C(s, s') = \sigma_r^2 \exp\{\phi_r^{-1} \|s - s'\|^2\}.$$

We refer to $C(\cdot, \cdot)$ as the covariance function, which calculates the covariance between two locations s and s'. Squared exponential covariance is a commonly used covariance function due to useful properties such as infinite smoothness and gradual decrease in covariance with distance. The parameter σ_r^2 is the spatial variance and ϕ_r is the spatial decay, which indicates how quickly correlation decreases with the squared distance.

We assign inverse-gamma conjugate priors for the variance parameters τ^2 and each σ_r^2. In addition, each phi_r is assigned a uniform prior. In summary,

$$\tau^2 \sim \text{Inv.Gamma}(\alpha_\tau, \beta_\tau),$$
$$\sigma_r^2 \sim \text{Inv.Gamma}(\alpha_r, \beta_r), \text{ and}$$
$$\phi_r \sim \text{Uniform}(l_r, u_r).$$

2.2 MCMC Algorithm

The primary function in the svc package, svclm, samples from the joint posterior distribution of $w_r(\boldsymbol{s}), \phi_r, \sigma_r^2$, and τ^2, for $r = 0, \ldots, p$ using a Gibbs sampler. We will explain how each parameter is sampled at iteration $t+1$.

2.2.1 Spatial Decay

The parameter $\phi_r^{(t+1)}$ is updated via a random walk Metropolis algorithm. Because $\phi_r^{(t)} \in (l_r, u_r)$, each proposal is calculated as

$$\phi_r' = f^{-1}(f(\phi_r^{(t)}) + U) = l_r + \frac{u_r - l_r}{1 + \exp\left\{\log\left(\frac{u_r - l_r}{\phi_r^{(t)} - l_r} - 1\right) + U\right\}},$$

where U is sampled from a normal distribution with mean 0.

The initial standard deviation for U is specified using the "phi_proposal_sd_start" argument or is 1 by default. Then the standard deviation is updated at each iteration via a Robust Adaptive Metropolis (RAM) algorithm (Vihola 2012).

2.2.2 Spatially Varying Coefficients

Let $\widetilde{Y}(\boldsymbol{s}) = Y(\boldsymbol{s}) - \sum_{j \neq r} X_j(\boldsymbol{s}) w_j^{(t)}(\boldsymbol{s})$. Then

$$\frac{\widetilde{Y}}{X_r}(\boldsymbol{s}) | w_r^{(t)}(\boldsymbol{s}), \tau^2 \sim \text{MVN}\left(w_r^{(t)}(\boldsymbol{s}), \tau^{2(t)} \text{diag}\left(\frac{1}{X_r^2(\boldsymbol{s})}\right)\right) := \text{MVN}\left(\mu, \Sigma\right).$$

Then assuming the prior $w_r(\boldsymbol{s}) \sim \text{MVN}\left(0, \sigma_r^{2(t)} C\left(\phi_r^{(t)}\right)\right) := \text{MVN}\left(0, \Sigma_0\right)$, we sample $w_r^{(t+1)}(\boldsymbol{s})$ from $\text{MVN}(\mu_1, \Sigma_1)$, where

$$\Sigma_1 = \left(\Sigma_0^{-1} + \Sigma^{-1}\right) \text{ and } \mu_1 = \Sigma_1 \Sigma^{-1} \frac{\widetilde{Y}}{X_r}(s).$$

2.2.3 Spatial Variance

We sample $\sigma_r^{2(t+1)}$ from $\text{Inv.Gamma}\left(\alpha_r + \frac{n}{2}, \beta_r + w_r^{(t)}(\boldsymbol{s})^\top C\left(\phi_r^{(t)}\right)^{-1} w_r^{(t)}(\boldsymbol{s})\right)$.

2.2.4 Nugget

We sample $\tau^{2(t+1)}$ from Inv.Gamma$\left(\alpha_t + \frac{n}{2}, \beta_t + \frac{1}{2}\left(Y(\boldsymbol{s}) - \sum_{j=0}^{p} X_r(\boldsymbol{s}) w_r^{(t)}(\boldsymbol{s})\right)\right)$.

3 Simulation

The purposes of the simulation are to

- assess the performance of SVCLM,
- compare the differences in performance between between full and low rank for SVCLM,
- compare the SVCLM with varycoef and spBayes.

3.1 Metrics

To achieve the purposes of the simulation study, we used the following metrics to evaluate SVCLM:

- Bias,
- RMSE,
- Computational cost (seconds).

3.2 Data generating mechanisms

We first generate a grid consisting of 21 evenly spaced longitudes and 21 evenly spaced latitudes for our coefficients to vary on.

Then, we focus on generating spatial random effects w, two regression coefficients β_1, β_2, and error term ϵ.

