NRCSE

Nonstationary spatial covariance modeling

My experience:

1. **Hourly ozone monitoring data at 100 sites** in the San Joaquin Valley, CA, for assessment of photochemical model predictions:

- Accounting for sub-grid scale spatial variability in point monitoring observations to assess grid average photochemical predictions
- Decomposing spatial difference fields—empirical grid cell estimates vs. model predictions—into components at different spatial scales
- Comparing empirical spatial covariance structure with model spatial covariance structure at different spatial scales.

2. **10-day aggregate precipitation** in Languedoc-Roussillon, France
3. **Daily ozone monitoring summaries** (max 8-hr ave) from 100's of monitoring sites across the country for spatial prediction at target locations.
4. Daily and 2-week average NO_x and $PM_{2.5}$ modeling for monitoring sites for MESA Air project.

Perspective: 1st and 2nd order stationarity is almost never a realistic assumption for any environmental monitoring data, except at small spatial scales.

Objectives for approaches to nonstationary spatial covariance modeling.

- **Characterizing spatially varying, locally (stationary) anisotropic structure.**
- **Scientific understanding/representation of covariance structure—not just a method of providing covariances for kriging.**

Capable of:

- reflecting effects of known explanatory environmental processes such as transport/wind, topography, point sources
- modeling effects of known explanatory environmental processes

Objectives (cont.)

- Application to purely spatial problems and/or problems with data sampled irregularly in space and time
- Application in context of dynamic models for space-time structure
- Application to “large” problems/data sets

❖ **Diagnostics for local and large-scale correlation structure:**

- is the spatial structure “right”
- is the nature/degree of nonstationarity (smoothness) right?

❖ **Evaluation of uncertainty in estimation (interpolation) of spatial covariance structure**

❖ **Incorporation in an approach to spatial estimation accounting for uncertainty in estimation of (parameters of) spatial covariance structure**

Selected Methods and References:

1. Basis function methods (Nychka, Wikle, …)

2. Kernel/smoothing methods (Fuentes, …)

3. Process convolution models (Higdon, Swall & Kern, …; Paciorek & Schervish)

4. Parametric models

5. Spatial deformation models

- Sampson & Guttorp, 1989, 1992, 1994
- Meiring, Monestiez, Sampson & Guttorp, 1997. "Developments..." *Geostatistics Wollongong '96*.
- Mardia & Goodall, 1993. in *Multivariate Environmental Statistics*.
- Smith, 1996. Estimating nonstationary spatial correlations. UNC preprint.

4. Spatial deformation models (cont.)

- Perrin & Meiring, 1999. Identifiability *J Appl Prob*.
- Perrin & Senoussi, 1998. Reducing nonstationary random fields to stationarity or isotropy, *Stat & Prob Letters*.
- Perrin & Monestiez, 1998. Parametric radial basis deformations. *GeoENV-II*.
- Iovleff & Perrin, 2004. Simulated annealing. *J Comp Graph Stat*
- Schmidt & O'Hagan, 2003. Bayesian inference for non-stationary spatial covariance structure via spatial deformations. *JRSS-B*
- Damian, Sampson & Guttorp, 2000. Bayesian estimation of semiparametric nonstationary covariance structures. *Environmetrics*.
- Damian, Sampson & Guttorp, 2003. Variance modeling for nonstationary spatial processes with temporal replications. *Journal of Geophysical Research*
- Related recent developments in the atmospheric science literature:

Review: **Descriptive characteristics of (stationary) spatial covariance expressed in a variogram**

The spherical correlation

$$\rho(\mathbf{v}) = \begin{cases} 1 - 1.5\mathbf{v} + 0.5\left(\frac{\mathbf{v}}{\phi}\right)^3; & \mathbf{h} < \phi \\ 0, \quad \text{otherwise} \end{cases}$$

Corresponding variogram

nugget → $\tau^2 + \frac{\sigma^2}{2}\left(3\frac{t}{\phi} - \left(\frac{t}{\phi}\right)^3\right); \quad 0 \le t \le \phi$

sill → $\tau^2 + \sigma^2; \quad t > \phi$ ← range

Consider two synthetic data sets; we will call them **A** and **B**. Some common descriptive statistics for these two data sets are given in Table 1.1.

	A	B
Count	15251	15251
Average	100.00	100.00
Standard Deviation	20.50	25.00
Median	100.35	106.92
10 Percentile	73.59	73.95
90 Percentile	125.61	124.72

Table 1.1 Some common descriptive statistics for the two example data sets.

The histograms for these two data sets are given in Figures 1.1 and 1.2. According to this evidence the two data sets are almost identical.

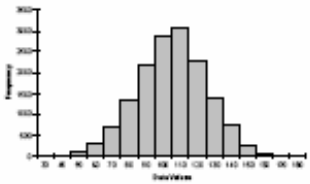


Figure 1.1 Data Set A Histogram

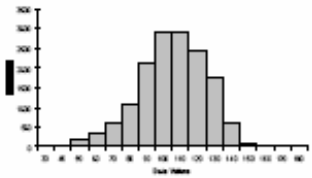


Figure 1.2 Data Set B Histogram

However, these two data sets are significantly different in ways that are not captured by the common descriptive statistics and histograms. As can be seen by comparing the associated contour plots (see Figures 1.3 and 1.4), data set **A** is rougher than data set B. Note that we can not say that data set **A** is "more variable" than data set B, since the standard deviations for the two data sets are the same, as are the magnitudes of highs and lows. The visually apparent difference between these two data sets is one of texture and not variability.

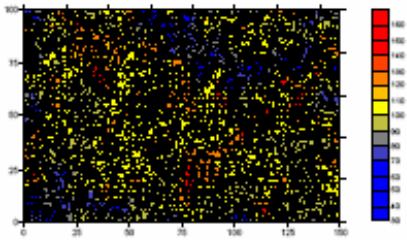


Figure 1.3 Data Set A Contour Plot

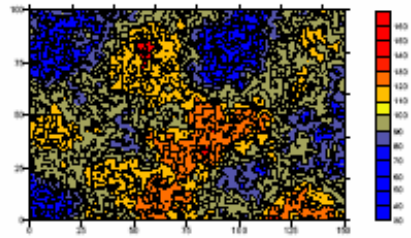


Figure 1.4 Data Set B Contour Plot

In particular, data set **A** changes more rapidly in space than does data set B. The continuous high zones (red patches) and continuous low zones (blue patches) are, on the average, smaller for data set **A** than for data set B. Such differences can have a significant impact on sample design, site characterization, and spatial prediction in general.

