

SVC: Bayesian Models for Spatially Varying Coefficients

Justice Akuoko-Frimpong, Jonathan Ta Edward Shao

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 $s_i \in \mathbb{R}^2$ for each $i = 1, \dots, n$ denote spatial locations where we observe response variables $Y(\mathbf{s})$ and covariates $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), \dots, X_p(\mathbf{s}))^\top$. The SVC model specification: We modeled spatial variation in outcomes through location-specific linear combinations of covariates, where each covariate’s effect is allowed to vary smoothly across space using Gaussian processes. This flexible approach captures complex, location-dependent relationships while accounting for measurement error through an independent noise term. Our computational framework combines three innovations to enable efficient Bayesian inference: (1) a low-rank Gaussian process approximation via strategic knot placement, reducing computational complexity

from cubic to near-linear scaling; (2) an adaptive Metropolis-Hastings sampler that automatically optimizes proposal distributions for spatial parameters during burn-in; and (3) optimized linear algebra operations using pre-computed distance matrices and Cholesky decompositions to accelerate covariance calculations. Together, these advances make spatially varying coefficient models practical for large-scale environmental datasets.

2 Methods

2.1 SVC Model

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

$$Y(\mathbf{s}) = \sum_{j=1}^p X_j(\mathbf{s})w_j(\mathbf{s}) + w_0(\mathbf{s}) + \epsilon(\mathbf{s}),$$

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

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

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

For each $w_r(\mathbf{s})$, 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 ϕ_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 $w_r(\mathbf{s})$, ϕ_r , σ_r^2 , and τ^2 , for $r = 0, \dots, p$ 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 $\tilde{Y}(\mathbf{s}) = Y(\mathbf{s}) - \sum_{j \neq r} X_j(\mathbf{s})w_j^{(t)}(\mathbf{s})$. Then

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

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

$$\Sigma_1 = (\Sigma_0^{-1} + \Sigma^{-1}) \quad \text{and} \quad \mu_1 = \Sigma_1 \Sigma^{-1} \frac{\tilde{Y}}{X_r}(\mathbf{s}).$$

2.2.3 Spatial Variance

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

2.2.4 Nugget

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

3 Simulation

The purposes of the simulation are to

- assess the performance of SVCLM,
- compare the differences in performance between full and low rank for SV-CLM,

- 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).

3.1.1 Spatial Priors

The spatial random effect w , and regression coefficients β_1 and β_2 , are modeled as Gaussian processes with constant mean vectors and spatial covariance matrices. These are defined using exponential covariance functions. Table 1 summarizes the prior distributions used.

Table 1: Priors for spatial random effects and regression coefficients.

Parameter	Variance (σ^2)	Range (ϕ)	Distribution
w	0.1	2	$\mathcal{N}(1_n, \sigma_w^2 \cdot \Phi_w(s, s'))$
β_1	0.1	2	$\mathcal{N}(1_n, \sigma_{\beta_1}^2 \cdot \Phi_{\beta_1}(s, s'))$
β_2	0.1	2	$\mathcal{N}(1_n, \sigma_{\beta_2}^2 \cdot \Phi_{\beta_2}(s, s'))$

Here, $\Phi(s, s')$ denotes a spatial covariance function (e.g., exponential or Matérn), and 1_n is a vector of ones of length n .

3.1.2 Covariate and Response Generation

The covariates and response variable are generated using the distributions and formulae shown in Table 2.

Table 2: Covariate and response generation.

Variable	Distribution / Formula
X_1	$\mathcal{N}(0, 1)$
X_2	$\mathcal{N}(5, 1)$
ϵ	$\mathcal{N}(0, \sigma_\epsilon^2)$, with $\sigma_\epsilon^2 = 0.1$
Y	$\beta_1 X_1 + \beta_2 X_2 + w + \epsilon$

3.2 Global Settings and Simulation Replication

To assess the performance of our estimation procedures and reduce sampling variability, the entire data generation and model fitting pipeline is replicated 1000 times. Table 3 summarizes the global settings used throughout the simulations.