Priors of w:

- $\sigma_w^2 = 0.1$
- $\phi_w = 2$
- $w \sim N(1_n, \sigma_w^2 \times \phi_w(s))$

Priors of β_1:

- $\sigma_{\beta_1}^2 = 0.1$
- $\phi_{\beta_1} = 2$
- $\beta_1 \sim N(1_n, \sigma_{\beta_1}^2 \times \phi_{\beta_1}(s))$

Priors of β_2:

- $\sigma_{\beta_2}^2 = 0.1$
- $\phi_{\beta_2} = 2$
- $\beta_2 \sim N(1_n, \sigma_{\beta_2}^2 \times \phi_{\beta_2}(s))$

Finally, we generate X and Y.

- $X_1 \sim N(0, 1)$
- $X_2 \sim N(5, 1)$
- $Y = \beta_1 \times X_1 + \beta_2 \times X_2 + w + \epsilon$.

3.3 Global settings and output

We will replicate the comparison for 1000 times to establish precession and reduce the variance in sampling draw.

4 Data Analysis

The real world data used in this section is a gridded satellite data with latitude, longitude, temperature, Normalized Difference Vegetation Index (NDVI); which measures vegetation health/greenness, and emissivity which represents a surface's efficiency in emitting thermal radiation. We used temperature as a univariate outcome and used emissivity and NDVI as covariates. Figure 1 below reveals the spatial distributions of the variables in the dataset.

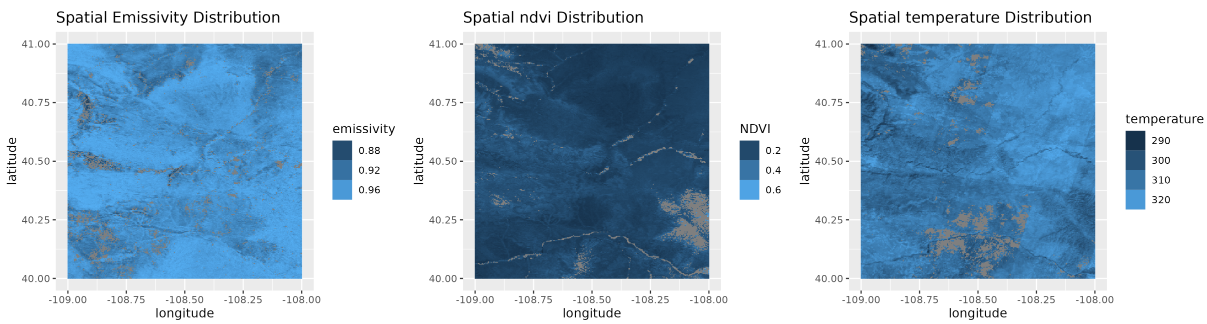

Figure 1: Spatial distributions of emissivity, NDVI, and land surface temperature in the study area.

Emissivity ranges from 0.88 to 0.96. Light blue zones ($\epsilon > 0.94$) displays dense vegetation/water bodies (high thermal emission efficiency) while dark blue clusters ($\epsilon \approx 0.88 - 0.90$) represents urbanized areas. Notice that notice that areas with high NDVI also tend to have higher emissivity which makes sense because vegetation emits efficiently. Cooler regions often align with greener (higher NDVI) areas — vegetation can reduce surface temperature through evapotranspiration.

We computed the empirical semivariograms for the variables in the dataset. The variables showed moderate spatial autocorrelation with low to moderate variance.

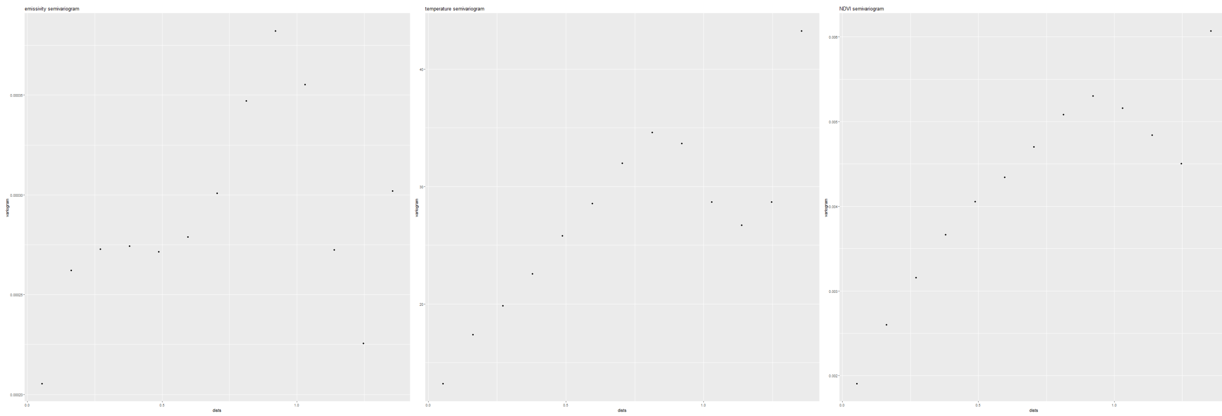

Figure 2: Empirical semivariograms for emissivity, temperature and NDVI

4.1 Model

4.2 Model fitting

Model Specification

We modeled land surface temperature (LST) using a spatially varying coefficient (SVC) framework:

$$\text{Temp}(s) = \beta_0(s) + \beta_{\text{NDVI}}(s) \cdot \text{NDVI} + \beta_{\text{Emis}}(s) \cdot \text{Emissivity} + \epsilon(s) \;\; \epsilon(s) \quad \sim \text{N}(0, \tau^2) \;\; \beta_r(s) \sim \text{GP}(0, \sigma_r^2 \exp(- \tag{1}$$

