Spatial smoothing using Gaussian processes

Chris Paciorek

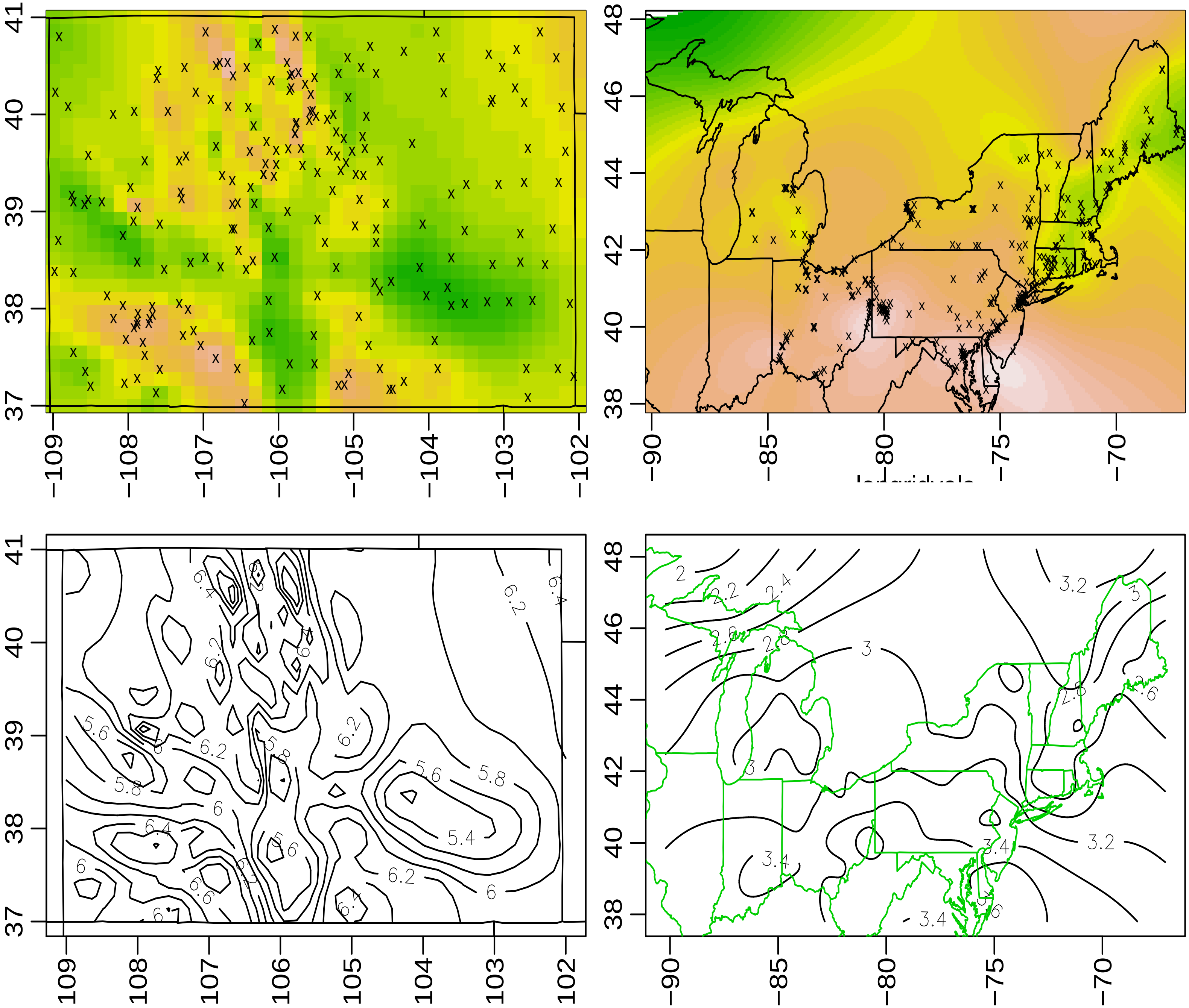
paciorek@hsph.harvard.edu

August 5, 2004

OUTLINE

- Spatial smoothing and Gaussian processes
- Covariance modelling
- Nonstationary covariance modelling
- Examples with simulated and real data

SPATIAL DATA



Colorado annual precipitation

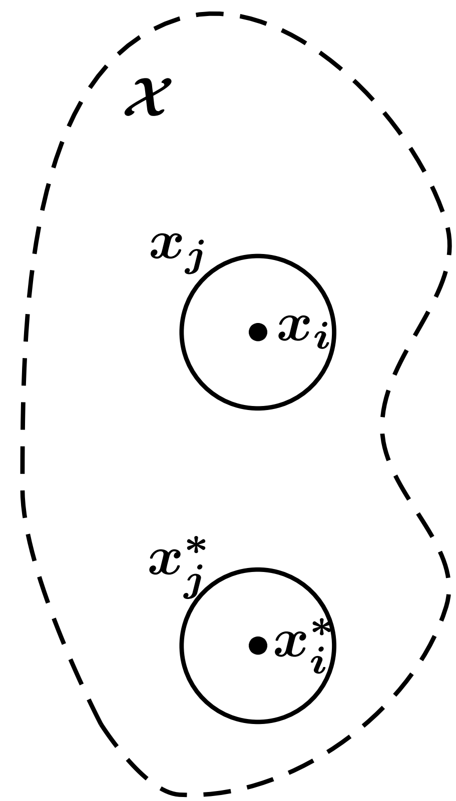
monthly PM10

Spatial data and models

- Features of spatial data
  - ❖ highly interactive, complex surfaces
  - ❖ high df:n; heterogeneous coverage
  - ❖ possible nonstationarity
  - ❖ spatial model is often a component of a larger model
    - ✥ spatiotemporal
    - ✥ model with covariates
    - ✥ possibly replicated
- Some standard models
  - ❖ Thin-plate splines
  - ❖ Basis function models with penalized coefficients
  - ❖ Gaussian processes (kriging or Bayesian)