It is not surprising that the common descriptive statistics and the histograms fail to identify, let alone quantify, the textural difference between these two example data sets. Common descriptive statistics and histograms do not incorporate the spatial locations of data into their defining computations.

The variogram is a quantitative descriptive statistic that can be graphically represented in a manner which characterizes the spatial continuity (i.e. roughness) of a data set. The variograms for these two data sets are shown in Figures 1.5 and 1.6. The difference in the initial slope of the curves is apparent.

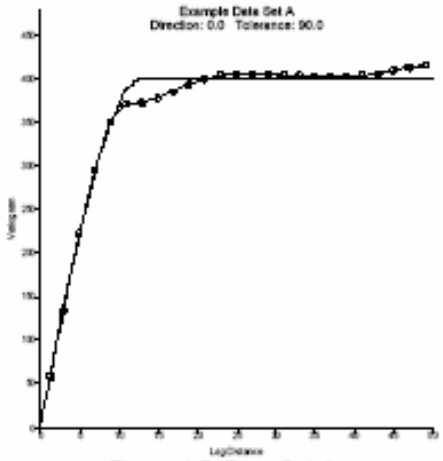


Figure 1.5 Data Set A Variogram and Model

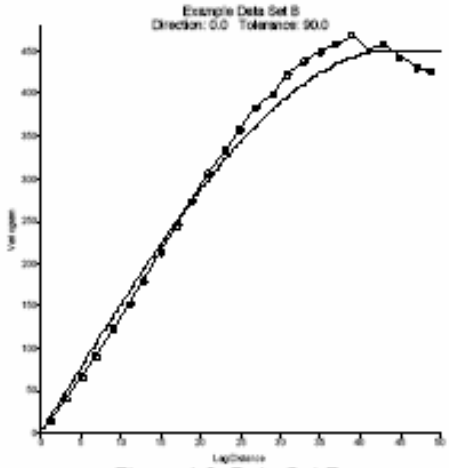
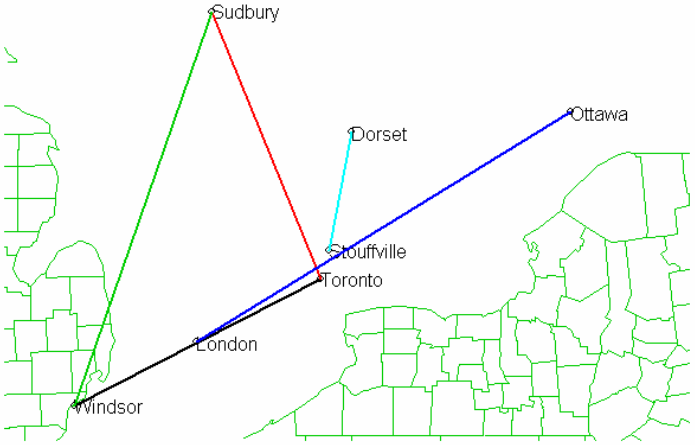


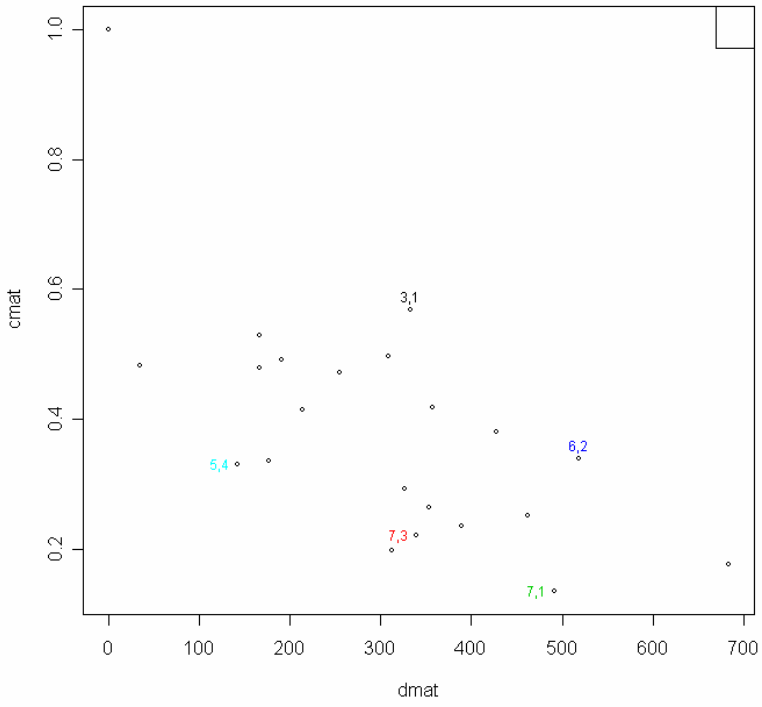
Figure 1.6 Data Set B Variogram and Model

Spatial continuity (roughness) of data set characterized by initial slope or range of variogram.

Correlation vs Distance for Ontario Ozone Data



Apparent <u>anisotropy</u>

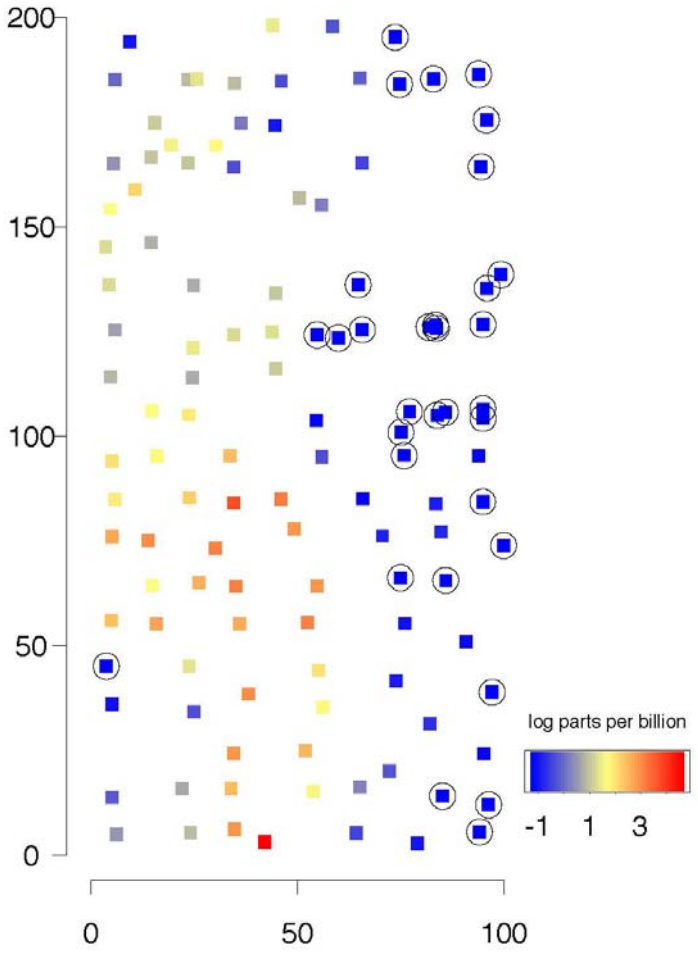


Nonstationary spatial covariance:

