# Nonstationary spatial process modeling
# Part II

**Paul D. Sampson --- Catherine Calder**
**Univ of Washington --- Ohio State University**

this presentation derived from that presented at the
Pan-American Advanced Study Institute on
Spatio-Temporal Statistics
Búzios, RJ, Brazil
June 16-26, 2014

# Nonstationary Models II

## A GENERAL MODELING FRAMEWORK

- Let $Z(\cdot)$ be a realization of a spatial stochastic process defined for all $\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^d$, where $d$ is typically equal to 2 or 3
- We observe the value of $Z(\cdot)$ at a finite set of locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n \in \mathcal{D}$ and wish to learn about the underlying process
- For all $\boldsymbol{s} \in \mathcal{D}$, let

$$Z(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Y(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

  where

  - $\mu(\cdot)$ is a deterministic mean function
  - $Y(\cdot)$ is a mean-zero latent spatial process
  - $\epsilon(\cdot)$ is a spatially independent error process, which is assumed to be independent of $Y(\cdot)$

# Nonstationary Models II

*Definition* A process is said to be second-order stationary if

$$\mathrm{E}[Y(\boldsymbol{s})] = \mathrm{E}[Y(\boldsymbol{s}+\boldsymbol{h})] = \mu$$

and

$$\mathrm{cov}[Y(\boldsymbol{s}), Y(\boldsymbol{s}+\boldsymbol{h})] = \mathrm{cov}[Y(\boldsymbol{0}), Y(\boldsymbol{h})] = C(\boldsymbol{h})$$

where the function $C(\boldsymbol{h})$, $\boldsymbol{h} \in \mathbb{R}^d$ is called the covariance function

→ Here, $Y(\cdot)$ is a nonstationary spatial process with covariance function $C(\boldsymbol{s}_1, \boldsymbol{s}_2) = \mathrm{cov}(Y(\boldsymbol{s}_1), Y(\boldsymbol{s}_2))$

# Nonstationary Models II

- We focus on modeling $C(\boldsymbol{s}_1, \boldsymbol{s}_2)$:
  1. has to be a valid covariance function
  2. has to be estimable (perhaps from only a single realization of the process)
- Following Sampson (2010)'s categorization, the following are a few approaches in the literature...
  1. Smoothing and weighted-average methods
  2. Basis function methods
  3. Process convolutions / spatially-varying parameters

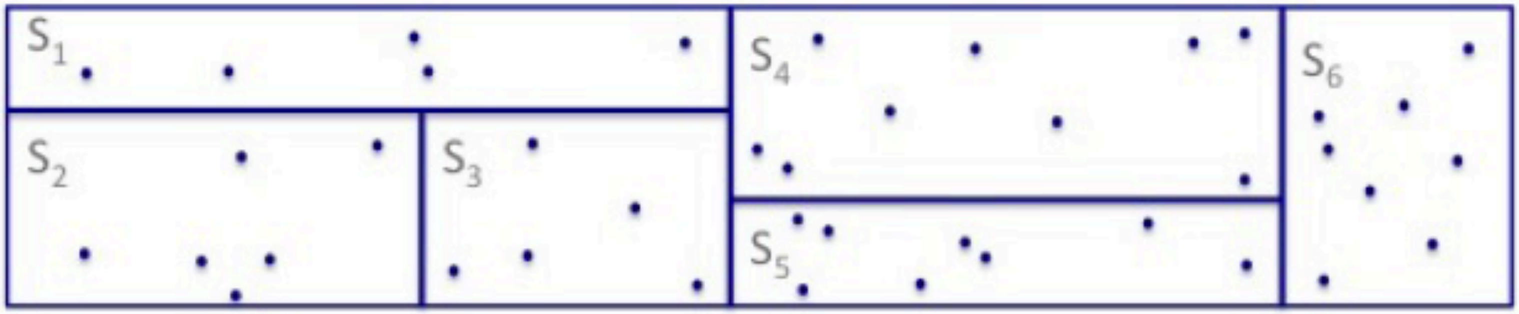
# Nonstationary Models II

## 1. SMOOTHING / WEIGHTED-AVERAGE METHODS

**Idea:** Construct a nonstationary spatial process by smoothing several locally stationary processes

**An example:** (Fuentes, 2001):

- Divide the spatial region $\mathcal{D}$ into $k$ disjoint subregions $S_i$, for $i = 1, \ldots, k$, such that $\mathcal{D} = \cup_{i=1}^{k} S_i$
- Let $Y_1(\cdot), Y_2(\cdot), \ldots, Y_k(\cdot)$ be stationary spatial processes associated with each of the subregions, with covariance functions estimated using the observations in each subregion



# Nonstationary Models II

- Construct a global nonstationary process as a weighted average of the locally stationary processes:

  $$Y(\boldsymbol{s}) = \sum_{i=1}^{k} w_i(\boldsymbol{s}) Y_i(\boldsymbol{s}),$$

  where $w_i(\boldsymbol{s})$ is weight function based on the distance between $\boldsymbol{s}$ and the 'center' of region $S_i$