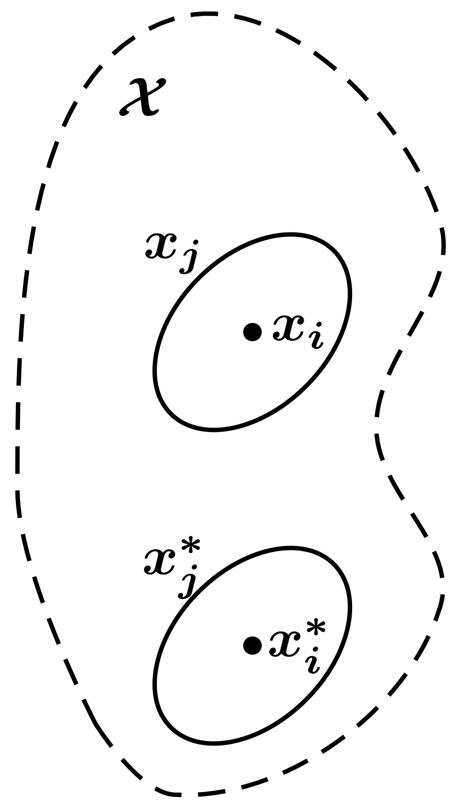
GAUSSIAN PROCESS DISTRIBUTION

- Infinite-dimensional joint distribution for $f(x), \; x \in \mathcal{X}$:
  - ❖ Example: $f(\cdot)$ a spatial process, $\mathcal{X} = \Re^2$
  - ❖ $f(\cdot) \sim \mathrm{GP}(\mu(\cdot), C(\cdot, \cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $C(f(x_i), f(x_j))$:
  - ❖ stationary, isotropic
  - ❖ stationary, anisotropic
  - ❖ nonstationary



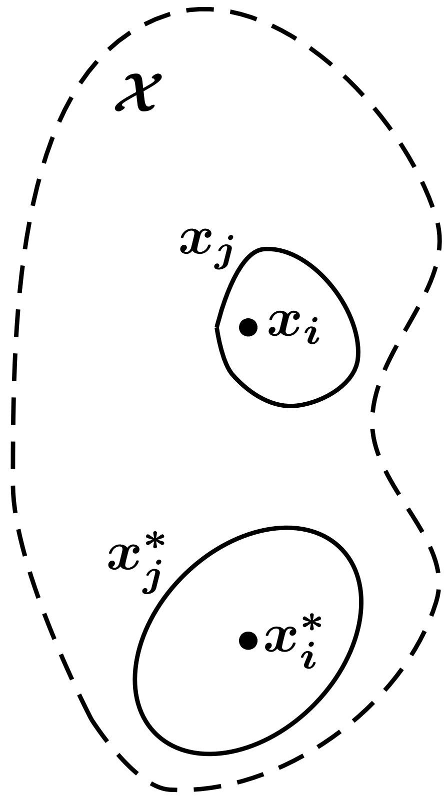
Gaussian process distribution

- Infinite-dimensional joint distribution for $\boldsymbol{f(x)},\ \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:
  - ❖ Example: $\boldsymbol{f}(\cdot)$ a spatial process, $\boldsymbol{\mathcal{X}} = \Re^{\mathbf{2}}$
  - ❖ $\boldsymbol{f}(\cdot) \sim \mathrm{GP}(\boldsymbol{\mu}(\cdot), \boldsymbol{C}(\cdot,\cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $\boldsymbol{C}(\boldsymbol{f}(\boldsymbol{x_i}), \boldsymbol{f}(\boldsymbol{x_j}))$:
  - ❖ stationary, isotropic
  - ❖ stationary, anisotropic
  - ❖ nonstationary



GAUSSIAN PROCESS DISTRIBUTION

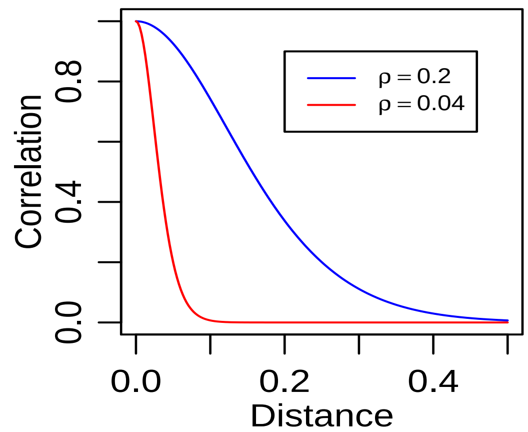
- Infinite-dimensional joint distribution for $\boldsymbol{f}(\boldsymbol{x}),\ \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:
  - ❖ Example: $\boldsymbol{f}(\cdot)$ a spatial process, $\boldsymbol{\mathcal{X}} = \Re^2$
  - ❖ $\boldsymbol{f}(\cdot) \sim \mathrm{GP}(\boldsymbol{\mu}(\cdot), \boldsymbol{C}(\cdot,\cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $\boldsymbol{C}(\boldsymbol{f}(\boldsymbol{x}_i), \boldsymbol{f}(\boldsymbol{x}_j))$:
  - ❖ stationary, isotropic
  - ❖ stationary, anisotropic
  - ❖ nonstationary