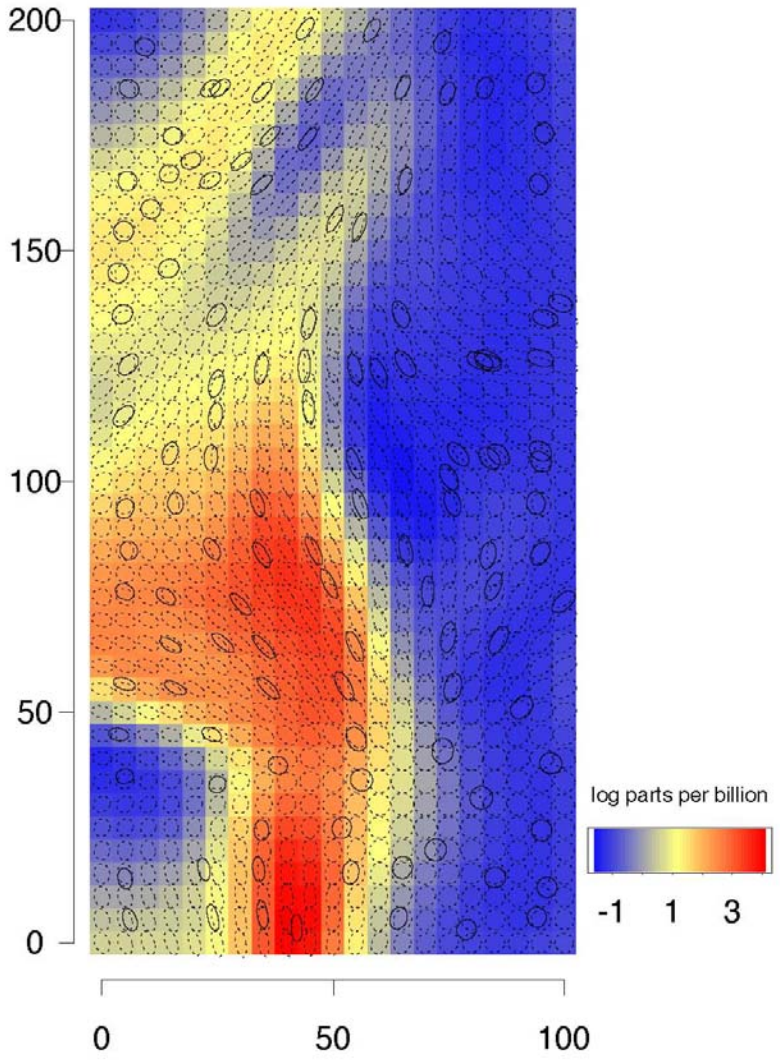
Basic idea: **the parameters of a local variogram model---nugget, range, sill, and anisotropy---vary spatially.**

Look at some pictures of applications from recent methodology publications.

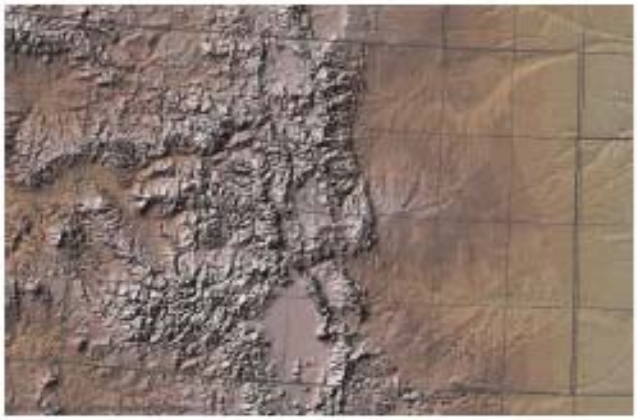
Swall & Higdon. Process convolution approach, Soil contamination example --- Piazza Rd site.

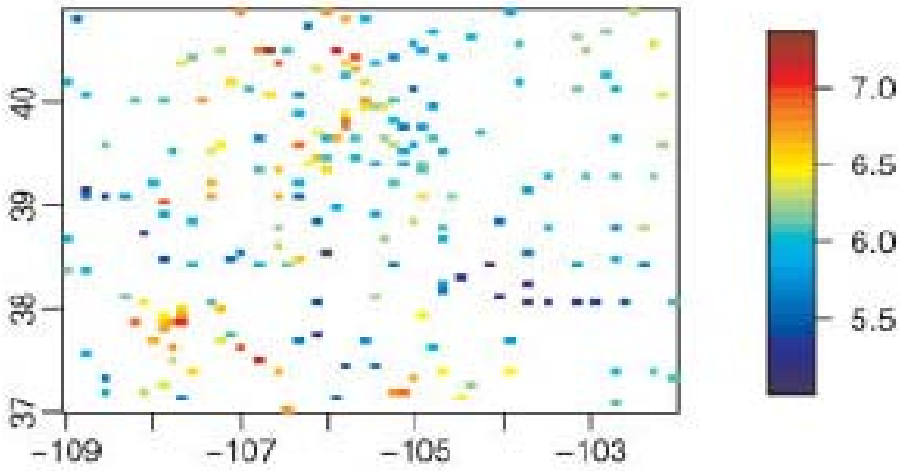


Swall & Higdon. Process convolution approach, Posterior mean and covariance kernel ellipses.



