# Accounting for spatially varying directional effects in spatial covariance structures

Joaquim Henriques Vianna Neto, Universidade Federal de Juiz de Fora, Brazil
Alexandra M. Schmidt
Universidade Federal do Rio de Janeiro, Brazil
Peter Guttorp
University of Washington, USA
Corresponding author: J. H. Vianna Neto (joaquim.neto@ufjf.edu.br)

**Abstract**

We discuss how to include wind directional information in the covariance function of spatial models. Our models are based on a constructive convolution approach, wherein a spatial process is described as a convolution between a spatially varying smoothing kernel and a white noise process.

**Keywords:** Gaussian processes; non-stationarity; process convolution.

## 1 Introduction

### 1.1 Motivation

Usually, air pollution data are observed at fixed points (monitoring stations) of a region of interest. In spatial statistics, one usually assumes that observations are partial realizations of a stochastic process $\{Z(s), s \in G\}$, with $G \in \mathbb{R}^C$, where commonly $C = 2$, and the components of the location vector $s$ are geographical coordinates.

When modelling air pollutant data, assumptions of stationarity and isotropy do not seem realistic. It seems reasonable to expect local effects of covariates not only in the mean but also in the covariance structure of the spatial process of interest. In this paper, we show how the incorporation of wind direction in the covariance structure of a Gaussian process for ground level ozone improves its predictive performance when compared to other models that do not use this information. Figure 1 shows the locations of the monitoring sites and the observed data.

The main aim of this paper is to extend the convolution approach to allow the convolution kernel function to depend on the direction the wind is blowing at a particular location.

## 2 Proposed Models

Assume that observations are a partial realization of a random process $\{Z(s), s \in G\}$, with $G \in \mathbb{R}^p$, where usually $p = 1$, 2, or 3. More specifically, let

$$Z(s) = \mu(s) + Y(s) + \epsilon(s), \tag{1}$$

where $\mu(s)$ represents the mean structure of the process and usually is a function of location $s$, $\epsilon(s)$ is a Gaussian white noise process, such that $\epsilon(s) \sim N(0, \tau^2)$ for

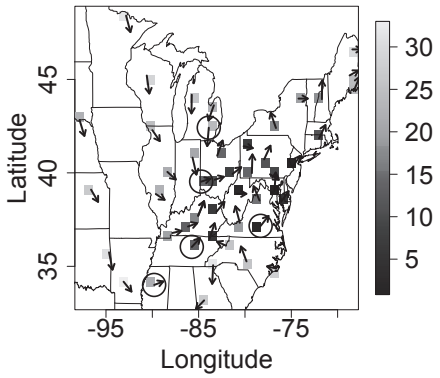



Figure 1: Observed values of ozone (solid squares in grayscale), and respective wind direction (arrows) for December, 11th, 2008, 3pm. For predictive purposes, circled locations were left out from the inference procedure.

any $s \in G$, and is usually present to capture measurement error and small scale dependence. We assume that $Y(s)$ is independent of $\epsilon(s)$, for all $s \in G$, and $Y(\cdot)$ represents a latent spatial process which captures any spatial structure left after adjusting for the effect of possible covariates included in $\mu(\cdot)$. We assume $Y(\cdot)$ is a Gaussian process described by

$$Y(s) = \int_G k_{s,x}(h)W(h)dh, \text{ for } s, h \in G \subset \mathbb{R}^2, \tag{2}$$

where $k_{s,x}(h)$ is a spatially varying kernel function that depends not only on the location $s$ but also on the observed covariate $x(s)$.

## 2.1 Accounting for wind direction in the nonstationary Matérn covariance function

Here we focus on the nonstationary version of the Matérn covariance function and introduce the covariate information in the kernel matrices, which we denote as $\Sigma(s, x)$. Let $x(s) = (u(s), v(s))'$ be a vector representing the directional covariate at location $s \in G$, with $||x(s)|| = 1$. We propose that the kernel matrices are modelled as

$$\begin{aligned} &\Sigma(s,x) = \Gamma(s,x)^T \Lambda \Gamma(s,x), \quad \text{where} \\ &\Gamma(s,x) = \begin{bmatrix} \cos\omega(x(s)) & \sin\omega(x(s)) \\ -\sin\omega(x(s)) & \cos\omega(x(s)) \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{bmatrix} \text{ and} \\ &\omega(x(s)) = \arctan\left(\frac{v(s)}{u(s)}\right). \end{aligned} \tag{3}$$

The matrix $\Gamma(s, x)$ plays the role of a rotation matrix, whose angle of rotation is given by the arctangent between the components of the wind vector at location $s$. This makes the Gaussian kernel at location $s$ coincide with the wind direction at $s$. On the other hand, the diagonal matrix $\Lambda$ captures the magnitude of the major and minor axes of the ellipses associated with the contours of the Gaussian kernel. As this process is based on the above matrix kernel we denote it as a *Locally Geometrically Anisotropic* (LGA) model.