- The number of subregions is chosen using BIC

# Nonstationary Models II

**Some other approaches:**

- Fuentes and Smith (2002) propose a continuous extension of the original model where

$$Y(\boldsymbol{s}) = \int_{\mathcal{D}} w(\boldsymbol{s} - \boldsymbol{u}) Y_{\theta(u)}(\boldsymbol{s}) d\boldsymbol{u}$$

- Nott and Dunsmuir (2002) propose letting

$$C(Y(\boldsymbol{s}_1), Y(\boldsymbol{s}_2)) = \Sigma_0 + \sum_{i=1}^{k} \underbrace{w_i(\boldsymbol{s}_1) w_i(\boldsymbol{s}_2) C_{\theta_i}(\boldsymbol{s}_1 - \boldsymbol{s}_2)}_{\text{local } \textit{residual} \text{ covariance structure}}$$

- Guillot et al. (2001) propose a nonparametric kernel estimator of a nonstationary covariance matrix

- Kim, Mallick, and Holmes (2005)'s approach automatically partitions the spatial domain into disjoint regions and then fits a piecewise Gaussian process model

# Nonstationary Models II

## 2. BASIS FUNCTION MODELS

**Idea:** decompose spatial covariance functions in terms of basis functions

**An example:** EOFs

- The Karhunen-Loéve (K-L) expansion of a covariance function is

$$C_Y(\boldsymbol{s}_1, \boldsymbol{s}_2) = \sum_{k=1}^{\infty} \lambda_k \phi_k(\boldsymbol{s}_1) \phi_k(\boldsymbol{s}_2)$$

where $\{\phi_k(\cdot) : k = 1, \ldots, \infty\}$ and $\{\lambda_k : k = 1, \ldots, \infty\}$ are the eigenfunctions and eigenvalues, respectively, of the Fredholm integral equation:

$$\int_{\mathcal{D}} C_Y(\boldsymbol{s}_1, \boldsymbol{s}_2) \phi_k(\boldsymbol{s}) d\boldsymbol{s} = \lambda_k \phi_k(\boldsymbol{s}_2)$$

# Nonstationary Models II

- Using this expansion, we can write the process as

$$Y(\boldsymbol{s}) = \sum_{k=1}^{\infty} a_k \phi_k(\boldsymbol{s}).$$

- It can be shown that the truncated decomposition

$$Y_p(\boldsymbol{s}) = \sum_{k=1}^{p} a_k \phi_k(\boldsymbol{s})$$

  is optimal in the sense that it minimizes the variance of the truncation error among all sets of basis function representations of $Y(\cdot)$ of order $p$.

- The $\phi_k(\boldsymbol{s})$s can be obtained numerically by solving the Fredholm integral equation (can be difficult).

# Nonstationary Models II

- An alternative solution when repeated observations of the spatial process (e.g., over time) are available: perform a principal components analysis of the empirical covariance matrix

  That is, if $\hat{\mathbf{\Sigma}}_Y$ is the empirical covariance matrix, we can solve the eigensystem

  $$\hat{\mathbf{\Sigma}}_Y \mathbf{\Phi} = \mathbf{\Phi} \mathbf{\Lambda},$$

  where

  - $\mathbf{\Phi}$ is the matrix of eigenvectors → called the "empirical orthogonal functions" or EOFs
  - $\mathbf{\Lambda}$ is the diagonal matrix with corresponding eigenvalues on the diagonal

# Nonstationary Models II

- We can use $\mathbf{\Phi}\boldsymbol{\alpha}$ in place of $\boldsymbol{Y} = (Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n))'$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)'$ are a collection of unknown parameters

  → typically, truncated version of this approach are used for dimension reduction

---

*Advantages* of using EOFs:

1. naturally nonstationary

*Disadvantages* of using EOFs:

1. prediction
2. measurement error

# Nonstationary Models II

**Some other examples:**

- Holland et al. (1998) represents a nonstationary spatial covariance function as the sum of a stationary model and a finite sum of EOFs
- Nychka (2002) uses multiresolution wavelets instead of EOFs for computational reasons. More recent work by Matsuo, Nychka, and Paul (2008) has extended the approach to handle irregularly spaced data
- Pintore and Holmes (2004) and Stephenson et al. (2005) induce nonstationarity by evolving the stationary power spectrum with a latent spatial power process
- Katzfuss (2014) propose a model with a low-rank representation of a nonstationary Matérn (with covariance tapering) model for computational considerations

# Nonstationary Models II

## 3. PROCESS CONVOLUTION MODELS / SPATIALLY-VARYING PARAMETERS

**Idea:** use a constructive specification of a (Gaussian) process to introduce nonstationarity

**An example:** (Higdon, 1998)

- Let $k(\cdot) : \mathbb{R}^d \to \mathbb{R}$ be a function satisfying

$$\int_{\mathbb{R}^d} k(\boldsymbol{u}) d\boldsymbol{u} < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} k^2(\boldsymbol{u}) d\boldsymbol{u} < \infty$$

  and $W(\cdot)$ denotes $d$-dimensional Brownian motion.

