# Qualitative and Quantitative Evaluation of Covariance Tapering in Large Kriging Problems

## Tapering and Kriging

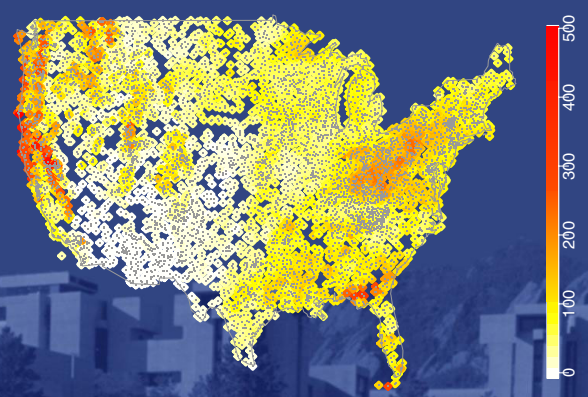
Introduce a sparseness structure in the covariance via tapering to gain computational advantages in large kriging problems constraint to maintaining asymptotic optimality.

In collaboration with Marc Genton and Doug Nychka.



## Motivation

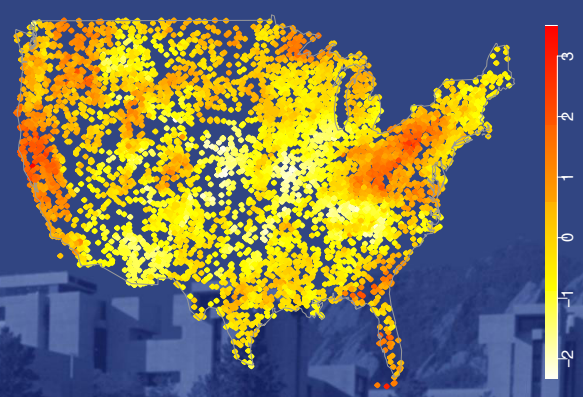
Monthly aggregated precipitation in April 1948






## Motivation

Precipitation anomaly in April 1948






## Best Linear Unbiased Predictor

Suppose a spatial process of the form

$$Z(\mathbf{x}) = \mathbf{m}(\mathbf{x})^{\mathsf{T}}\beta + \varepsilon(\mathbf{x}), \quad \mathrm{E}(\varepsilon) = \mathbf{0}, \quad \mathrm{Cov}(\varepsilon) = \mathbf{C}$$

with observations $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)^{\mathsf{T}}$, $\mathbf{x}_i \in \mathcal{D} \subset \mathbb{R}^d$.

The kriging estimator (BLUP) is

$$\hat{Z}(\mathbf{x}_0) = \lambda^{\mathsf{T}}\mathbf{Z}$$

where

$$\lambda = \mathbf{C}^{-1}\big(\mathbf{I} - \mathbf{M}(\mathbf{M}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{M})^{-1}\big(\mathbf{M}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{c} - \mathbf{m}(\mathbf{x}_0)\big)\big)$$

with $\mathbf{c}_i = \mathrm{Cov}\big(Z(\mathbf{x}_0), Z(\mathbf{x}_i)\big)$, $\mathbf{M} = (\mathbf{m}(\mathbf{x}_1), \ldots, \mathbf{m}(\mathbf{x}_n)^{\mathsf{T}}$.



## Best Linear Unbiased Predictor

Suppose a spatial process of the form

$$Z(\mathbf{x}) = \mathbf{m}(\mathbf{x})^{\mathsf{T}}\beta + \varepsilon(\mathbf{x}), \quad \mathrm{E}(\varepsilon) = \mathbf{0}, \quad \mathrm{Cov}(\varepsilon) = \mathbf{C}$$

with observations $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)^{\mathsf{T}}$, $\mathbf{x}_i \in \mathcal{D} \subset \mathbb{R}^d$.

The kriging estimator (BLUP) is

$$\hat{Z}(\mathbf{x}_0) = \lambda^{\mathsf{T}}\mathbf{Z}$$

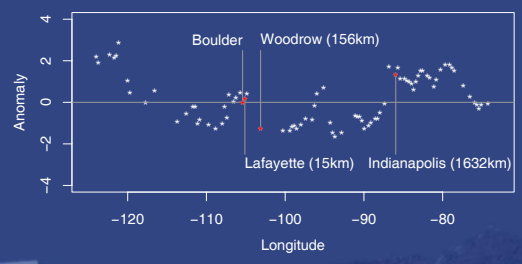
where

$$\lambda = \mathbf{C}^{-1}\big(\mathbf{I} - \mathbf{M}(\mathbf{M}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{M})^{-1}\big(\mathbf{M}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{c} - \mathbf{m}(\mathbf{x}_0)\big)\big) = \mathbf{C}^{-1}\mathbf{c}$$

with $\mathbf{c}_i = \mathrm{Cov}\big(Z(\mathbf{x}_0), Z(\mathbf{x}_i)\big)$, and $\mathbf{M} = (\mathbf{m}(\mathbf{x}_1), \ldots, \mathbf{m}(\mathbf{x}_n)^{\mathsf{T}}$.

