
Nonparametric Regression Using Nonstationary Gaussian Processes

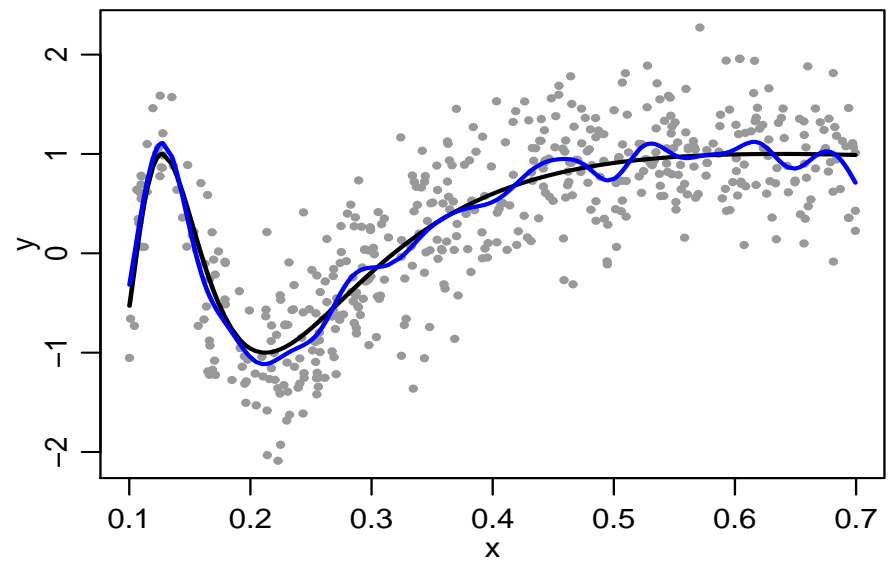
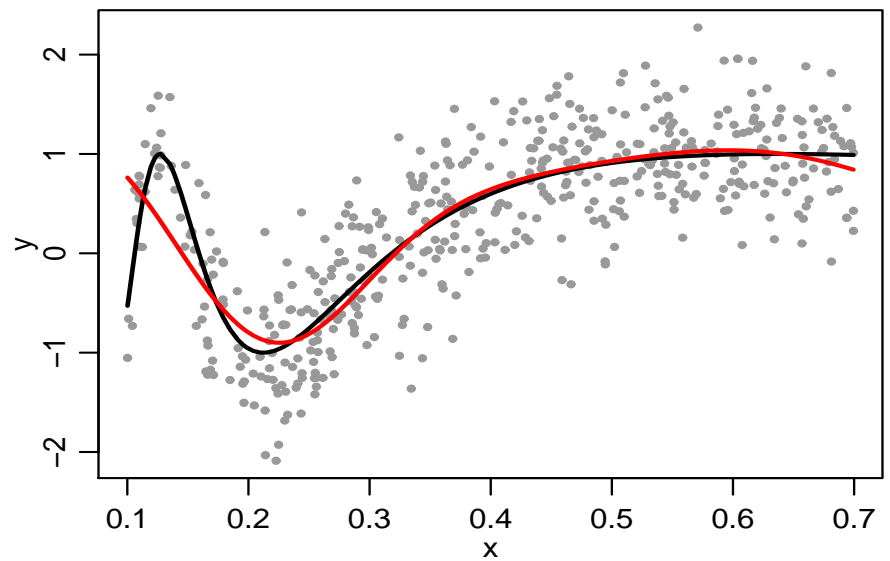
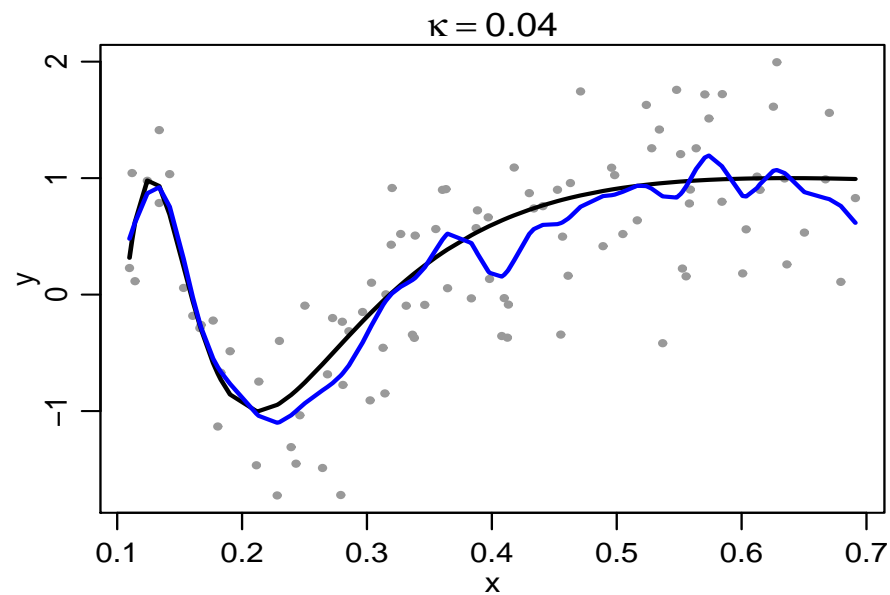
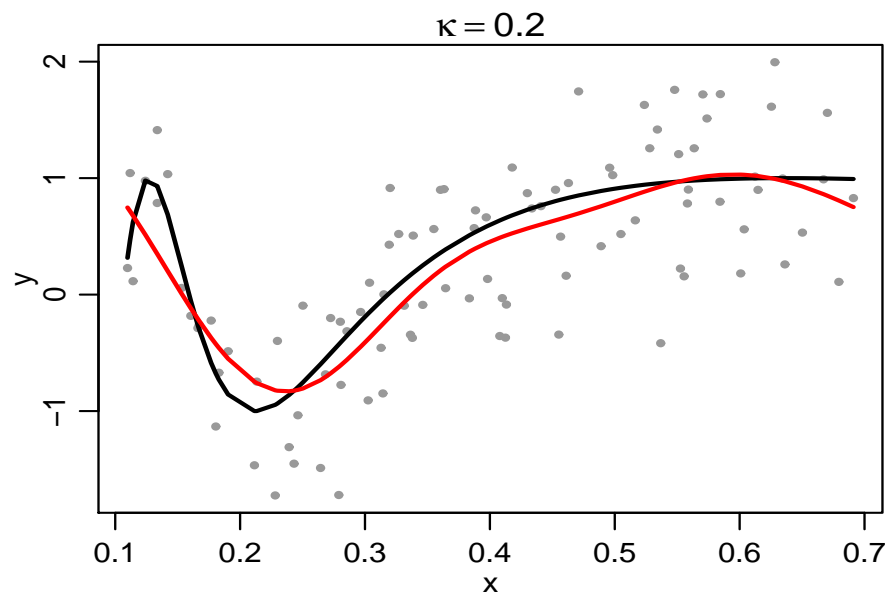
Chris Paciorek