Paciorek & Schervish, 2006 – Colorado 1981 annual precip (log)





Paciorek & Schervish, 2006 – kernels (ellipses of constant Gaussian density) representing estimated correlation structure

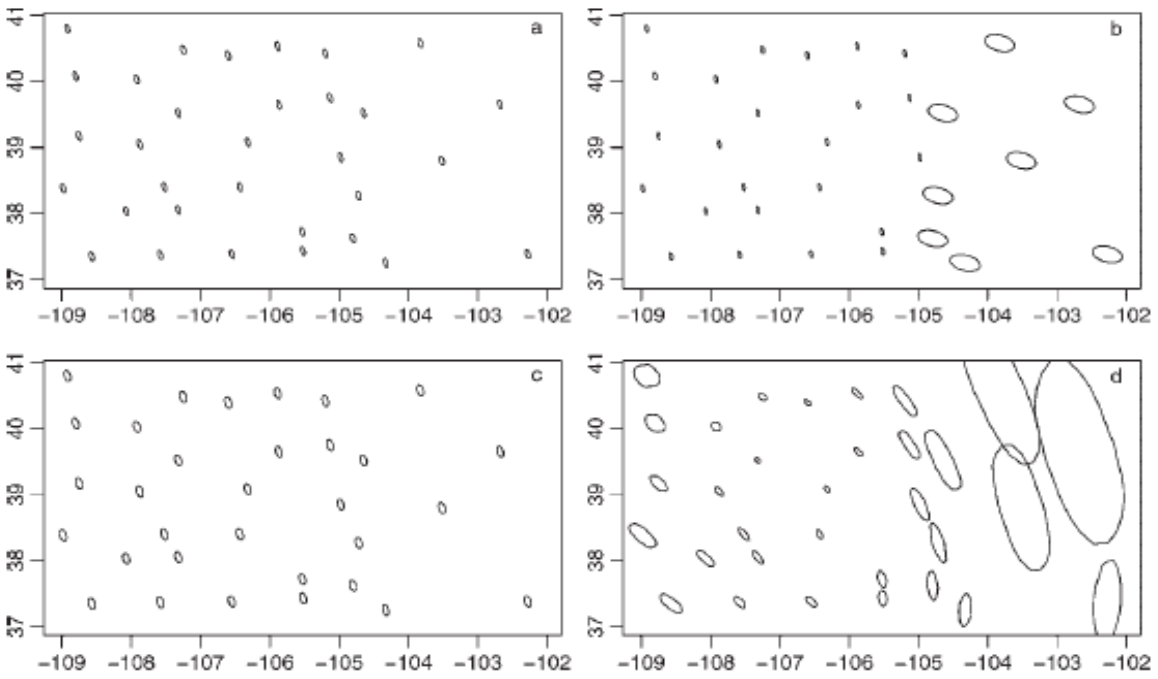


Figure 4. Kernels (ellipses of constant probability density of Gaussian densities) representing the estimated correlation structure for (a) stationary kriging, (b) nonstationary kriging based on two regions, (c) the fully Bayesian stationary GP model, and (d) the nonstationary GP model. For the Bayesian models, the ellipse-like figures are the posterior means of constant probability density ellipse values at a sequency of angles, $0, \ldots, 2\pi$

Pintore & Holmes, 2005. Spatially adaptive non-stationary covariance functions via spatially adaptive spectra.

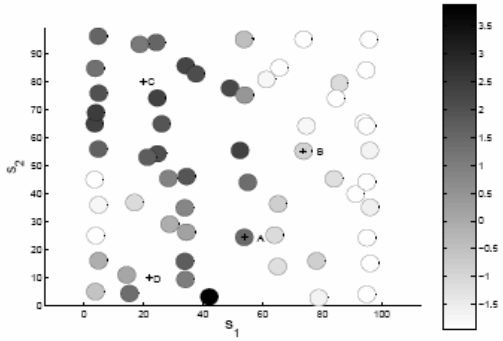


Figure 19: Plot 1: Bubble patch of the measurements of concentrations at the $n = 60$ data locations, Section 10.

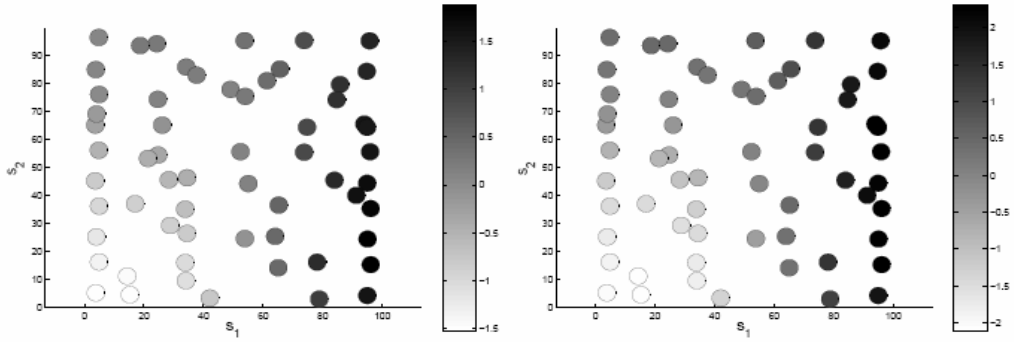


Figure 20: From the left, Plot 1: Bubble patch of $\log \eta_i$'s for the non-stationary model using the Karhunen-Loève expansion. Plot 2: Bubble patch of $\log \eta_i$'s for the non-stationary model using the Fourier representation., Section 10.