Table 3: Global simulation settings.

Setting	Value / Description
Spatial grid size	$21 \times 21 = 441$ locations
Number of replications	1000
Error variance σ_ϵ^2	0.1
Covariance function	Exponential kernel

3.3 Comparisons with Other Packages

We compare the performance of our proposed method with `varycoef` and `spBayes` using bias, RMSE, and computational time (in seconds), averaged over 1000 simulation replications. The true values for β_1 , β_2 , and w are all set to 1.

Table 4: Comparison of methods based on bias, RMSE, and computation time (seconds). Placeholder values shown.

Method	β_1			β_2			w		
	Bias	RMSE	Time	Bias	RMSE	Time	Bias	RMSE	Time
Proposed Method	0.000	0.000	0.0	0.000	0.000	0.0	0.000	0.000	0.0
<code>varycoef</code>	0.000	0.000	0.0	0.000	0.000	0.0	0.000	0.000	0.0
<code>spBayes</code>	0.000	0.000	0.0	0.000	0.000	0.0	0.000	0.000	0.0

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.

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 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 sur-

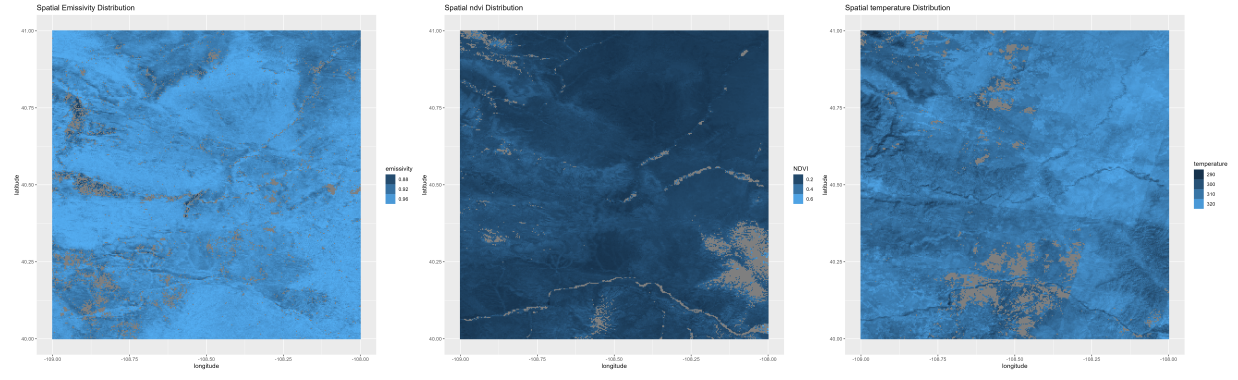


Figure 1: Spatial distributions of emissivity, NDVI, and land surface temperature in the study area.

face temperature through evapotranspiration.

We computed the empirical semivariograms for the variables in the dataset. The variables