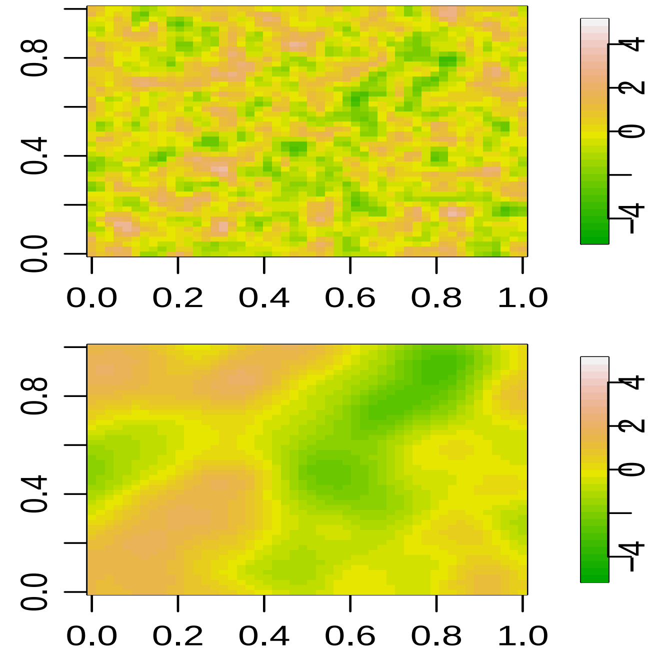


Correlation scale

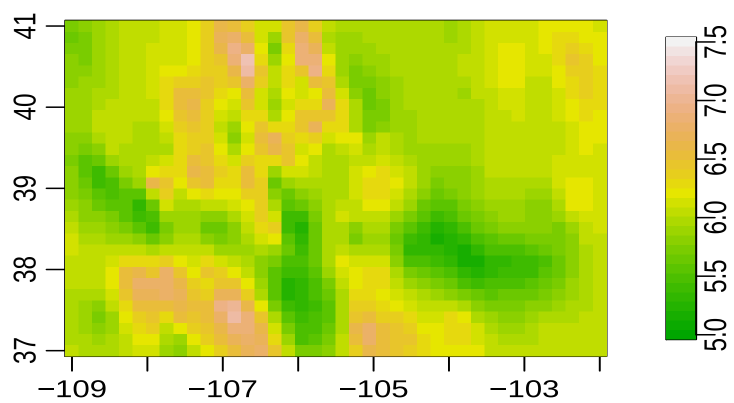
$$R(f(x_i), f(x_j)) = R\left(\tau_{ij}^2\right)$$

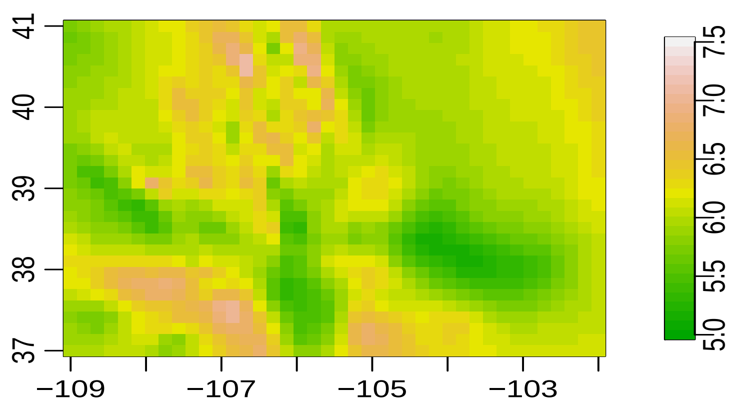
$$\tau_{ij}^2 = \frac{(x_i - x_j)^T (x_i - x_j)}{\rho^2}$$





Nonstationarity: correlation a function of location





NONSTATIONARY COVARIANCE

- spatial deformation (Sampson, Guttorp & co-workers)
- mixtures of Gaussian processes (Fuentes) or splines (Wood et al.)
- kernel convolution method (Higdon, Swall, and Kern)