# Nonstationary Models II

- It can be shown that the process

$$Y(\boldsymbol{s}) = \int_{\mathbb{R}^d} k_{\boldsymbol{s}}(\boldsymbol{u}) W(d\boldsymbol{u})$$

is Gaussian with $\mathrm{E}[Y(\boldsymbol{s})] = 0$ and

$$C_Y(\boldsymbol{s}_1, \boldsymbol{s}_2) = \mathrm{cov}[Y(\boldsymbol{s}_1), Y(\boldsymbol{s}_2)] = \int_{\mathbb{R}^d} k_{\boldsymbol{s}_1}(\boldsymbol{u}) k_{\boldsymbol{s}_2}(\boldsymbol{u}) d\boldsymbol{u}$$

for $\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^d$

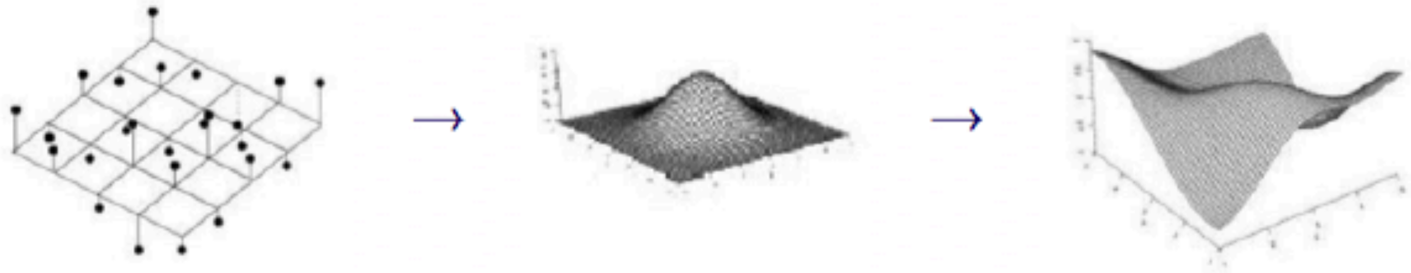
If the kernels $k_s(u)$ are of fixed shape, such as Gaussian kernels varying only in location, the covariance is stationary, a function only of $| s_1 - s_2 |$.
If the parameters of the kernels, such as orientation and anisotropy of elliptical contours, vary in space, we have a nonstationary model.

# Nonstationary Models II

- Higdon (1998) proposes a discrete approximation to a nonstationary Gaussian process:

$$Y(\boldsymbol{s}) = \sum_{i=1}^{k} k_{\boldsymbol{s}}(\boldsymbol{u}_i)\, x_i$$

where the $x_i$'s are i.i.d. $\mathrm{N}(0, \lambda^2)$ random variables associated with each knot location $\boldsymbol{u}_i$.



# Nonstationary Models II

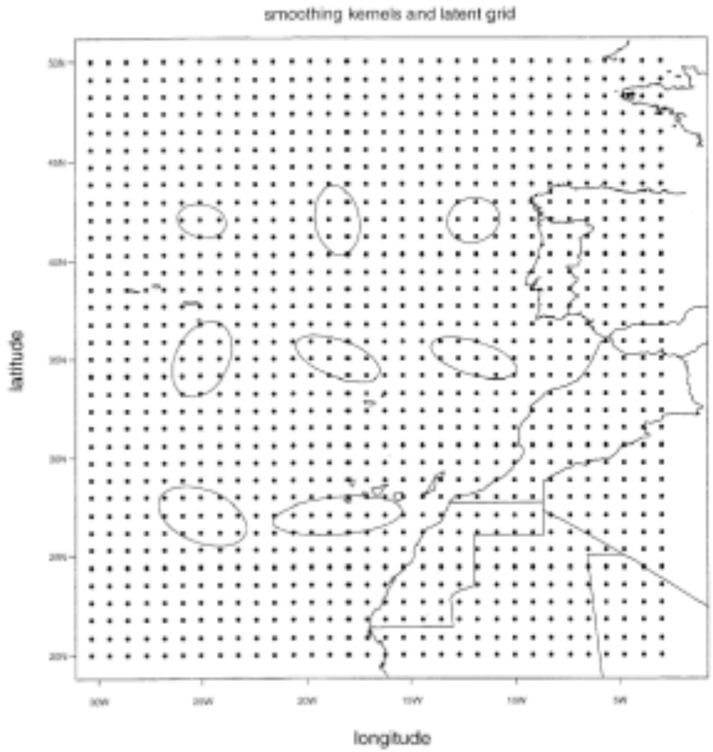
- Higdon (1998) proposes using this model for North Atlantic ocean temperatures. In this model, the kernels were weighted averages of fixed ‘basis kernels’

$$Y(\boldsymbol{s}) = \sum_{i=1}^{k} k_{\boldsymbol{s}}(\boldsymbol{u}_i)\, x_i$$

where

$$k_{\boldsymbol{s}}(u_i) = \sum_{j=1}^{8} w_j(\boldsymbol{s}) k_{\boldsymbol{s}_j^*}(\boldsymbol{u}_i)$$

$$w_j(\mathbf{s}) \propto \exp\left(-\frac{1}{2}||\boldsymbol{s} - \mathbf{s}_\mathbf{j}^*||^2\right)$$