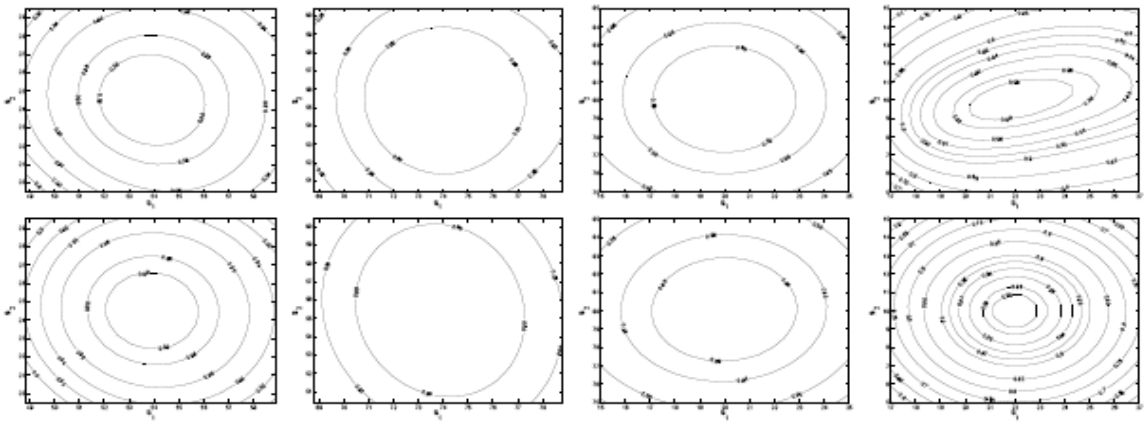
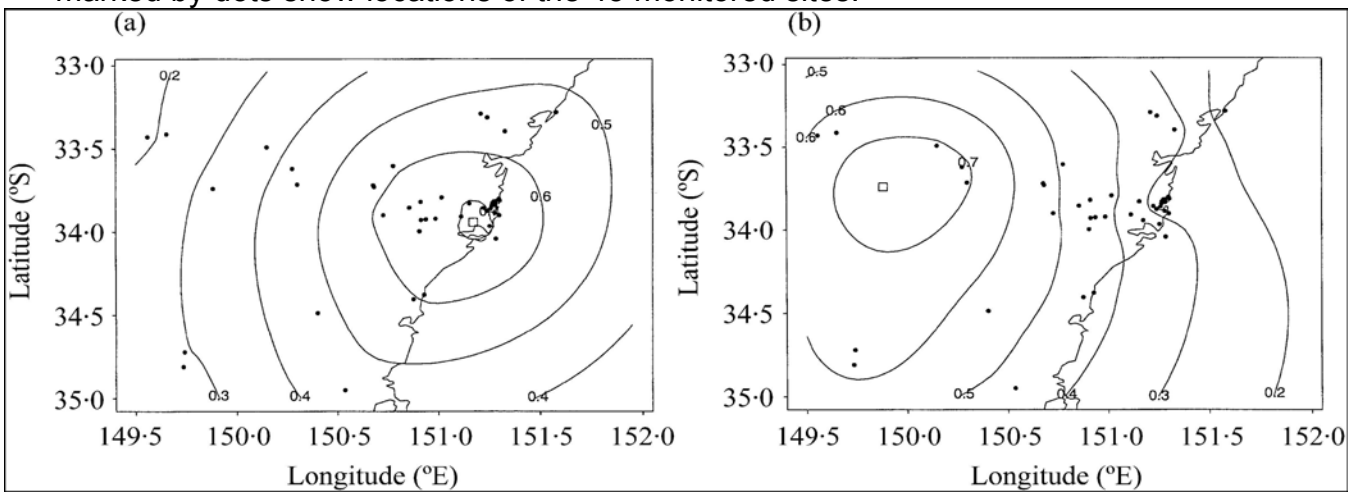
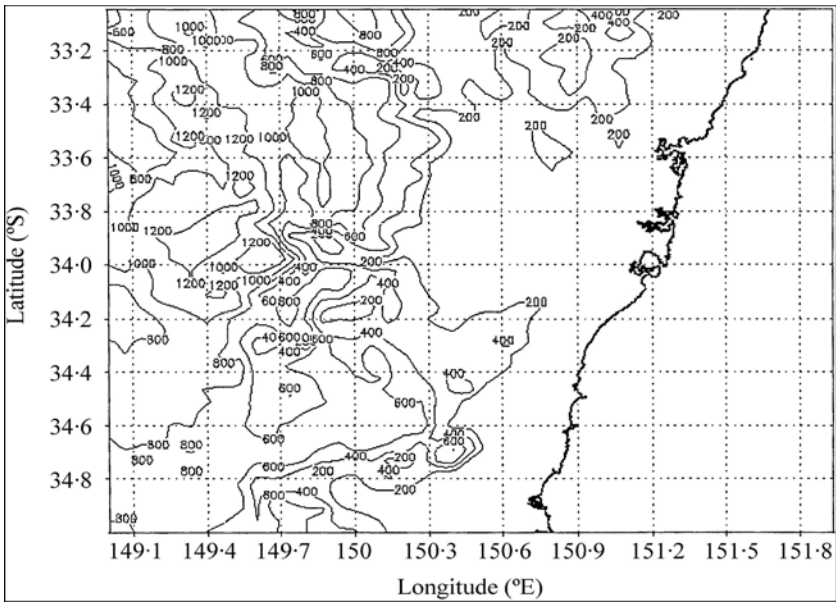


Figure 21: From the left, Top row plots: Correlation structure around points A, B, C and D using the Karhunen-Loève representation. From the left, Bottom row plots: Correlation structure around points A, B, C and D using the Fourier representation, Section 10.

Nott & Dunsmuire, 2002, Biometrika. Fig. 2. Sydney wind pattern data. Contours of equal estimated correlation with two different fixed sites, shown by open squares: (a) location 33·85°S, 151·22°E, and (b) location 33·74°S, 149·88°E. The sites marked by dots show locations of the 45 monitored sites.





Kim, Mallock & Holmes, JASA 2005. Piecewise Gaussian model for groundwater permeability data