## 2.2 An alternative kernel based on a projection measurement

We now propose an alternative way to account for a directional covariate in the kernel function $k_{s,x}(\cdot)$ of equation (2). We start by proposing a measurement that provides the degree of concordance between any two vectors.

**Projection of the average wind direction as a degree of concordance between vectors** Let $x(s)$ be defined as before, and define $r(x(s), x(s^*)) = (x(s) + x(s^*))/2$ to be the average vector when considering locations $s$ and $s^*$ in $G$. The projection of the average vector over the set $\{b \times (s - s^*) + s^* : b \in \mathbb{R}\}$ (a straight line that passes through $s$ and $s^*$) is

$$proj_x(s, s^*) = \frac{\langle r\left(x(s), x(s^*)\right), (s - s^*)\rangle}{\langle (s - s^*), (s - s^*)\rangle} \times (s - s^*),$$

where $\langle a_1, a_2\rangle$ represents the inner product between vectors $a_1$ and $a_2$. Figure 2 depicts how the projection captures the concordance between the vectors.

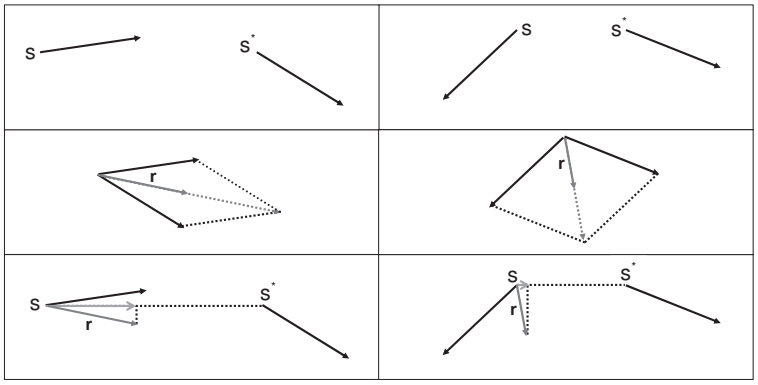



Figure 2: Diagram with directional covariates, black arrows (first row), average direction $r$ (second row), and resultant projection, gray arrows (third row).

Here, the kernel function is modelled as

$$k_{s,x}(h) = \frac{\sigma \alpha_{s,x}(h)}{\sqrt{\int\limits_G \alpha_{s,x}(h)^2 \, dh}}, \tag{4}$$

where $\alpha_{s,x}(\cdot)$ is a function that is specified below. We assume

$$\alpha_{s,x}(h) = \begin{cases} \exp\left(-\frac{\|s-h\|}{\phi_1 + \phi_2 \|proj_x(s,h)\|}\right), & \text{if } s \neq h \\ 1, & \text{if } s = h. \end{cases} \tag{5}$$

In order to guarantee an exponential decay of the kernel we assume $\phi_1 > 0$ and $\phi_2 > 0$. This way, $\phi_1$ measures the effect of the Euclidean distance, and $\phi_2$ the effect of the directional covariate in the correlation structure of $Y(\cdot)$. We call a process based on this kernel $\alpha_{s,x}(\cdot)$ a *Projection Model*.

The kernel in (5) does not provide a closed form of the covariance structure of $Y(\cdot)$. For this reason, we resort to the discretized version of this process and assume $Y(s) = \sum_{l=1}^{m} k_{s,x}(h_l) W(h_l)$. Therefore, all the integrals are substituted by their discretized versions.

### 2.3 An alternative model for $\Sigma(s)$ of the nonstationary Matérn covariance function

It is reasonable to compare this approach with those proposed by Higdon et al. (1999) or Paciorek and Schervish (2006), which are able to capture more flexible covariance structures. Following Paciorek and Schervish (2006), we construct each $\Sigma(s)$ using the spectral decomposition, that is, we assume $\Sigma(s) = \Gamma(s)^T \Lambda(s) \Gamma(s)$. The matrix $\Lambda(s)$ is a diagonal matrix of eigenvalues, $\Lambda(s) = \begin{bmatrix} \lambda_1(s) & 0 \\ 0 & \lambda_2(s) \end{bmatrix}$. Different from Paciorek and Schervish (2006), we let $\Gamma(s) = \begin{bmatrix} \cos(\theta(s)) & \sin(\theta(s)) \\ -\sin(\theta(s)) & \cos(\theta(s)) \end{bmatrix}$.