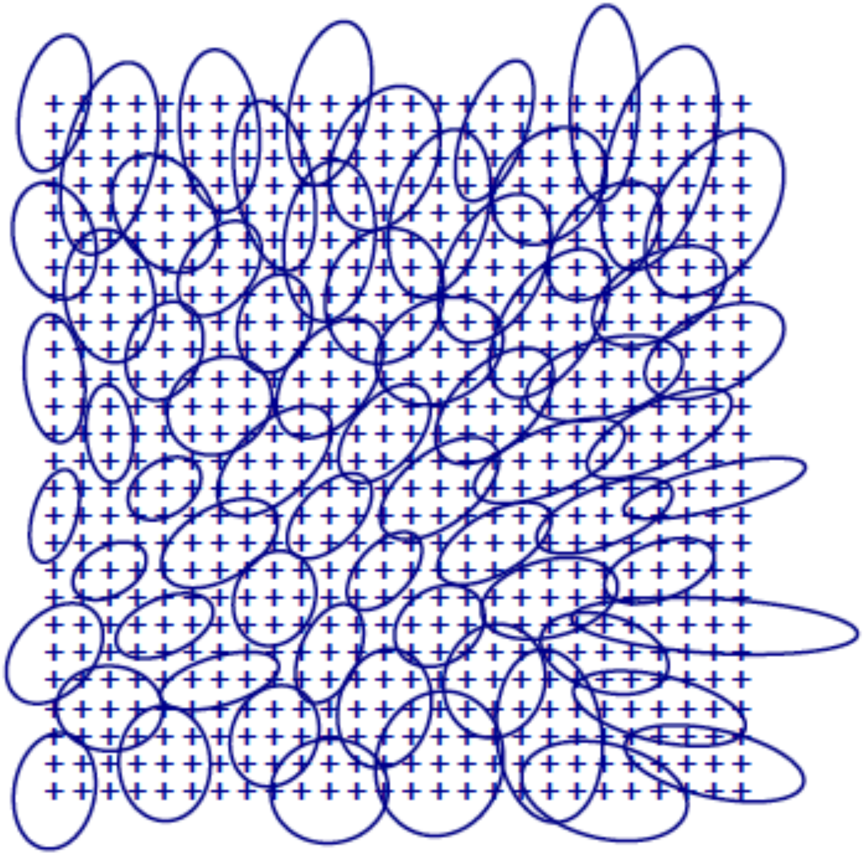
(Higdon, 1998)

$$k_{\boldsymbol{s}_j^*}(\boldsymbol{u}_i) = \frac{1}{\sqrt{2\pi}} |\boldsymbol{\Sigma}_{\boldsymbol{s}_j^*}|^{-1} \exp\left(-\frac{1}{2}(\boldsymbol{s}_j^* - \boldsymbol{u}_i)' \boldsymbol{\Sigma}_{\boldsymbol{s}_j^*}^{-1} (\boldsymbol{s}_j^* - \boldsymbol{u}_i)\right)$$

# Nonstationary Models II

**Some other examples:**

- Kernel parameters can vary smoothly in space (Higdon, Swall, and Kern, 1999; Paciorek and Schervish, 2006):



# Nonstationary Models II

- A famous result (Thiebaux 1976; Thiebaux and Pedder 1987) uses a parametric class of Gaussian kernel functions in Equation 2 to give a closed-form covariance function; this result was later extended (Paciorek 2003; Paciorek and Schervish 2006; Stein 2005) to show that

$$C^{NS}(\mathbf{s}, \mathbf{s}'; \boldsymbol{\theta}) = \sigma(\mathbf{s})\sigma(\mathbf{s}') \frac{|\boldsymbol{\Sigma}(\mathbf{s})|^{1/4} |\boldsymbol{\Sigma}(\mathbf{s}')|^{1/4}}{\left|\frac{\boldsymbol{\Sigma}(\mathbf{s})+\boldsymbol{\Sigma}(\mathbf{s}')}{2}\right|^{1/2}} g\left(\sqrt{Q(\mathbf{s}, \mathbf{s}')}\right),$$

  is a valid, nonstationary, parametric covariance function on $R^d$; $d \geq 1$, when $g()$ is chosen to be a valid correlation function on $R^d$; $d \geq 1$. Note that this equation no longer requires kernel functions to be speci ed.
  $\theta$ is a generic parameter vector, $\sigma(\cdot)$ represents a spatially-varying standard deviation, $\Sigma(\cdot)$ is a $d \times d$ matrix that represents the spatially-varying local anisotropy (controlling both the range and direction of dependence), and

$$Q(\mathbf{s}, \mathbf{s}') = (\mathbf{s} - \mathbf{s}')^\top \left(\frac{\boldsymbol{\Sigma}(\mathbf{s}) + \boldsymbol{\Sigma}(\mathbf{s}')}{2}\right)^{-1} (\mathbf{s} - \mathbf{s}')$$