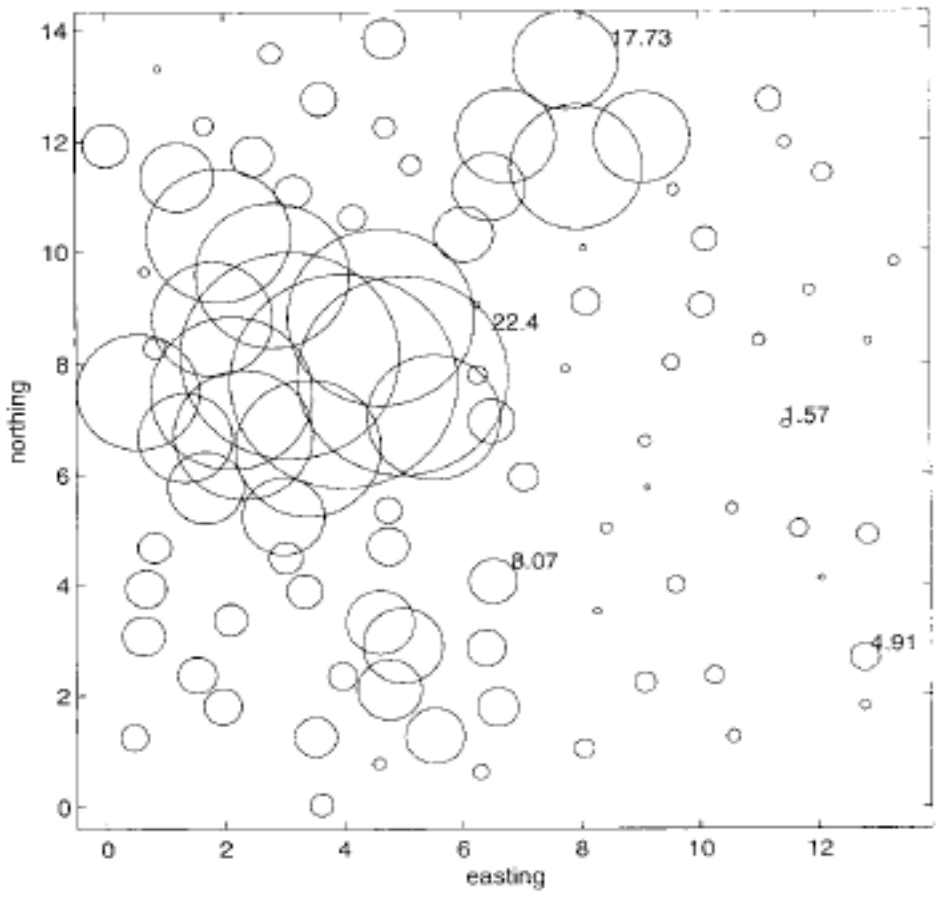


Figure 9. Bubble Plot of the Transformed Data.

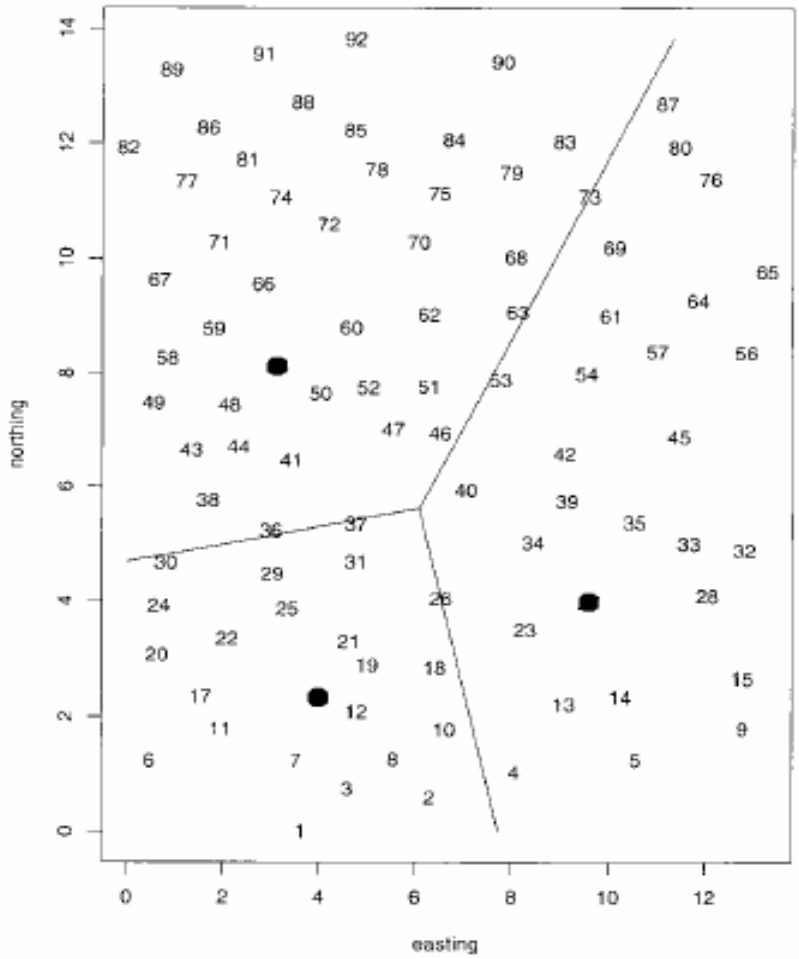


Figure 10. The Posterior Modal Sample of Voronoi Tessellation (with numbered locations).

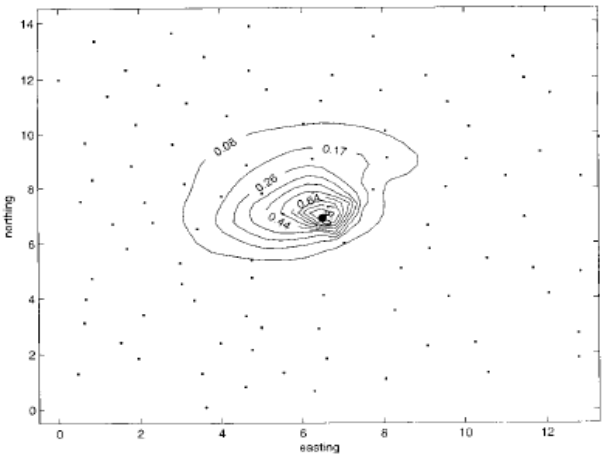


Figure 15. Marginal Correlation Surface for the 46th Location (marked with a bullet).

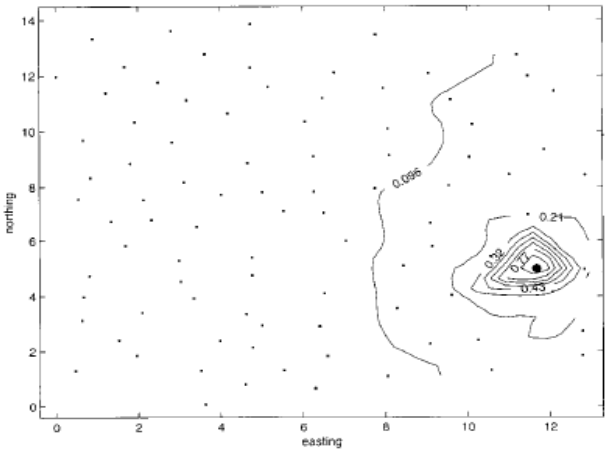


Figure 14. Marginal Correlation Surface for the 33rd Location (marked with a bullet).