❖ $R^{NS}(f(x_i), f(x_j)) = c_{ij} \int_{\Re^P} k_{x_i}(u) k_{x_j}(u) du$

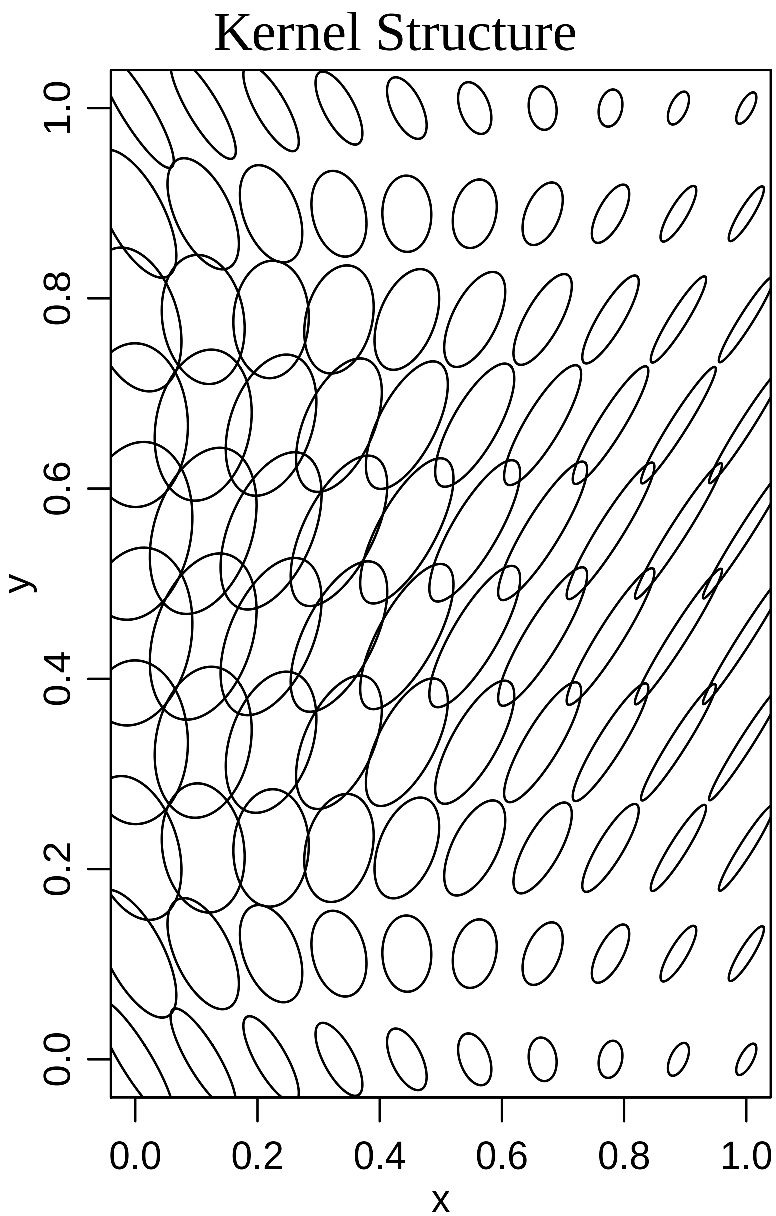
❖ Guaranteed positive definite

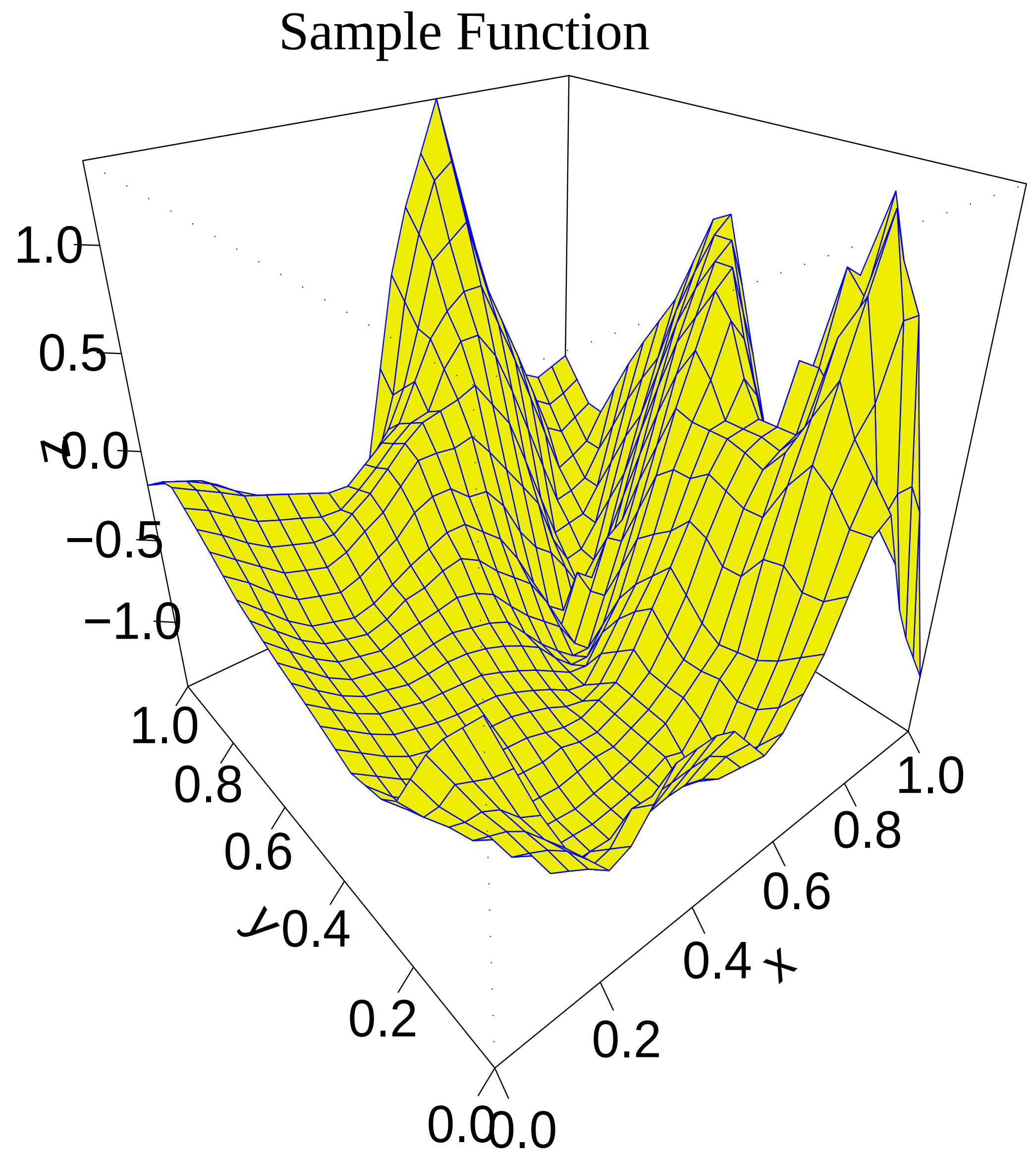
❖ Gaussian kernels:

$$\begin{aligned} k_{x_i}(u) &\propto \exp\left(-(u - x_i)^T \Sigma_i^{-1} (u - x_i)\right) \\ R^{NS}(f(x_i), f(x_j)) &= c_{ij} \exp\left(-(x_i - x_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (x_i - x_j \right. \end{aligned}$$