# Nonstationary Models II

$Q(s, s')$ above is a Mahalanobis distance. Furthermore, choosing $g(\cdot)$ to be the Matern correlation function also allows for the introduction of $\kappa(s)$, a spatially-varying smoothness parameter (Stein 2005; in this case, the Matern correlation function in in the above equation has smoothness $[\kappa(s) + \kappa(s')]$/2.

While this equation no longer requires the notion of kernel convolution, we refer to $\Sigma(\cdot)$ as the kernel matrix, since it was originally defined as the covariance matrix of a Gaussian kernel function (Thiebaux 1976).

- Kleiber and Nychka (2012) further extend this model to the multivariate setting.
- Calder (2007, 2008) proposes space-time versions of the Higdon model.
- Heaton (2014) extends process convolution models to spherical domains.

# Nonstationary Models II

## SUMMARY

- lots of models → some have been well studied, some haven't
- very little work on model comparison
- with the exception of the basis function models, computation is a BIG challenge
- no general software
- recent work has focused on understanding the reasons for nonstationarity (e.g., covariates)

# Nonstationary Models II

To appear in the *Journal of Statistical Software*:

Local Likelihood Estimation for Covariance Functions with Spatially-Varying Parameters: The convoSPAT Package for R

Mark D. Risser

Catherine A. Calder

# Nonstationary Models II

$$Z(s) = x(s)'\boldsymbol{\beta} + Y(s) + \epsilon(s)$$

where $Y(s)$ is a spatially dependent, mean zero, Gaussian process with covariance function $C^{NS}$ defined above, and $\epsilon(s) \sim N(0, \tau^2(s))$ is measurement error with (possibly) spatially varying variance.

Let $\boldsymbol{\theta}$ represent the vector of all the variance-covariance parameters for $Y(s)$ and $\epsilon(s)$. Then

$$\boldsymbol{Z}|\boldsymbol{Y}, \boldsymbol{\beta}, \boldsymbol{\theta} \sim N_n(\boldsymbol{X\beta} + \boldsymbol{Y}, \boldsymbol{D}(\boldsymbol{\theta}))$$

where the ith row of $\boldsymbol{X}$ is $x(s_i)$ and $\boldsymbol{D}(\boldsymbol{\theta})$ is the diagonal matrix with elements $\tau^2(s_i)$. Integrate out the latent process $\boldsymbol{Y}$ and we have the *marginal likelihood* of the observed data $\boldsymbol{Z}$ given all the parameters

$$\boldsymbol{Z}|\boldsymbol{\beta}, \boldsymbol{\theta} \sim N_n(\boldsymbol{X\beta}, \boldsymbol{D}(\boldsymbol{\theta}) + \boldsymbol{\Omega}(\boldsymbol{\theta}))$$

where $\boldsymbol{\Omega}(\boldsymbol{\theta})$ has elements

$$\Omega_{ij}(\boldsymbol{\theta}) = C^{NS}(s_i, s_j; \boldsymbol{\theta})$$

the latter being specified by the parameters of the spatial correlation function $g(\cdot)$ and the spatially varying $\Sigma(\cdot), \sigma(s), \tau^2(s)$, and/or $\kappa(s)$, if the Matern is used.

For a particular application, the practitioner can specify the underlying correlation structure (through the correlation function $g(\cdot)$) as well as determine which of $\Sigma(\cdot), \sigma(s), \tau^2(s),$ and/or $\kappa(s)$ should be fixed or allowed to vary spatially.

However, some care should be taken in choosing which quantities should be spatially-varying: for example, Anderes and Stein (2011) note that allowing both $\Sigma(\cdot)$ and $\kappa(s)$ to vary over space leads to issues with identiability.

# Nonstationary Models II

To reduce the computational demands of fitting this model, Risser & Calder use the *discretized basis kernel* approach of Higdon (1998). The estimated Gaussian kernel function at any specified location is a weighted average of "basis" kernel functions, estimated locally over the region of interest.