*(Department of Biostatistics, Harvard University
and Department of Statistics, Carnegie Mellon University)*

Outline

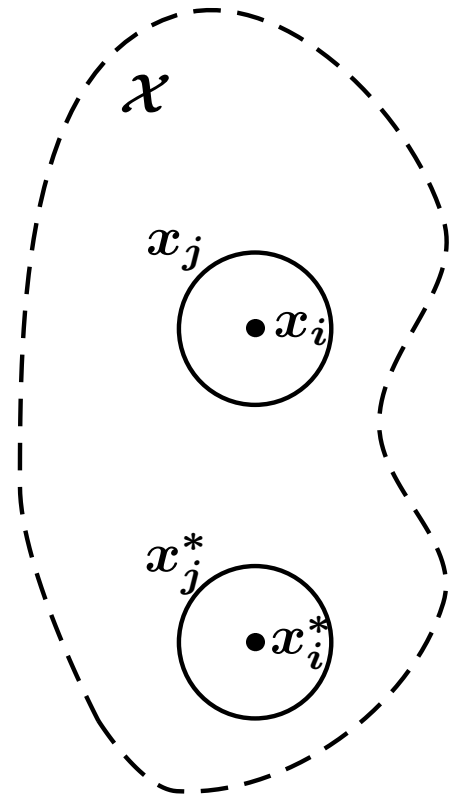
- Spatial adaptation in nonparametric regression
- Gaussian process (GP) modelling
- Nonstationary covariance model
- Empirical comparison with other adaptive smoothing methods