The kernel matrices are expected to vary smoothly across the spatial domain. In our proposal this is guaranteed through the prior specification of $\lambda_j(s)$, for $j = 1, 2$, and $\theta(s)$. We assume $\log \lambda_j(\cdot)$ follow independent Gaussian processes, with mean $\mu_{\lambda_j}$, variance $\sigma^2_\lambda$ and a squared exponential covariance function, that is, $Cov(\log \lambda_j(s), \log \lambda_j(s^*)) = \sigma^2_\lambda \exp\left\{-(||s - s^*||/\phi_\lambda)^2\right\}$. Next, we model $\theta(s)$ as

$$\theta(s) = \frac{\pi}{2} \Phi(\gamma(s)), \tag{6}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. We assume $\gamma(s)$ follows a Gaussian process with mean $\mu_\gamma$ and a squared exponential covariance function, that is, $Cov(\gamma(s), \gamma(s^*)) = \sigma^2_\gamma \exp\left\{-(||s - s^*||/\phi_\gamma)^2\right\}$.

Note that $\theta(s)$ varies smoothly across the region induced by the GP for $\gamma(s)$. Under the parametrization in equation (6), $\theta(s) \in [0, \pi/2]$.

## 3 Data Analysis

We fit 5 different models to this dataset.

**Fitted Models** The notation of the fitted models is the following:

M1 Isotropic model with Matérn covariance function

M2 Elliptical anisotropic model with Matérn covariance function, that is

$$Cov(Y(s_i), Y(s_j)) = \sigma^2 \left(2^{\nu-1}\Gamma(\nu)\right)^{-1} \left(\frac{\sqrt{\|u\|^T \Lambda \|u\|}}{\varphi}\right)^{\nu} \kappa_\nu \left(\frac{\sqrt{\|u\|^T \Lambda \|u\|}}{\varphi}\right),$$

where

$$\Lambda = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

where $0 < \theta < \pi/2$, $\kappa_\nu(\cdot)$ denotes the modified Bessel function of the third type and order $\nu$, $\Gamma(\cdot)$ is the usual gamma function, $\varphi > 0$ is a parameter related to the decay of the correlation and $\nu > 0$ determines the smoothness of the process.

M3 Nonstationary Matérn covariance function with $\Sigma(s, x)$ as in equation (4)

M4 Covariance function based on the Projection model

M5 Nonstationary Matérn covariance function with $\Sigma(s)$

For all models we fixed the smoothness parameter of the Matérn covariance function ($\nu$) at 1.

| Model | Computational time | No. Iterations per minute |
|---|---|---|
| M1 | 7 min | 7142.85 |
| M2 | 1 h 26 min | 581.39 |
| M3 | 1 h 55 min | 434.78 |
| M4 | 21 min | 2380.95 |
| M5 | 144 h 30 min | 80.73 |

Table 1: Computational time, and number of iterations per minute, to run the MCMC algorithm for 50,000 iterations for models M1, M2, M3, M4, and 700,000 iterations for model M5.

Table 2 shows the values of four different model comparison criteria, PPL, DIC, the predictive likelihood based on the circled locations shown in Figure 1, and the mean squared error (MSE). The MSE was computed using the mean of the posterior predictive distribution as the fitted value.

| Model | DIC | | | PPL | | | Predictive likelihood | MSE |
|---|---|---|---|---|---|---|---|---|
| | $\overline{D}$ | $p_D$ | DIC | G | P | $D_1$ | | |
| M1 | 326.06 | 3.04 | 329.10 | 458.77 | 1395.79 | 1625.18 | $2.33 \times 10^{-08}$ | 131.10 |
| M2 | 318.70 | -1.80 | 316.37 | 251.37 | 1050.26 | 1175.95 | $6.18 \times 10^{-08}$ | 100.53 |
| M3 | 309.46 | 3.88 | 313.34 | 87.80 | 639.06 | 682.96 | $6.49 \times 10^{-07}$ | 45.33 |
| M4 | 289.79 | 4.03 | 293.82 | 90.38 | 464.93 | 510.12 | $4.93 \times 10^{-06}$ | 25.80 |
| M5 | 302.29 | 3.03 | 305.32 | 59.97 | 539.60 | 569.59 | $9.98 \times 10^{-08}$ | 70.46 |