Define *mixture component locations*, typically on a regular grid, with parameters $\{\phi_k = (\Sigma_k, \sigma^2{}_k, \tau^2{}_k, \kappa_k): k = 1, \cdots, K\}$. Then the parameter set for arbitrary location $s$ is calculated as:

$$\phi(s) = \sum_{k=1}^{K} \omega_k(s)\ \phi_k,$$

$$\omega_k(s) \propto exp\left\{-\frac{\|s - b_k\|^2}{2\lambda_\omega}\right\}$$

For example, the kernel matrix for location $s$ is $\Sigma(s) = \sum_{k=1}^{K} \omega_k(s)\, \Sigma_k$ .

We must specify the tuning parameter $\lambda_\omega$ as well as the size and spacing of the grid of mixture locations. The modeler chooses which parameters should be spatially-varying: the kernel matrices, the process variance, the nugget variance, and the smoothness.

Having done so, the number of parameters is linear in $K$, the number of mixture component locations, instead of $n$, the sample size.

# Nonstationary Models II

Prediction proceeds by the usual conditional Gaussian calculations using plug-in estimates of the parameters $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\theta}}$, computed by local likelihood estimation, as explained below.

Risser and Calder consider a number of out-of-sample evaluation criteria:

$$MSPE = \frac{1}{m}\sum_{j=1}^{m}\left(z_j^* - \hat{z}_j^*\right)^2$$

$$pMSDR = \frac{1}{m}\sum_{j=1}^{m}\frac{\left(z_j^* - \hat{z}_j^*\right)^2}{\widehat{\sigma}_j}$$

And a continuous rank probability score, CRPS (Gneiting and Raftery 2007), which measures the fit of the predictive density. Larger CRPS (smaller negative values) indicates better model fit.

# Nonstationary Models II

## Computationally efficient inference.

Fast and efficient inference for a nonstationary process convolution model has yet to be made readily available for general use. The equation for the nonstationary spatial covariance, $C^{NS}(s, s'; \theta)$, requires some kind of constraints and has suffered from a lack of widespread use due to the complexity of the requisite model fitting and limited pre-packaged options. Focusing on the spatially-varying local anisotropy matrices, the covariance function requires a kernel matrix at every observation and prediction location of interest.

Paciorek and Schervish (2006) accomplish this by modeling $\Sigma(\cdot)$ as itself a (stationary) stochastic process, assigning Gaussian process priors to the elements of the spectral decomposition of $\Sigma(\cdot)$; alternatively, Katzfuss (2013) uses a basis function representation of $\Sigma(\cdot)$. Both of these models are highly parameterized and require intricate Markov chain Monte Carlo methods for model fitting.

# Nonstationary Models II

## Computationally efficient inference.

Risser and Calder achieve computational efficiency using

(1) the discrete mixture representation above, and its requisite specification of the size/spacing of the basis grid, and the tuning parameter for the weight function, and

(2) the idea of using local likelihood estimation (Tibshirani and Hastie, 1987), rather than aim to optimize the full log-likelihood.

# Nonstationary Models II

## Computationally efficient inference.

First, recall REML estimation. The full log likelihood is

$$\mathcal{L}^F(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{Z}) = -\frac{1}{2}\log|\boldsymbol{\Omega} + \mathbf{D}| - \frac{1}{2}(\mathbf{Z} - \mathbf{X}\boldsymbol{\beta})^\top(\boldsymbol{\Omega} + \mathbf{D})^{-1}(\mathbf{Z} - \mathbf{X}\boldsymbol{\beta}),$$

The “restricted” log likelihood (based on “n-p” linear combinations having expected value zero for all possible parameter values, can be written.

$$\mathcal{L}^R(\boldsymbol{\theta}; \mathbf{Z}) = -\frac{1}{2}\log|\boldsymbol{\Omega} + \mathbf{D}| - \frac{1}{2}\log|\mathbf{X}^\top(\boldsymbol{\Omega} + \mathbf{D})^{-1}\mathbf{X}| - \frac{1}{2}\mathbf{Z}^\top\mathbf{P}\mathbf{Z},$$

(See the paper for specification of the matrix P.) We maximize this and estimate $\beta$ by

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^\top(\hat{\boldsymbol{\Omega}} + \hat{\mathbf{D}})^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top(\hat{\boldsymbol{\Omega}} + \hat{\mathbf{D}})^{-1}\mathbf{Z},$$

# Nonstationary Models II

## Computationally efficient inference.

With Local Likelihood Estimation (LLE), rather than optimize the restricted likelihood directly, we maximize it on neighborhoods for each mixture component $b_k$, a neighborhood depending on a radius $r$, the "span" or window size for each mixture component:

$$N_k = \{s_i \in \{s_1, \cdots, s_n\} : \|s_i - b_k\| \leq r\}$$

and

$$Z_{N_k} = \{Z(s) : s \in N_k\}$$

Note: The restricted likelihood to be optimized for each neighborhood will be based on a stationary version of the spatial model:

$$\widetilde{Z}(\mathbf{s}) = \mathbf{x}(\mathbf{s})^\top \widetilde{\boldsymbol{\beta}} + \widetilde{Y}(\mathbf{s}) + \widetilde{\epsilon}(\mathbf{s}),$$

where $\tilde{Y}$ is a stationary process with covariance function

$$C^S(\mathbf{s} - \mathbf{s}') = \sigma^2 g\left(||\boldsymbol{\Sigma}^{-1/2}(\mathbf{s} - \mathbf{s}')||\right)$$