DEGREE OF SMOOTHING



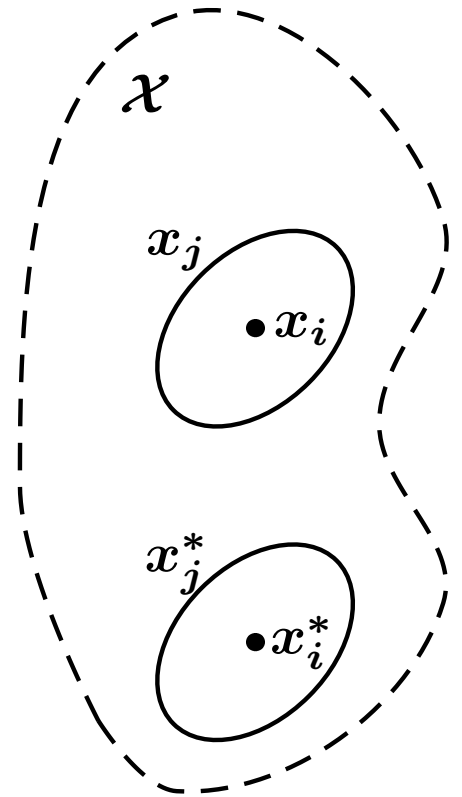
GAUSSIAN PROCESS DISTRIBUTION

- Infinite-dimensional joint distribution for $Z(x)$, $x \in \mathcal{X}$:
 - ❖ Example: $Z(\cdot)$ a regression function, $\mathcal{X} = \mathbb{R}^P$
 - ❖ $Z(\cdot) \sim \mathbf{GP}(\mu(\cdot), C(\cdot, \cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $C(x_i, x_j)$:
 - ❖ stationary, isotropic
 - ❖ stationary, anisotropic
 - ❖ nonstationary



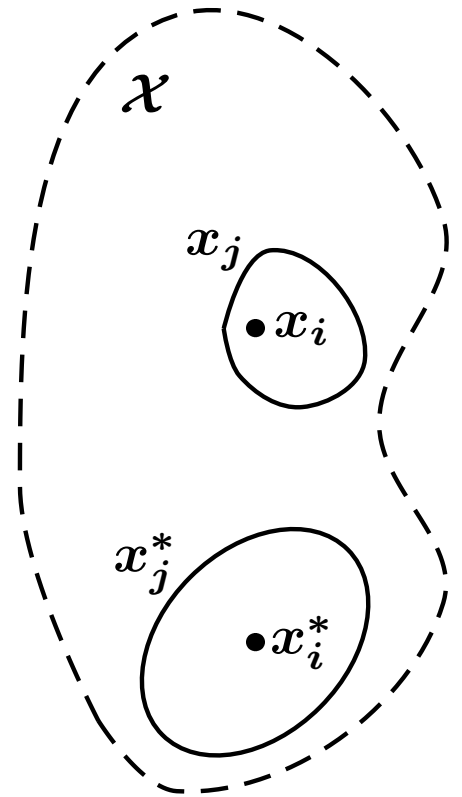
GAUSSIAN PROCESS DISTRIBUTION

- Infinite-dimensional joint distribution for $Z(x)$, $x \in \mathcal{X}$:
 - ❖ Example: $Z(\cdot)$ a regression function, $\mathcal{X} = \mathbb{R}^P$
 - ❖ $Z(\cdot) \sim \mathbf{GP}(\mu(\cdot), C(\cdot, \cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $C(x_i, x_j)$:
 - ❖ stationary, isotropic
 - ❖ stationary, anisotropic
 - ❖ nonstationary



GAUSSIAN PROCESS DISTRIBUTION

- Infinite-dimensional joint distribution for $Z(x)$, $x \in \mathcal{X}$:
 - ❖ Example: $Z(\cdot)$ a regression function, $\mathcal{X} = \mathbb{R}^P$
 - ❖ $Z(\cdot) \sim \mathbf{GP}(\mu(\cdot), C(\cdot, \cdot))$
- Finite-dimensional marginals are normal
- Types of covariance functions, $C(x_i, x_j)$:
 - ❖ stationary, isotropic
 - ❖ stationary, anisotropic
 - ❖ nonstationary