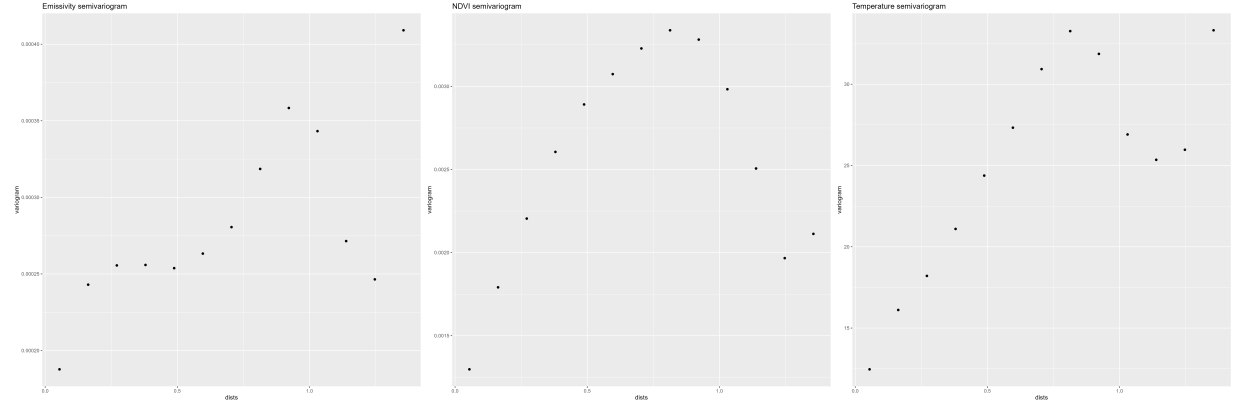


Figure 2: Empirical semivariograms for emissivity, temperature and NDVI

showed moderate spatial autocorrelation with low to moderate variance.

4.1 Model

4.2 Model fitting

Model Specification

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

framework:

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

Prior specifications: We specified weakly informative priors. Spatial variance; $\sigma_r^2 \sim \text{Inv-Gamma}(0.001, 0.001)$. Spatial range; $\phi_r \sim \text{Uniform}(0.001, 500)$ and the 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.

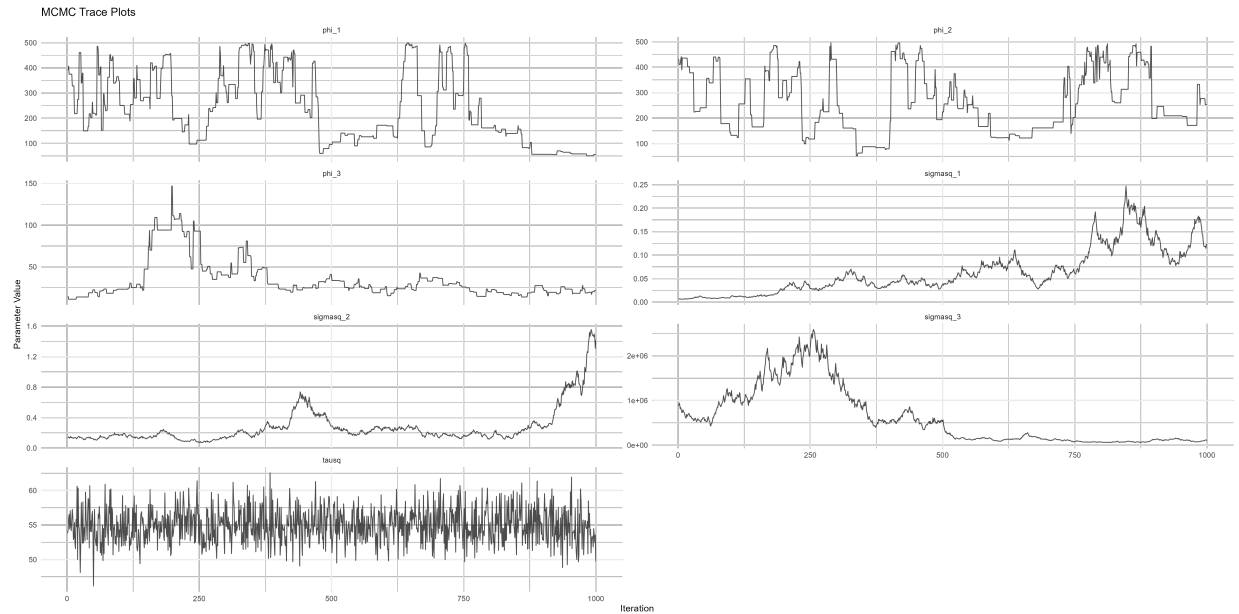


Figure 3: Trace plots for the parameters in the model

Some variables 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.