# Nonstationary Models II

## Computationally efficient inference.

Note: The kernel matrices are parameterized in terms of of the eigenvalues and angle of rotation of its spectral decomposition

$$\mathbf{\Sigma} = \begin{bmatrix} \cos(\eta) & -\sin(\eta) \\ \sin(\eta) & \cos(\eta) \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos(\eta) & \sin(\eta) \\ -\sin(\eta) & \cos(\eta) \end{bmatrix}$$

The full model can be fit after plugging REML estimates into the covariance function $C^{NS}(s, s'; \theta)$ using the discrete basis representation
to calculate the likelihood for the observed data.

# Nonstationary Models II

## Computationally efficient inference.

This is a very nice computational method, but it requires specification of many moving parts:

- the *number* and *placement* of *mixture component locations*,
- selecting *which of the spatial dependence parameters should be fixed* or allowed to vary spatially,
- the *tuning parameter* $\lambda$ for the weighting function $w$,
- the *fitting radius* r for the local likelihood estimation.

Parameter estimates for this model are likely to be sensitive to the choice of $K$ and the placement of mixture component locations. Tibshirani and Hastie (1987) discuss the importance of choosing the radius *r,* suggesting that the model should be fit using a range of $r$ values, and use a global criterion such as the maximized overall likelihood, cross-validation, or Akaike's Information Criterion to choose the final model. This strategy could either be implemented on a trial-and-error basis or in an automated scheme. Of course, regardless of the number and locations of the mixture component centroids, the radius r should be chosen such that a large enough number of data points are used to estimate a local stationary model.