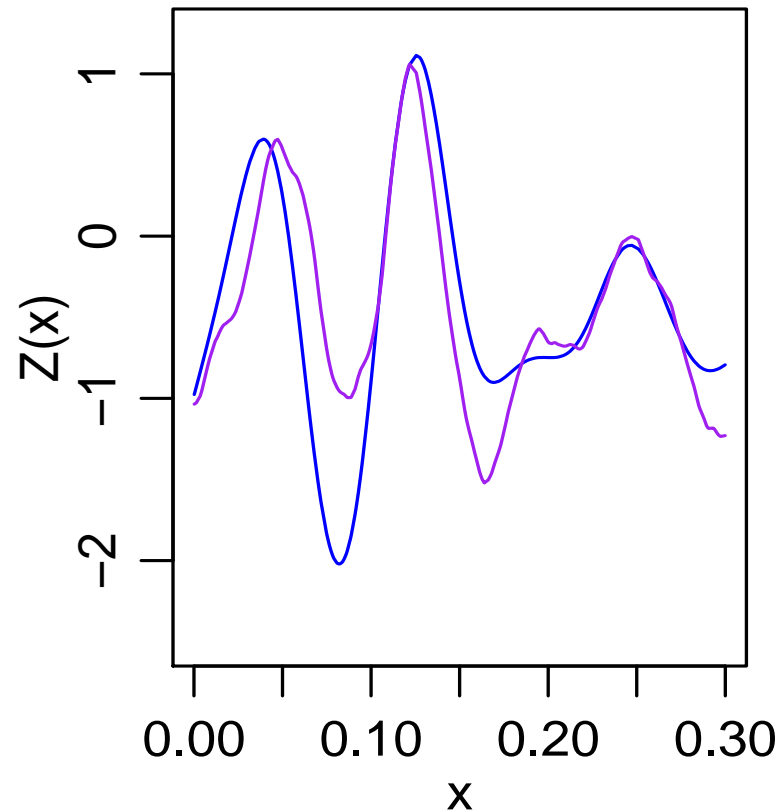
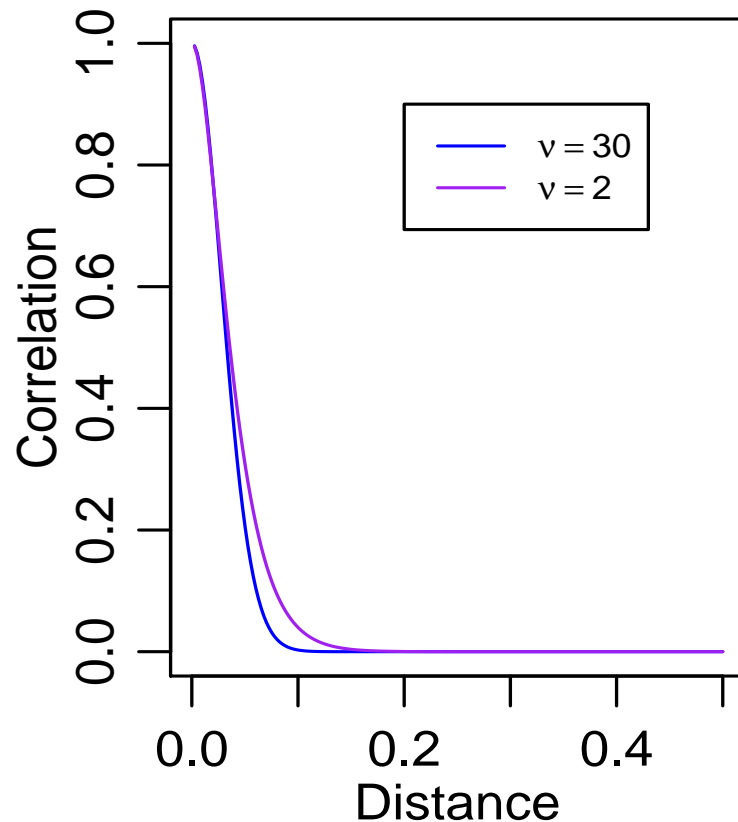
NONPARAMETRIC REGRESSION MODEL

- covariates $x_i \in \mathfrak{R}^P$
- Bayesian model

$$\begin{aligned} Y_i &\sim N(f(x_i), \eta^2), \\ f(\cdot) &\sim \text{GP}(\mu, \sigma^2 R(\cdot, \cdot; \nu, \theta)) \end{aligned}$$

STATIONARY CORRELATION FUNCTIONS

$$\text{Matérn form: } R(\tau) = \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{2\sqrt{\nu}\tau}{\kappa} \right)^{\nu} K_{\nu} \left(\frac{2\sqrt{\nu}\tau}{\kappa} \right)$$



- Differentiability controlled by ν , asymptotic advantages (Stein)

GENERALIZED NONSTATIONARY CORRELATION

- Generalizing Higdon, Swall, and Kern (1999)

$$R^{NS}(x_i, x_j) = \frac{|\Sigma_{x_i}|^{\frac{1}{4}} |\Sigma_{x_j}|^{\frac{1}{4}}}{\left| \frac{\Sigma_{x_i} + \Sigma_{x_j}}{2} \right|^{\frac{1}{2}}} R(\sqrt{Q_{ij}}),$$