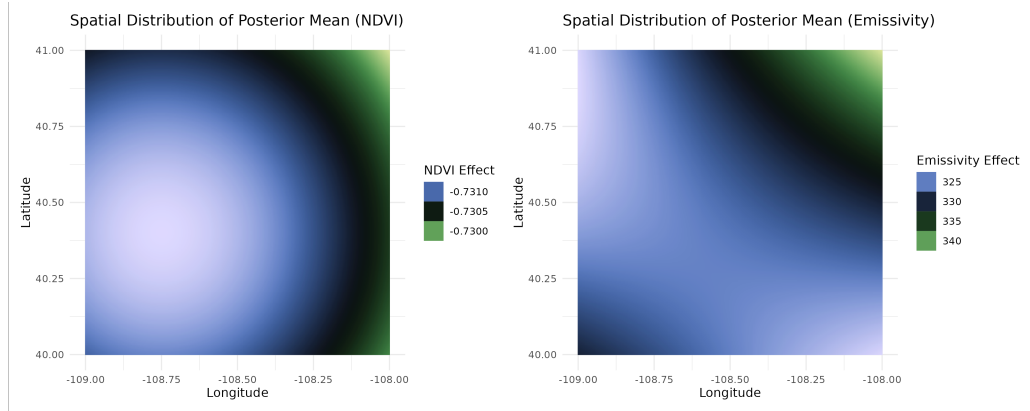


Figure 4: Spatially varying coefficient surfaces for (A) NDVI and (B) Emissivity effects on land surface temperature

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 $0.73^{\circ}C$ 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.

Predictions

We generated predicted temperature values across the entire spatial domain using the posterior mean estimates of the spatially varying coefficients. These predictions were computed for all locations with complete predictor data, regardless of whether the actual temperature was observed. The predicted values were obtained by multiplying the mean coefficient estimates with the corresponding covariate values at each location. We then visualized both

the observed and predicted temperature surfaces side by side to assess model performance and spatial prediction accuracy.

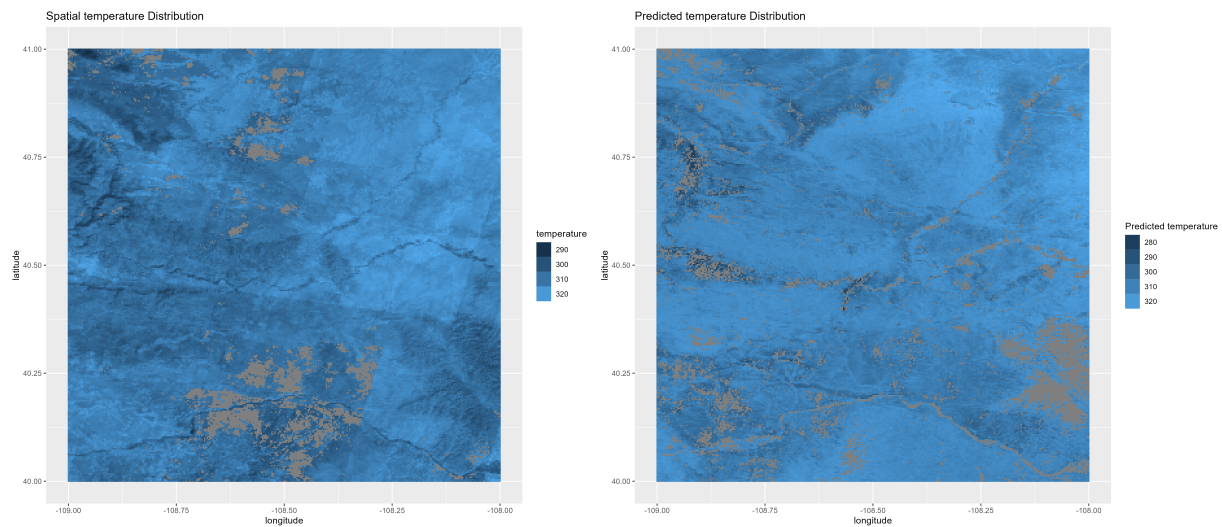


Figure 5: Side by side plot of tempearture and their predicted values across space

The comparison reveals the model’s ability to capture spatial patterns in temperature, even in regions where observations were missing but predictor data were available.

5 Conclusion

Conclusion