# The `convoSPAT` package for R

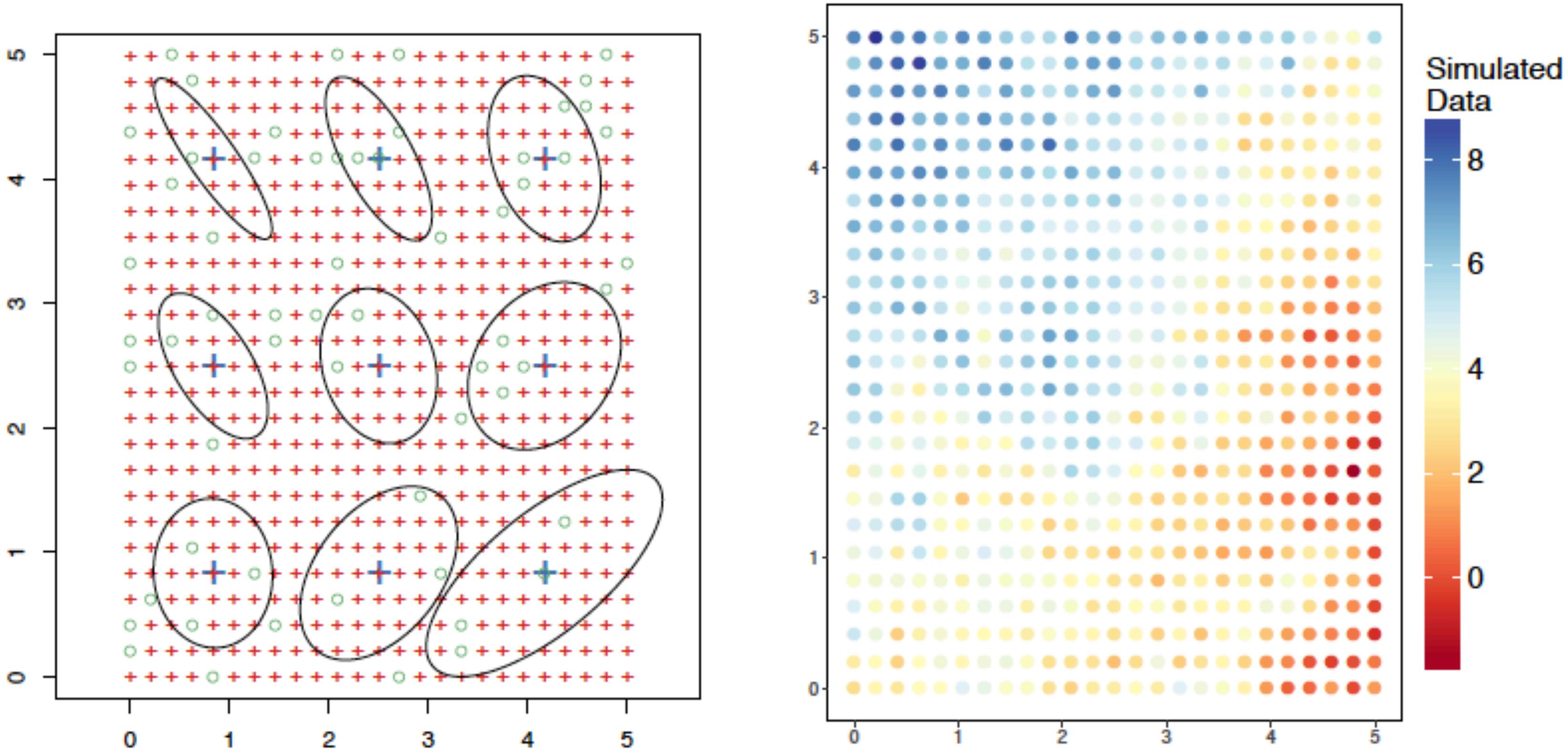


Figure 1: Left: true mixture component ellipses with observation locations (red) and validation locations (green). Right: simulated data.

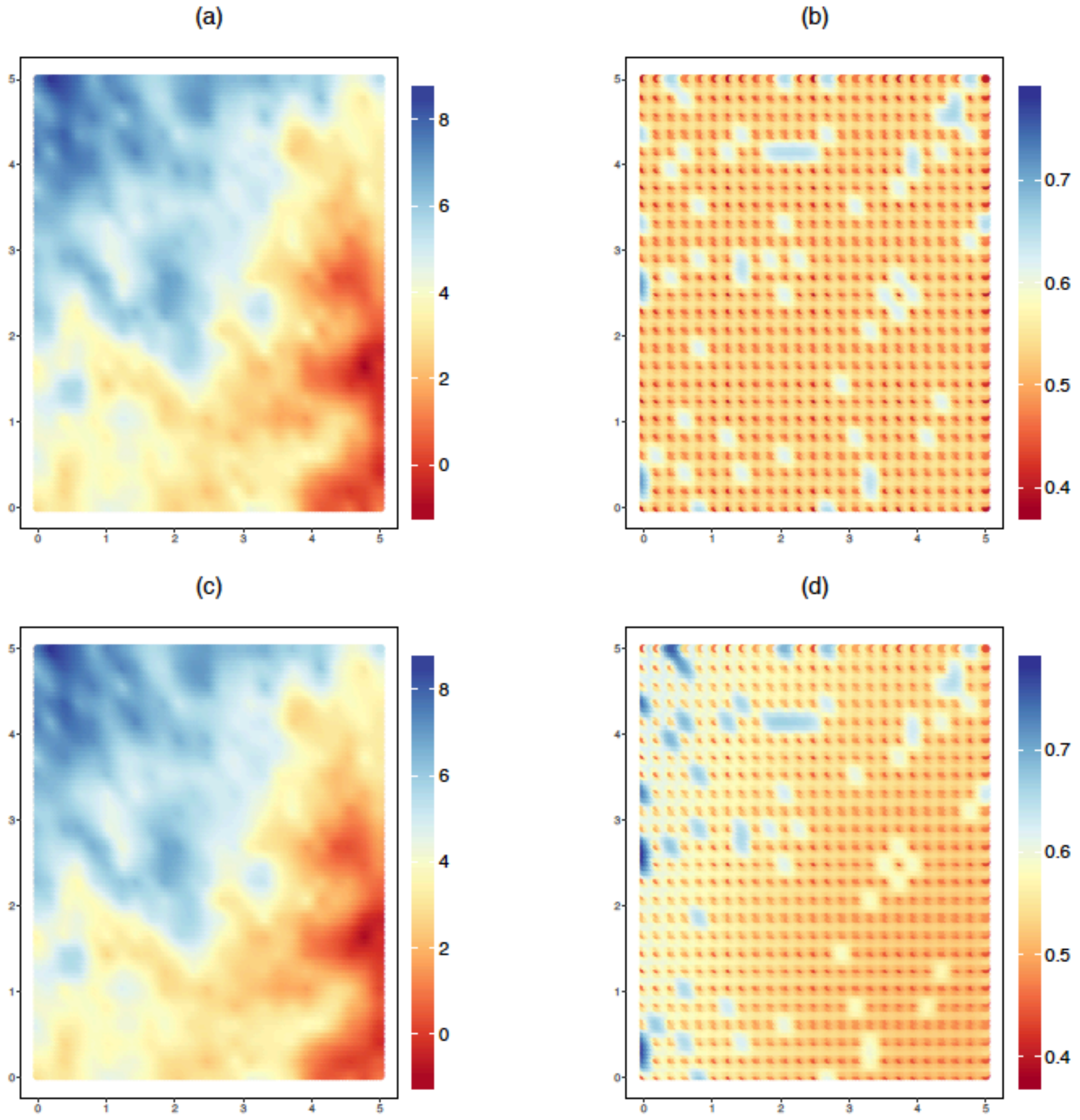


Figure 2: Predictions and prediction errors from the stationary model (a. and b.) and the nonstationary model (c. and d.).