where

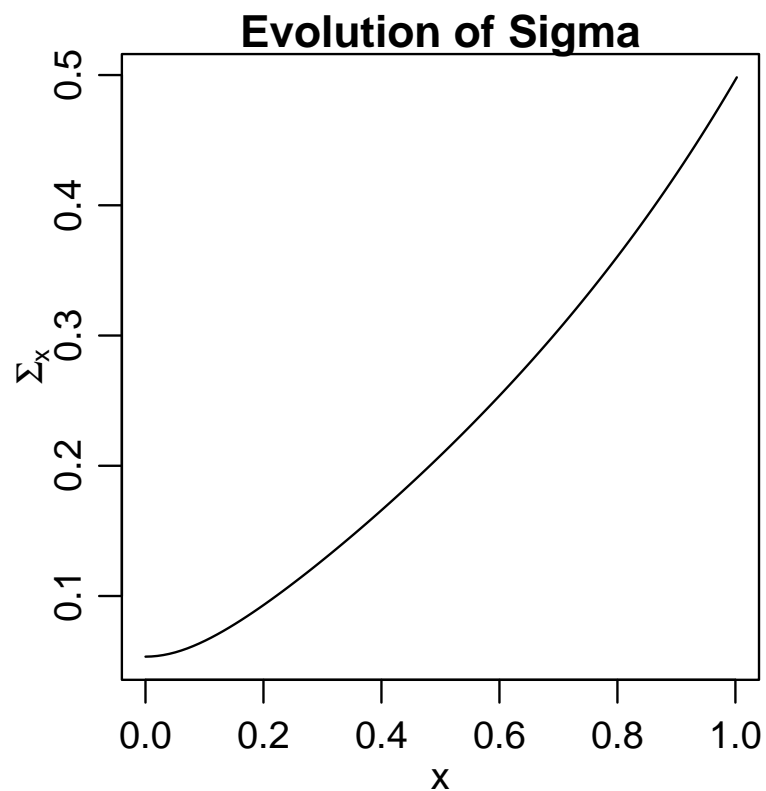
$$R(\sqrt{Q_{ij}}) = \frac{1}{\Gamma(\nu)2^{\nu-1}} (2\sqrt{\nu Q_{ij}})^{\nu} K_{\nu}(2\sqrt{\nu Q_{ij}})$$

and

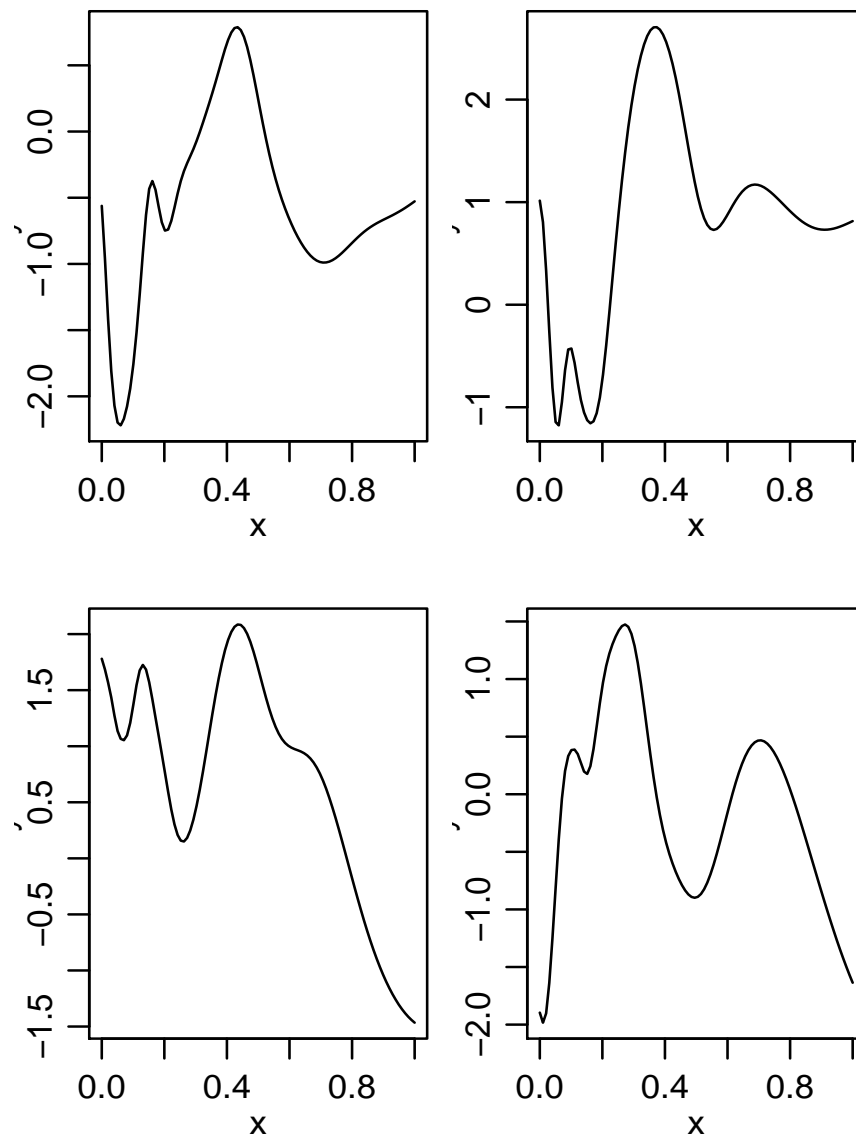
$$Q_{ij} = (x_i - x_j)^T \left(\frac{\Sigma_{x_i} + \Sigma_{x_j}}{2} \right)^{-1} (x_i - x_j)$$