Deformation-based Nonstationary covariance models

- P. Guttorp and P. D. Sampson (1994): Methods for estimating heterogeneous spatial covariance functions with environmental applications. In G. P. Patil, C. R. Rao (editors): *Handbook of Statistics XII: Environmental Statistics:* 663-690. New York: North Holland/Elsevier.
- W. Meiring, P. Guttorp, and P. D. Sampson (1998): Space-time Estimation of Grid-cell Hourly Ozone Levels for Assessment of a Deterministic Model, *Environmental and Ecological Statistics* 5: 197-222.
- P.D. Sampson (2001). Spatial Covariance. In: *Encyclopedia of Environmetrics*.
- P.D. Sampson, D. Damian, and P. Guttorp (2001). Advances in Modeling and Inference for Environmental Processes with Nonstationary Spatial Covariance. In: *GeoENV 2000: Geostatistics for Environmental Applications*, P. Monestiez, D. Allard, R. Froidevaux, eds., Dordrecht: Kluwer, pp. 17-32.
- P.D. Sampson, D. Damian, P. Guttorp, and D.M. Holland (2001). Deformation—based nonstationary spatial covariance modelling and network design. In: *Spatio-Temporal Modelling of Environmental Processes*, Colec·ció «Treballs D'Informàtica I Tecnologia», Núm. 10., J. Mateu and F. Montes, eds., Castellon, Spain: Universitat Jaume I, pp. 125-132.
- D. Damian, P.D. Sampson and P. Guttorp (2003): Variance Modeling for Nonstationary Spatial Processes with Temporal Replications, *Journal of Geophysical Research – Atmosphere*, 108 (D24).
- F. Bruno, P. Guttorp, P.D. Sampson, & D/ Cocchi. (2004). Non-separability of space-time covariance models in environmental studies. In The ISI International Conference on Environmental Statistics and Health: conference proceedings (Santiago de Compostela, July, 16-18, 2003)", a cura di Jorge Mateu, David Holland, Wenceslao González-Manteiga, Universidade de Santiago de Compostela, Santiago de Compostela 2003, pp. 153-161
- John Kent: Statistical Methodology for Deformations

Geometric anisotropy

- **Recall that if $C(x,y) = C(\|x-y\|)$ we have an *isotropic* covariance (circular isocorrelation curves).**

- **If $C(x,y) = C(\|Ax-Ay\|)$ for a linear transformation A, we have *geometric anisotropy* (elliptical isocorrelation curves).**

- **General nonstationary correlation structures are typically locally geometrically anisotropic.**

The deformation idea

In the geometric anisotropic case, write

$$C(x,y) = C(\|f(x) - f(y)\|)$$

where $f(x) = Ax$**. This suggests using a general nonlinear transformation**

$$f: \ R^2 \rightarrow R^d$$

G-plane → D-space

Usually d **= 2 or 3.**

We do not want f **to fold.**

Remark: Originally introduced as a multidimensional scaling problem: find Euclidean representation with intersite distances monotone in spatial dispersion, D(x,y)

Space-time Model with Spatial Deformation

Damian et al., 2000 (*Environmetrics*), 2003 (*JGR*)

$$Z(x,t) = \mu(x,t) + \nu(x)^{1/2} H_t(x) + \varepsilon(x,t)$$

$\mu(x,t)$ spatio-temporal trend
parametric in time; mv spatial process

$\nu(x)$ temporal variance at x,
log-normal spatial process

$\varepsilon(x,t)$ msmt error and short-scale variation
$N(0,\sigma_\varepsilon^2)$, independent of $H_t(x)$

$H_t(x)$ mean 0, var 1, 2nd-order cont. spatial process
$C(x,y) = \text{Cov}(H_t(x), H_t(y)) \xrightarrow[x \to y]{} 1.$

$$\text{Cov}(Z(x,t), Z(y,t)) = \begin{cases} \sqrt{\nu(x)\nu(y)}\, C(x,y) & x \neq y \\ \nu(x) + \sigma_\varepsilon^2 & x = y \end{cases}$$

Model (cont.)

$H_t(x)$ mean 0, var 1, 2nd-order cont. spatial process

$\mathrm{Cov}(H_t(x), H_t(y)) \xrightarrow[x \to y]{} 1.$

$$\mathrm{Cor}\left(H_t(x), H_t(y)\right) = \rho_\theta\left(\|f(x) - f(y)\|\right)$$

$f : G \to D$ smooth, bijective
(Geographic $\to$ Deformed plane)

$\rho_\theta(d)$ isotropic correlation function
in a known parametric family
(exponential, power exp, Matern)

i.e. The correlation structure of the spatial process is an (isotropic) function of Euclidean distances between site locations after a bijective transformation of the geographic coordinate system.

Model (cont.)

The spatial deformation f encodes the nonstationarity: spatially varying local anisotropy.

We model this in terms of observation sites $x_1, x_2, \ldots, x_N$ as a pair of thin-plate splines:

$$f(x) = c + \mathbf{A}x + \mathbf{W}^{\mathrm{T}}\sigma(x)$$

$c + \mathbf{A}x$ Linear part: global/large scale anisotropy $\quad c_{2\times 1},\ \mathbf{A}_{2\times 2}$

$\mathbf{W}^{\mathrm{T}}\sigma(x)$ Non-linear part, decomposable into components of varying spatial scale: $\quad \mathbf{W}_{N\times 2},\ \sigma(x)_{N\times 1}$

$$\sigma(x) = \begin{bmatrix} \sigma(x - x_1) \\ \vdots \\ \sigma(x - x_N) \end{bmatrix} \qquad \sigma(h) = \begin{cases} \|h\|^2 \log(\|h\|) & \|h\| > 0 \\ 0 & \|h\| = 0 \end{cases}$$

⇒ **Model parameters**: $f: \{c, \mathbf{A}, \mathbf{W}\},\ \mu,\ \theta,\ \sigma_\varepsilon^2,\ \nu: \{\tilde{\mu}, \tilde{\theta}, \tilde{\sigma}^2\}$

Implementation

Consider observations at sites $x_1, \ldots, x_n$. Let $\hat{C}_{ij}$ be the empirical covariance between sites x_i and x_j. Minimize

$$(\theta, f) \mapsto \sum_{i,j} w_{ij} \left(\hat{C}_{ij} - C\left(\left| f(x_i) - f(x_j) \right|; \theta \right) \right)^2 + \lambda J(f)$$

where $J(f)$ is a penalty for non-smooth transformations, such as the bending energy

$$J(f) = \iint \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] dx dy$$

When f is computed as a thin-plate spline, the minimization above can be considered in terms of the deformed coordinates, $\xi_i = f(x_i)$, or the parameters of the analytic representation of the thin-plate spline, $\{c, \mathbf{A}, \mathbf{W}\}$