Prior specifications: We specified weakly informative priors

- Spatial variance: $\sigma_r^2 \sim \text{Inv-Gamma}(0.001, 0.001)$
- Spatial range: $\phi_r \sim \text{Uniform}(0.001, 500)$ km
- Nugget effect: $\tau^2 \sim \text{Inv-Gamma}(0.001, 0.001)$

The computational implementation featured Knot-based Approximation. We selected 1 knot per 10 grid cells ($k = 10$) using the `simpleknots()` function. Removed knots with missing data.

4.3 Results

Convergence Diagnostics

We made trace plots after burn-in to check for convergence.

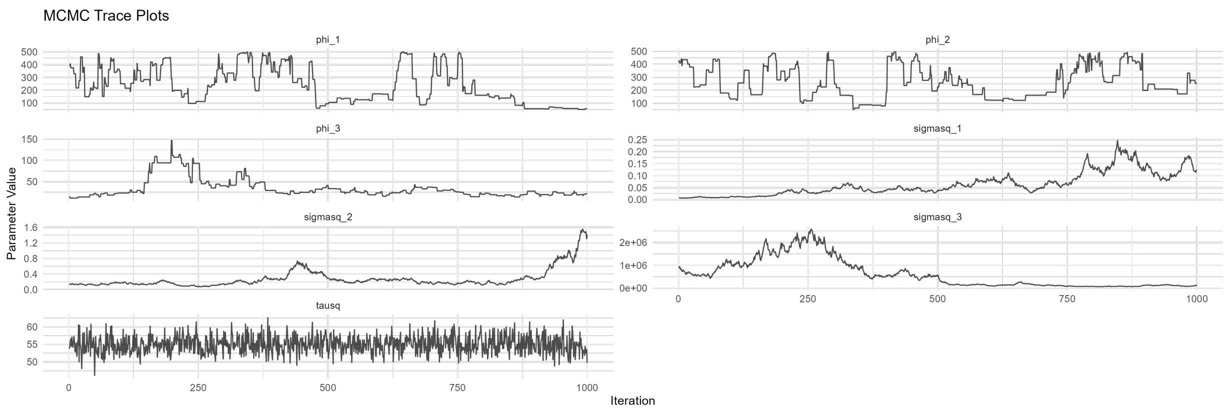

Figure 3: Trace plots for the parameters in the model

Some parameters display stable and well-mixed behavior, suggesting convergence, while others show trends or irregular jumps which was expected in our setting.

Interpretation of Coefficient Surfaces

Below is a plot of the posterior means of the coefficients over space.

The posterior mean NDVI coefficients show Strong negative effects (cooling) across the entire study region, with values ranging from -0.7310 to -0.7300. Each unit increase in NDVI corresponds to about $0.73°C$ decrease in land surface temperature across space. The emissivity coefficients reveal strong positive associations (mean $\beta_{Emis} \approx 330$) which implies that across space, a unit increase in emissivity increase temperature.

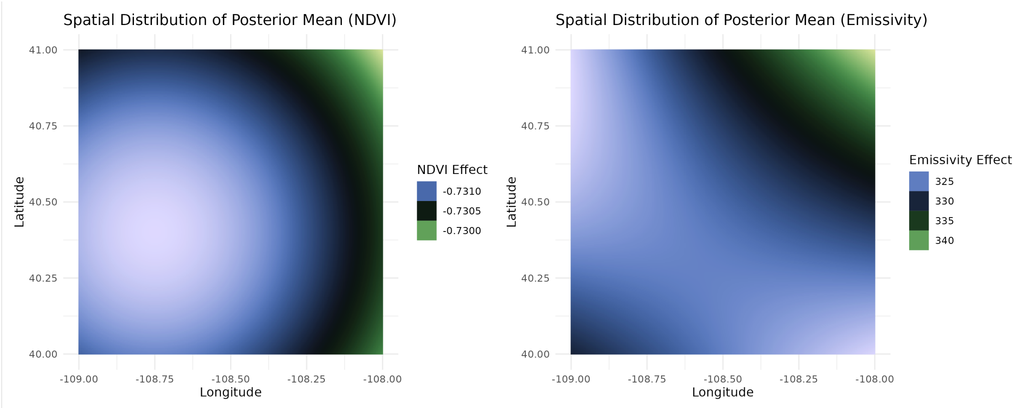

Figure 4: Spatially varying coefficient surfaces for (A) NDVI and (B) Emissivity effects on land surface temperature

5 Conclusion

Conclusion