- ❖ Based on Matérn stationary correlation function
- ❖ Differentiability of sample processes will depend on ν , provided kernel matrices vary sufficiently smoothly in covariate space
- ❖ Evolution of 'Mahalanobis' matrices, Σ_x , in space determines nonstationarity

NONSTATIONARY GPs IN 1D



Some sample functions



BAYESIAN SPECIFICATION FOR THE NONSTATIONARITY

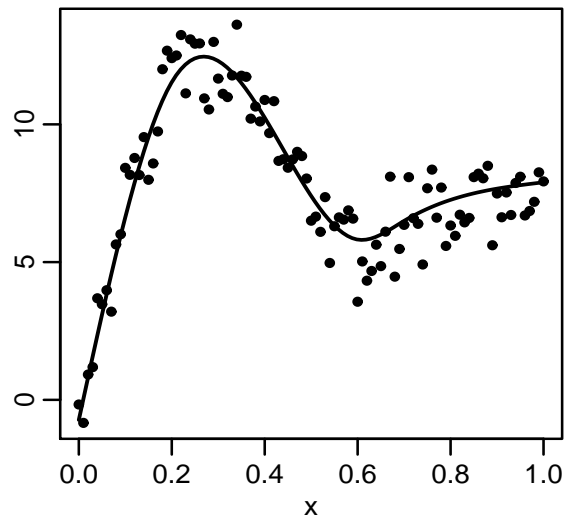
- Spatial field of matrices, Σ_x , needs to be modelled
- Approach:
 - ❖ Use a set of GP priors so that Σ_x vary smoothly in covariate space
 - ❖ Constrain so that Σ_x is positive definite for all x
 - ❖ Estimate Σ_x in the MCMC
- Unwieldy and highly-parameterized with many covariates

NONPARAMETRIC REGRESSION COMPARISON

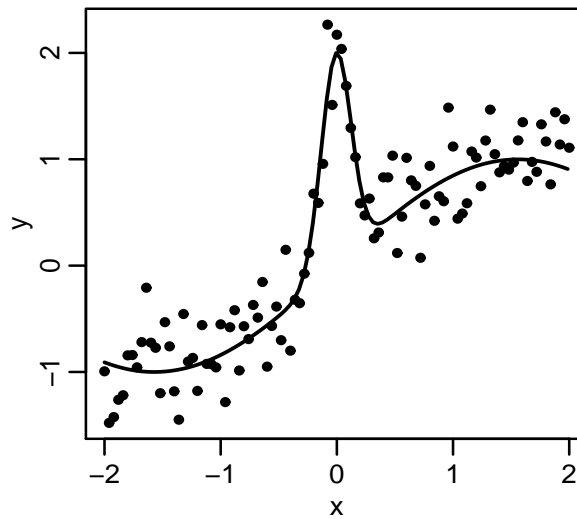
- Compare performance to Bayesian models in which f is a spline
 - ❖ $x_i \in \mathbb{R}^1$: BARS (DiMatteo, Genovese & Kass 2002)
 - ❖ $x_i \in \mathbb{R}^P, P > 1$:
 - ❖ BMARS (Denison, Mallick & Smith 1998) - tensor products of univariate splines
 - ❖ BMLS (Holmes & Mallick 2001) - multivariate linear splines (continuous piecewise planes)