❖ $f(\cdot) \sim \mathrm{GP}(\mu, \sigma^2 R^{NS}(\cdot, \cdot))$

NONSTATIONARY GPS IN 2D





- Squared exponential form:

$$\exp\left(-\left(\frac{\tau_{ij}}{\rho}\right)^2\right) \Rightarrow c_{ij}\exp\left(-(x_i-x_j)^T\left(\frac{\Sigma_i+\Sigma_j}{2}\right)^{-1}(x_i-x_j)\right)$$

- Theorem 1: if $R(\tau)$ is positive definite for $\Re^P$, $P = 1, 2, \ldots$, then

$$R^{NS}(f(x_i), f(x_j)) = c_{ij}R\left(\sqrt{Q_{ij}}\right)$$

  is positive definite for $\Re^P$, $P = 1, 2, \ldots$

- Theorem 2: Smoothness properties of original stationary correlation retained

- Specific case of Matérn nonstationary covariance

$$\frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{2\sqrt{\nu\tau_{ij}^2}}{\rho}\right)^{\nu}K_{\nu}\left(\frac{2\sqrt{\nu\tau_{ij}^2}}{\rho}\right) \Rightarrow \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(2\sqrt{\nu Q_{ij}}\right)^{\nu}K_{\nu}\left(2\sqrt{\nu Q_{ij}}\right)$$

  More flexible form, differentiability not constrained, possible asymptotic advantages

A BASIC BAYESIAN SPATIAL MODEL

- Model:

$$\begin{aligned} \boldsymbol{Y_i} &\sim \boldsymbol{N(f(x_i), \eta^2)},\ \boldsymbol{x_i} \in \boldsymbol{\Re^2} \\ \boldsymbol{f(\cdot)} &\sim \mathrm{GP}(\boldsymbol{\mu}, \boldsymbol{\sigma^2 R^{NS}}(\cdot, \cdot; \boldsymbol{\nu}, \boldsymbol{\Sigma}_{(\cdot)})) \end{aligned}$$

- Let $\boldsymbol{R^{NS}}$ be the nonstationary Matérn correlation
- Kernels ($\boldsymbol{\Sigma_x}$) constructed based on stationary GP priors

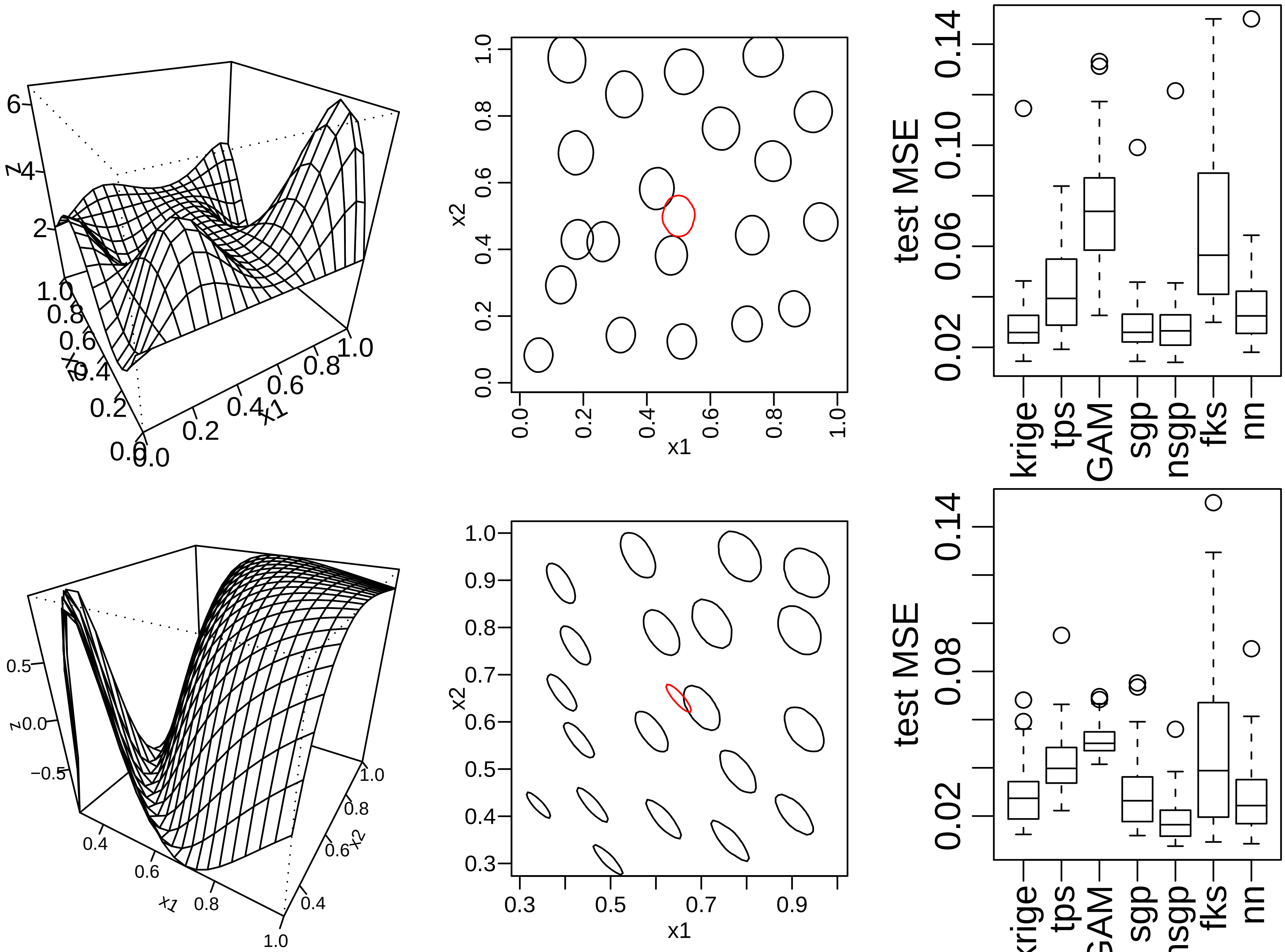
  Define multiple kernel matrices, $\boldsymbol{\Sigma_x},\ \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$