## Motivation

Precipitation anomaly along 40° latitude






## Matérn Covariogram 1 2

We need a broad, flexible class of covariograms to describe spatial processes.

We recommend the Matérn class given by

$$c(\mathbf{h}) \propto \left(\alpha\|\mathbf{h}\|\right)^{\nu}\mathcal{K}_{\nu}\left(\alpha\|\mathbf{h}\|\right)$$

and with spectral density

$$f(\omega) \propto \frac{1}{\left(\alpha^2 + \|\omega\|^2\right)^{\nu+d/2}}$$

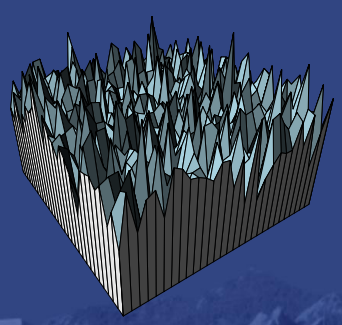
Differentiability at the origin of the covariogram is related to the tail behavior of the spectrum, i.e. the smoothness of the process.

The process is $m$ times mean squared differentiable iff $m < \nu$.



## Matérn Covariogram 1 2

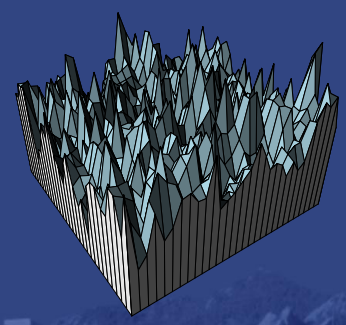
$\nu = 0.0$ (white noise)
effective range is 0.2





## Matérn Covariogram 1 2

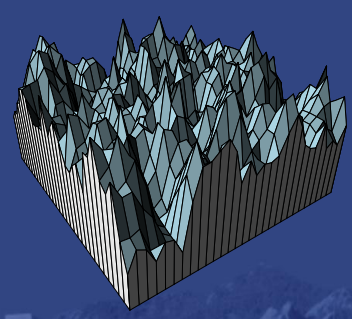
$\nu = 0.5$
effective range is 0.2





## Matérn Covariogram 1 2

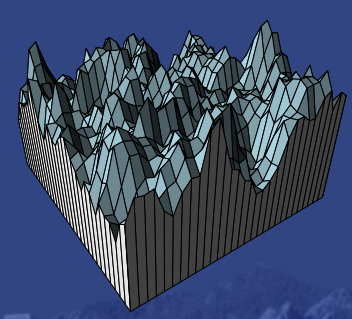
$\nu = 1.0$
effective range is 0.2





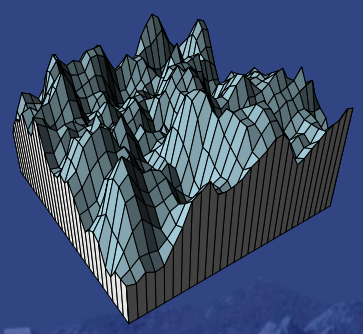
## Matérn Covariogram 1 2

$\nu = 1.5$
effective range is 0.2



## Matérn Covariogram 1 2

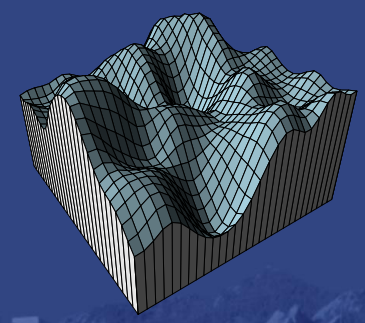
$\nu = 2.0$
effective range is 0.2





## Matérn Covariogram 1 2

$\nu = 5.0$
effective range is 0.2





## Misspecified Covariances

In a series of (Annals) papers, Stein gives asymptotic results for misspecified covariances.

Suppose the true spectrum $f_0$. If we krig with the misspecified covariance characterized by $f_1$ then under the conditions

$$\frac{f_1(\boldsymbol{\omega})}{f_0(\boldsymbol{\omega})} = 1 \text{ as } \|\boldsymbol{\omega}\| \to \infty \qquad \text{and } \ldots$$

we have

$$\frac{\text{MSE pseudo-BLUP}(f_1)}{\text{MSE BLUP}(f_0)} \to 1$$

$$\frac{\text{presumed-MSE pseudo-BLUP}(f_1)}{\text{MSE BLUP}(f_0)} \to 1$$



## Tapered Covariances

Tapering, i.e. multiplying a covariance function $c(\cdot)$ with $k(\cdot)$, is a form of misspecification if

$$\frac{\mathcal{F}\Big(c(\|\mathbf{h}\|)k(\|\mathbf{h}\|)\Big)}{\mathcal{F}\Big(c(\|\mathbf{h}\|)\Big)} \to 1 \text{ as } \|\boldsymbol{\omega}\| \to \infty$$

Which taper satisfies this condition?