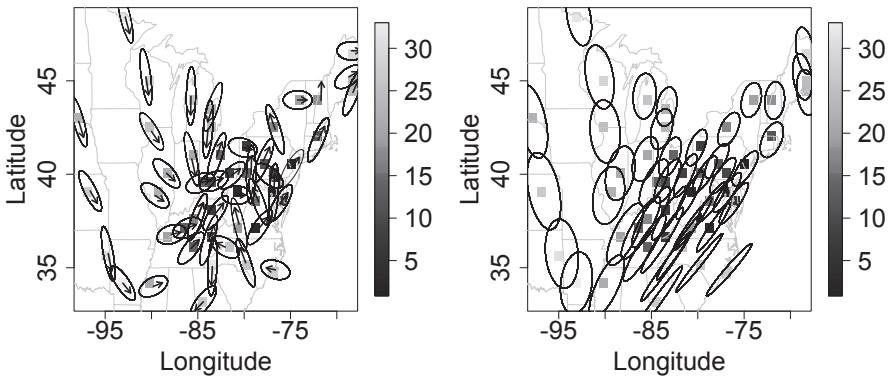
Table 2: Model comparison criteria: DIC, PPL, the predictive likelihood based on the circled locations in Figure 1, and the mean squared error, under each fitted model.

Figure 3 shows the posterior mean of the resultant ellipses obtained under models M3 and M5. This helps to understand how the spatial correlations are changing across the region under models M3 and M5.

The behavior of the proposed correlation functions in Sections 2.1 and 2.2 can be visualized through the panels of Figure 4.

# 4 Discussion

We have suggested two different approaches to consider wind information in the covariance structure of environmental processes observed at fixed locations. The challenge is how to account for the influence of the wind direction while still guaranteeing that the resultant spatial covariance structure is positive definite. Borrowing ideas from the convolution approach of Higdon et al. (1999), and Paciorek and Schervish (2006), we include the directional covariate in the kernel function of their convolution approach.




(a) Estimated ellipses under M3 (b) Estimated ellipses under M5

Figure 3: Left panel: wind vectors and respective estimated ellipses under model M3. Right panel: Estimated ellipses under model M5.

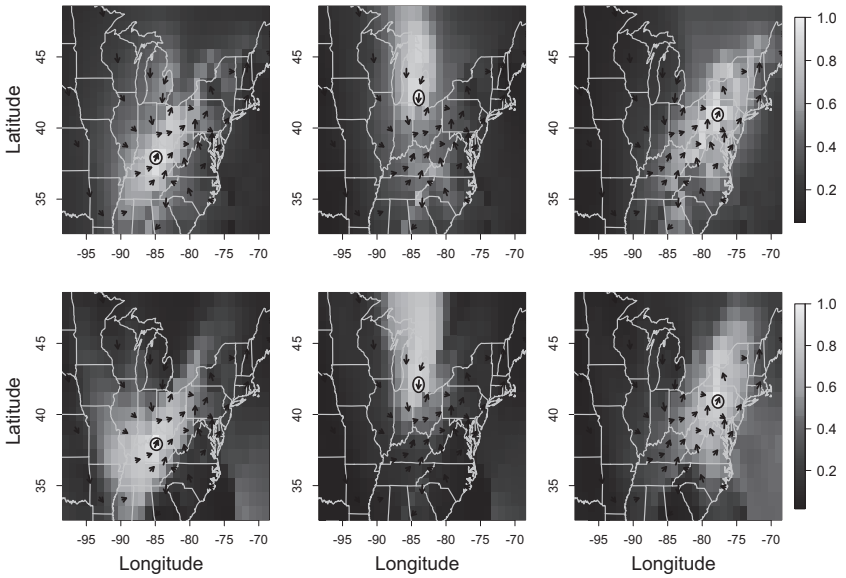



Figure 4: Posterior mean of the estimated correlation between the location marked by a circle and all the others in the grid, under models M3 (first row) and M4 (second row).

## References

Higdon, D., Swall, J. and Kern, J. (1999). Non-stationary spatial modeling, *Bayesian Statistics 6* pp. 761–768.

Paciorek, C. J. and Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions, *Environmetrics* **17**(5): 483–506.