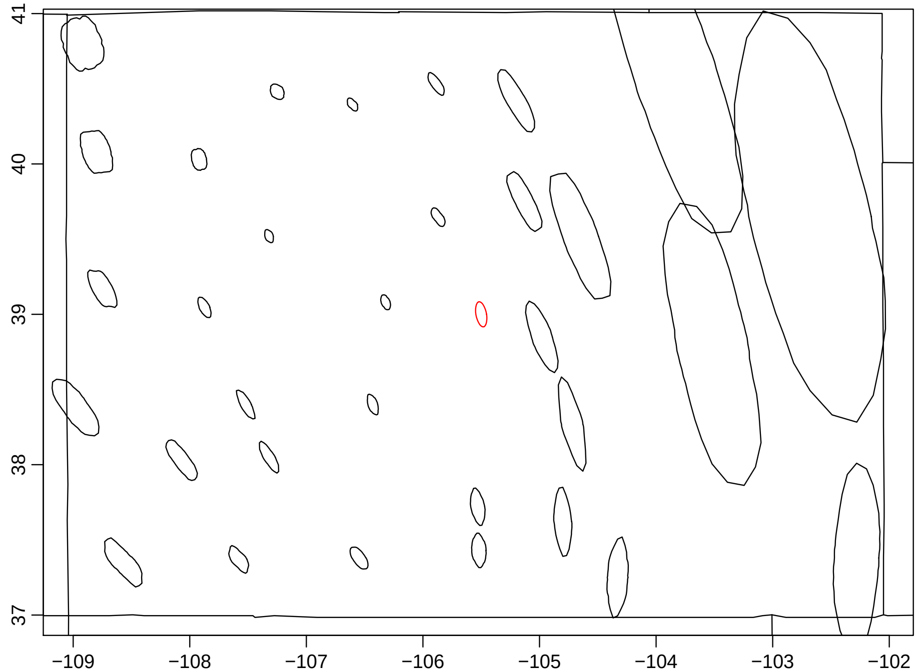
  Smoothly-varying (element-wise) in covariate space

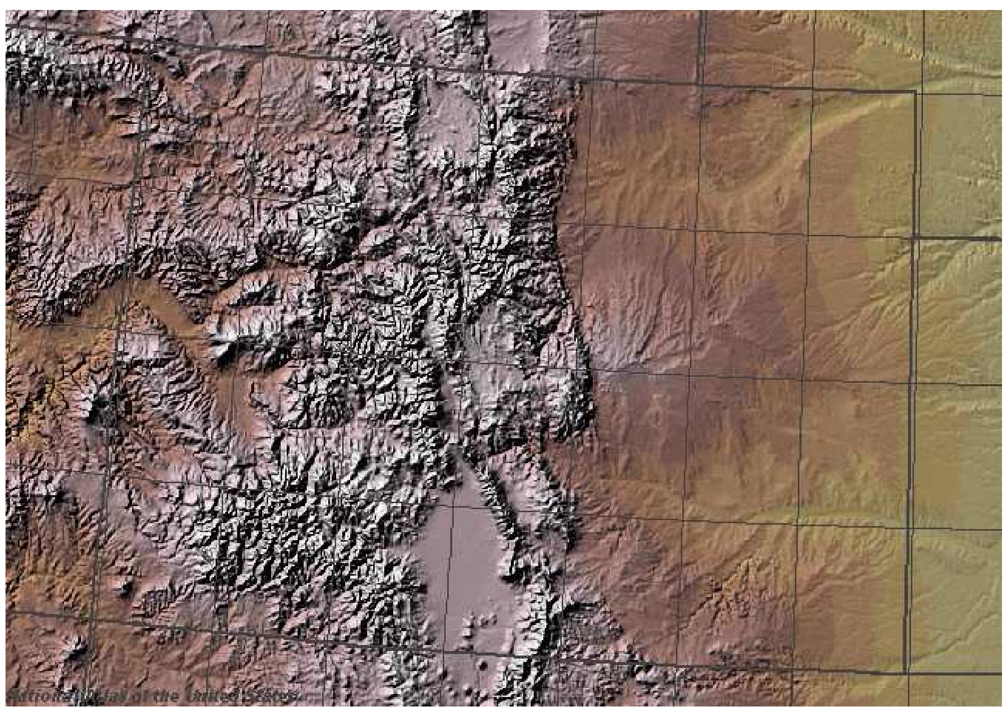
  Positive definite

- MCMC sampling of all parameters including those determining $\boldsymbol{\Sigma}_{(\cdot)}$