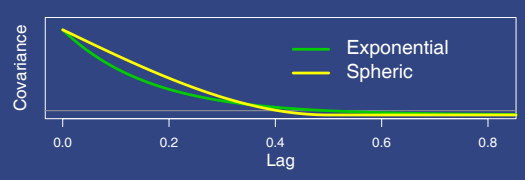
The answer lies in the Principal Irregular Term.

The taper

- has to be as differentiable at the origin as the original covariogram
- has to be more differentiable throughout the domain than at the origin
- has to be inflated (factor obtained via PIT)

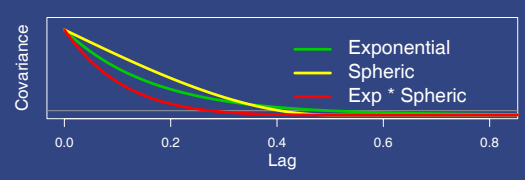


## Examples

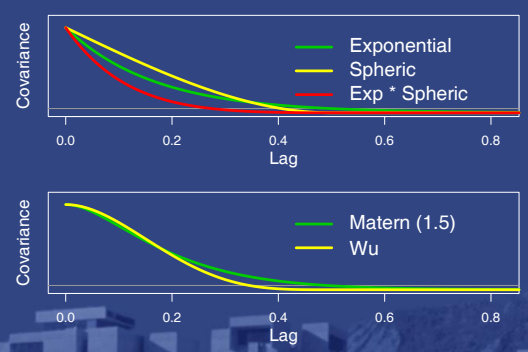





## Examples

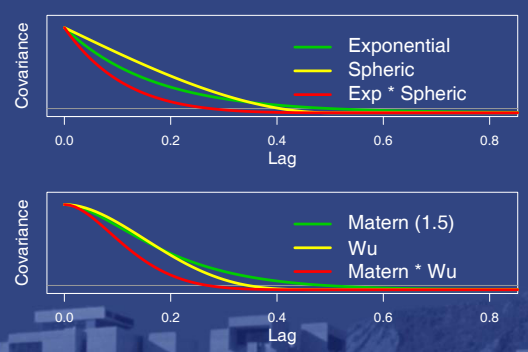


## Examples



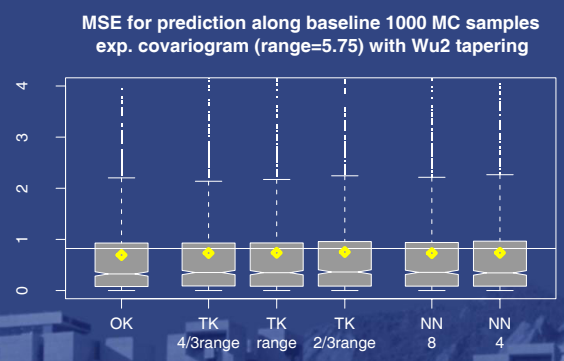



## Examples






## Numerical Study






## Recap

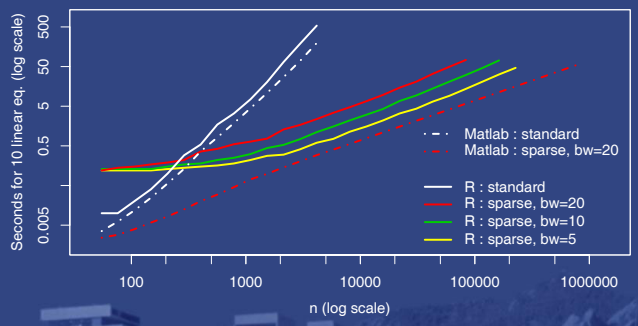
To summarize we have discussed

- the BLUP, requiring solutions of linear systems
  $\rightsquigarrow$ one point $\mathbf{c}^T\mathbf{C}^{-1}\mathbf{Z}$, field $\widetilde{\mathbf{C}}^T\mathbf{C}^{-1}\mathbf{Z}$
- optimal covariances, decay exponentially
  $\rightsquigarrow$ no zeros in $\mathbf{C}$, $\mathbf{c}$ or $\widetilde{\mathbf{C}}$
- that tapering preserves optimality,
  $\rightsquigarrow$ covariance matrices are sparse

Is there a computational gain?



## Computational Efficiency






## It would be coffee time...