REGRESSION RESULTS - 1D

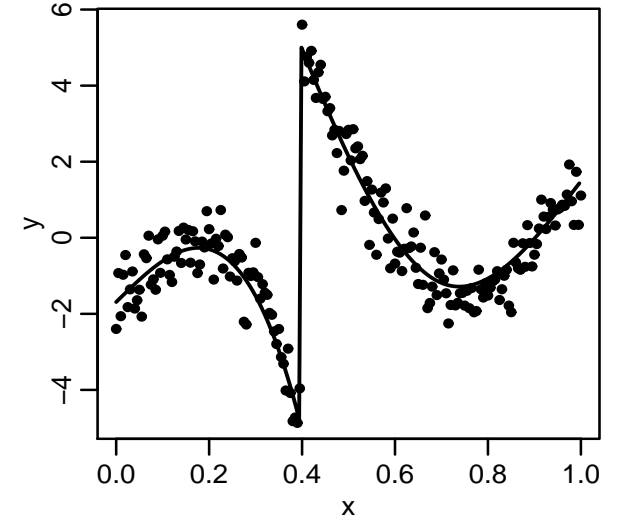
test function 1



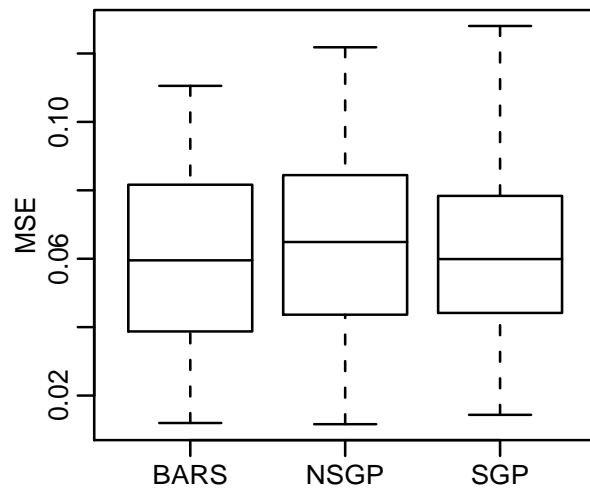
test function 2



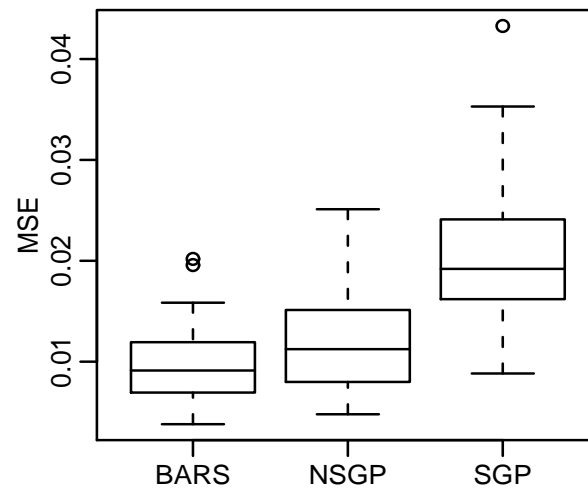
test function 3



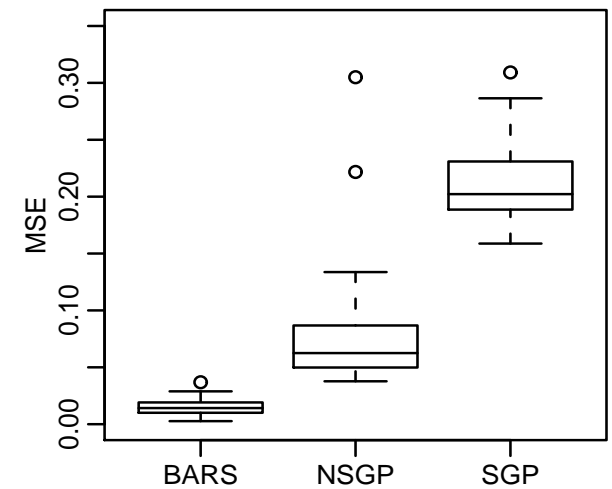
MSE



MSE

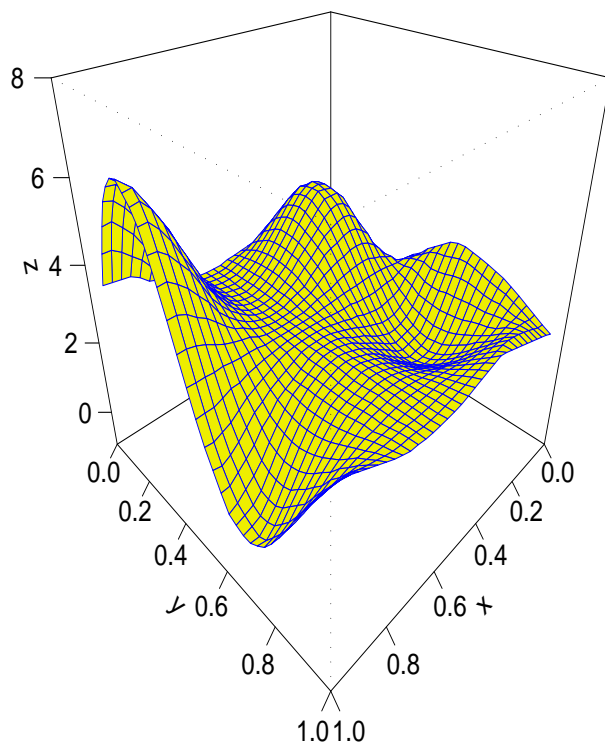


MSE

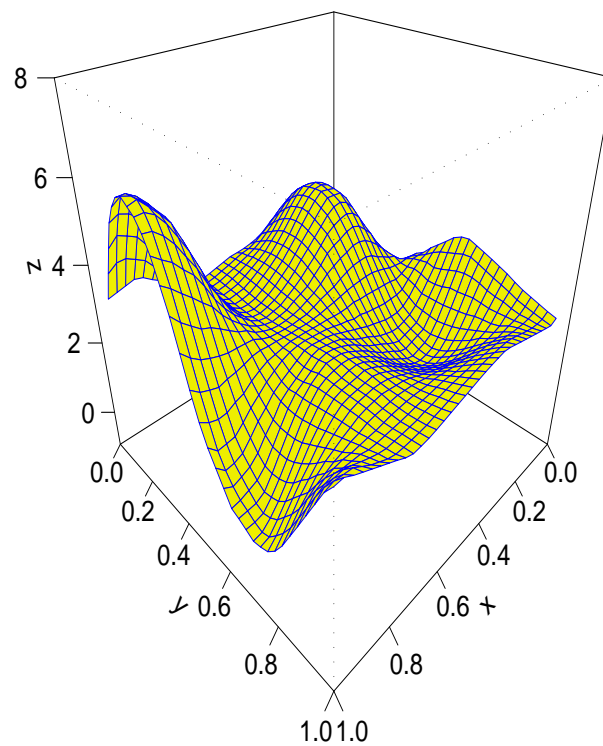


REGRESSION RESULTS - 2D

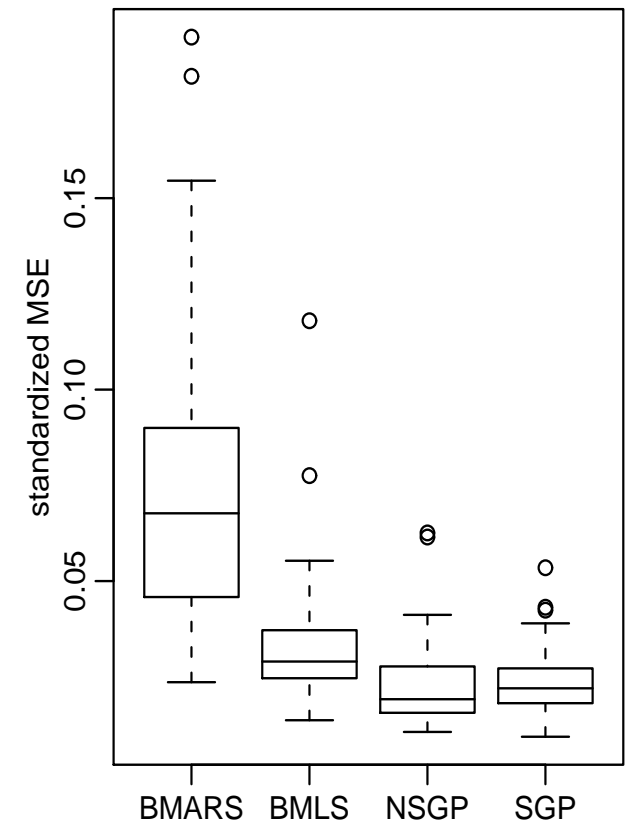
True function



NSGP estimate



standardized MSE



test function: $P = 2, n = 225$

REGRESSION RESULTS - REAL DATA

- Dec. 1993 mean temperatures in Americas, $n = 109$

$P = 2$: longitude, latitude

- daily ozone in NY, $n = 111$

$P = 3$: radiation, temperature, wind speed

- cross-validated MSE

model	temperature	ozone
Lin Regr	—	0.021
GAM	—	0.020
BMARS	1.74	0.0062
BMLS	2.40	0.0062
SGP	1.40	0.0062
NSGP	1.10	0.0054

COMMENTS AND CONCLUSIONS

- Nonstationary GP holds promise for spatially adaptive multivariate regression modelling
- Methodology extends readily to non-Gaussian data (e.g., binomial, Poisson)
- Fitting and mixing are slow using MCMC - further research needed
- Parameterizing and fitting the Σ_x field are complicated with more than 2-3 covariates