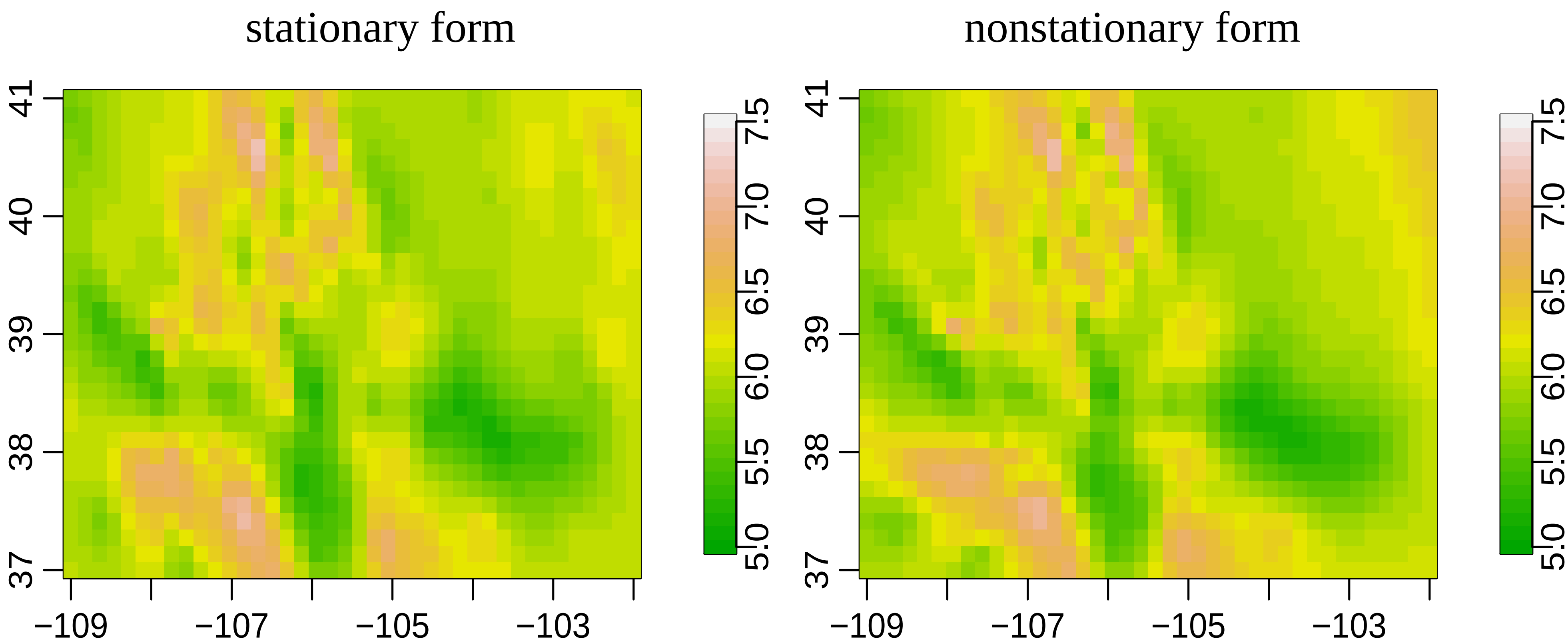
EMPIRICAL ASSESSMENT - SIMULATED FUNCTIONS



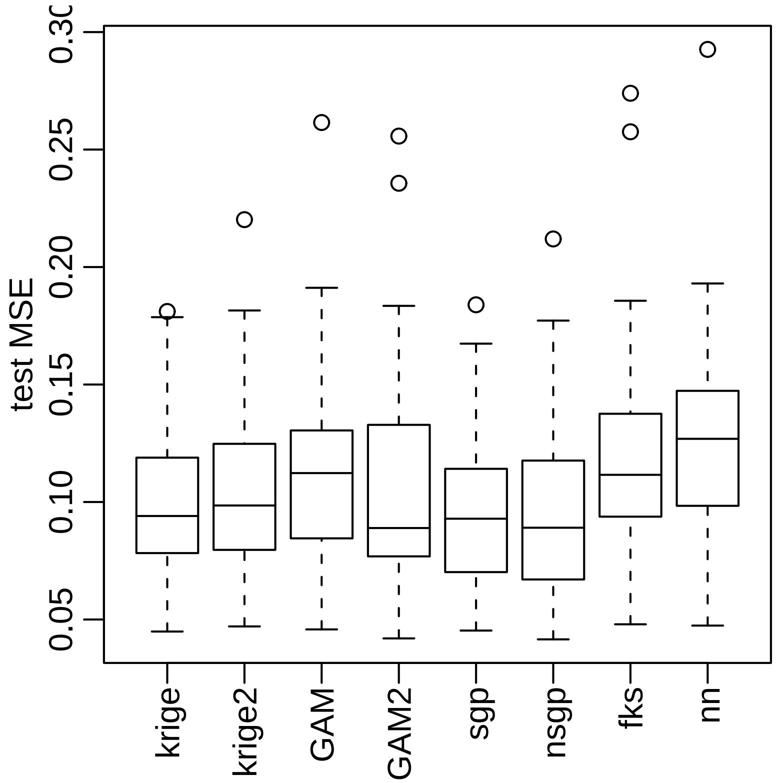

41
40
39
38
37
−109
−108
−107
−106
−105
−104
−103
−102



Colorado precipitation surfaces



Empirical Assessment - Colorado precipitation



- Comments

  surface is complex relative to number of obs; high df:n ratio

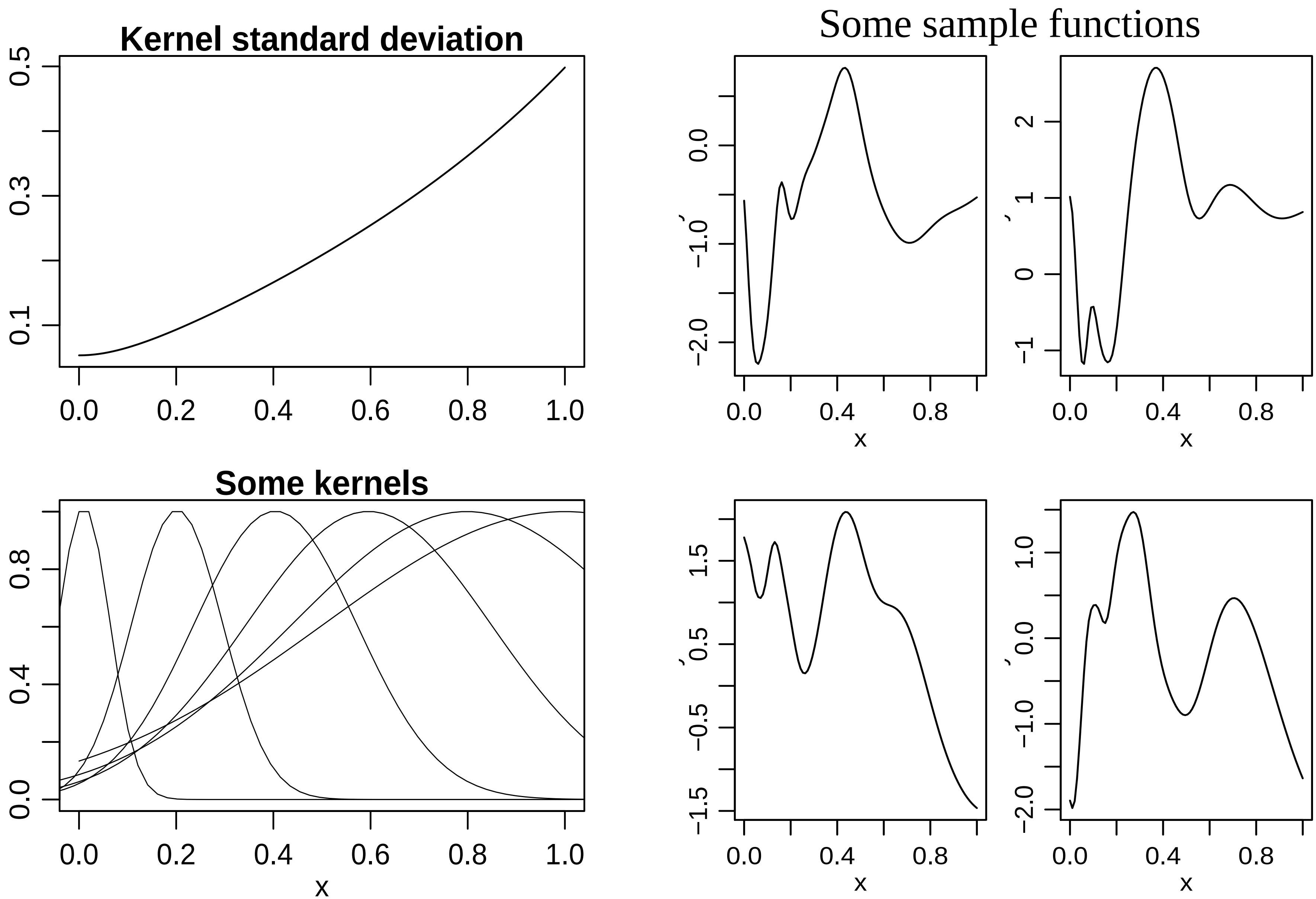
  bias-variance tradeoff

  simple two-region nonstationary models offer little improvement

Future Work

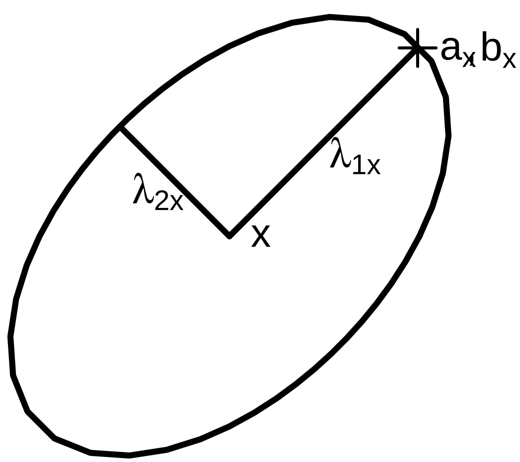
- Further implementation of nonstationary model:
  - ❖ Approaches for constraining hyperparameters for better fits and mixing
  - ❖ Computationally tractable approaches for multiple realizations
  - ❖ Ad hoc approaches for fitting the nonstationary covariance structure
- criteria for choosing a spatial smoother based on data spacing, df:n, signal to noise ratio
- Fourier basis representation for spatial processes for fast Bayesian estimation with non-normal data or complicated models

NONSTATIONARY GPS IN 1D



SMOOTHLY-VARYING KERNEL MATRICES

- Spectral decomposition ($\Re^P$): $\boldsymbol{\Sigma_x = \Gamma_x^T \Lambda_x \Gamma_x}$
  - ❖ in $\Re^{\mathbf{2}}$, stationary GP priors on unnormalized eigenvector coordinates $(\boldsymbol{a_x}, \boldsymbol{b_x})$ and on logarithm of second eigenvalue $(\boldsymbol{\lambda_{x,2}})$
  - ❖ efficient parameterizations of $\boldsymbol{\Phi(\cdot) \in \{a(\cdot), b(\cdot), \lambda_2(\cdot)\}}$ using basis function approximation to a stationary GP



Degree of Smoothing