More on the **equations** of the thin-plate spline

$$f(x) = \left(f_1(x), f_2(x)\right)^T : \mathbb{R}^2 \to \mathbb{R}^2$$

minimizing "bending energy" subject to interpolation constraints

$$f_j(x_i) = \xi_{ij}, \; 1 \le i \le N; \; j = 1, 2,$$

is an equation of the form

$$f(s) = c + \mathbf{A}s + \mathbf{W}^T \tilde{\sigma}(s)$$

where the coefficients $\mathbf{W}$ satisfy $1^T \mathbf{W} = 0, \;\; \mathbf{X}^T \mathbf{W} = 0.$

I.e. the columns W_1 and W_2 of $\mathbf{W}$ are vectors in the subspace spanned by $\{1, X_1, X_2\} : \mathbf{V} = \left\{v \in \mathbb{R}^N : v^T 1 = 0, \; v^T X_1 = 0, \; v^T X_2 = 0\right\}.$

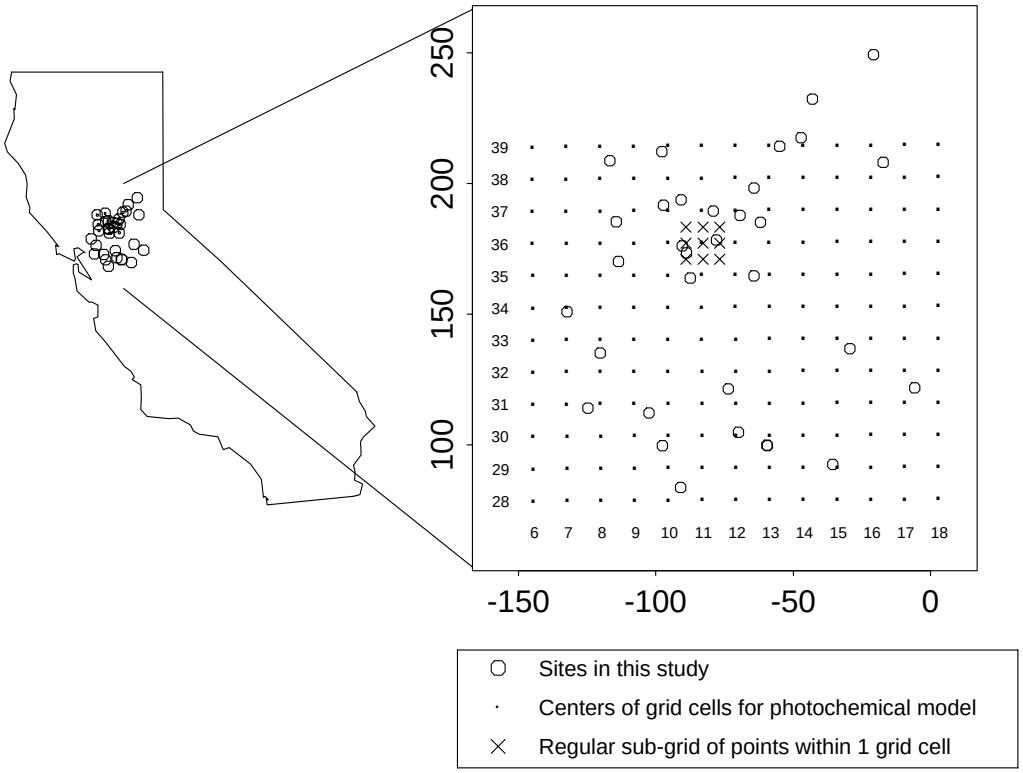
The system of equations for computation of a thin-plate spline is

$$\begin{bmatrix} \boldsymbol{\Xi} \\ 0 \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{\mathbf{S}} & 1 & \mathbf{X} \\ 1^T & 0 & 0 \\ \mathbf{X}^T & 0 & 0 \end{bmatrix}}_{\Gamma} \begin{bmatrix} \mathbf{W} \\ c^T \\ \mathbf{A}^T \end{bmatrix},$$ where $\tilde{\mathbf{S}}$ is $N \times N$ with elements

$\tilde{\mathbf{S}}_{ij} = \sigma(x_i = x_j),$ and the "bending energy" is $J(f) = \mathrm{tr}(\mathbf{W}^T \tilde{\mathbf{S}} \mathbf{W})$

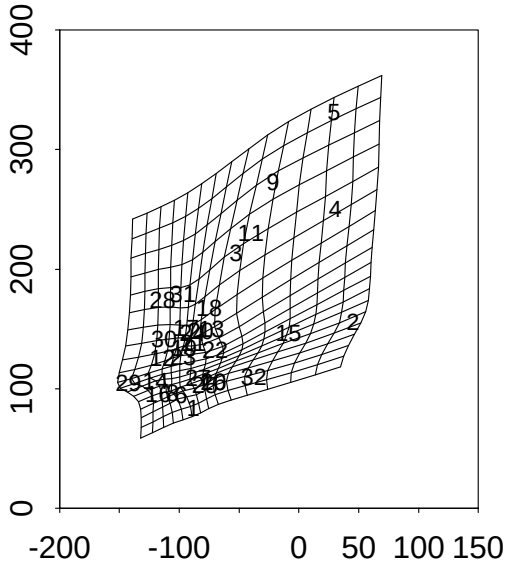
SARMAP

An ozone monitoring exercise in California, summer of 1990, collected data on some 130 sites.



Transformation

This is for hr. 16 in the afternoon



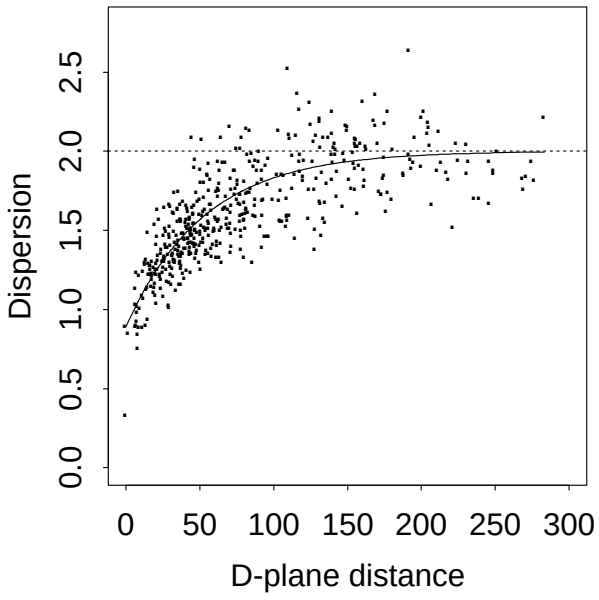
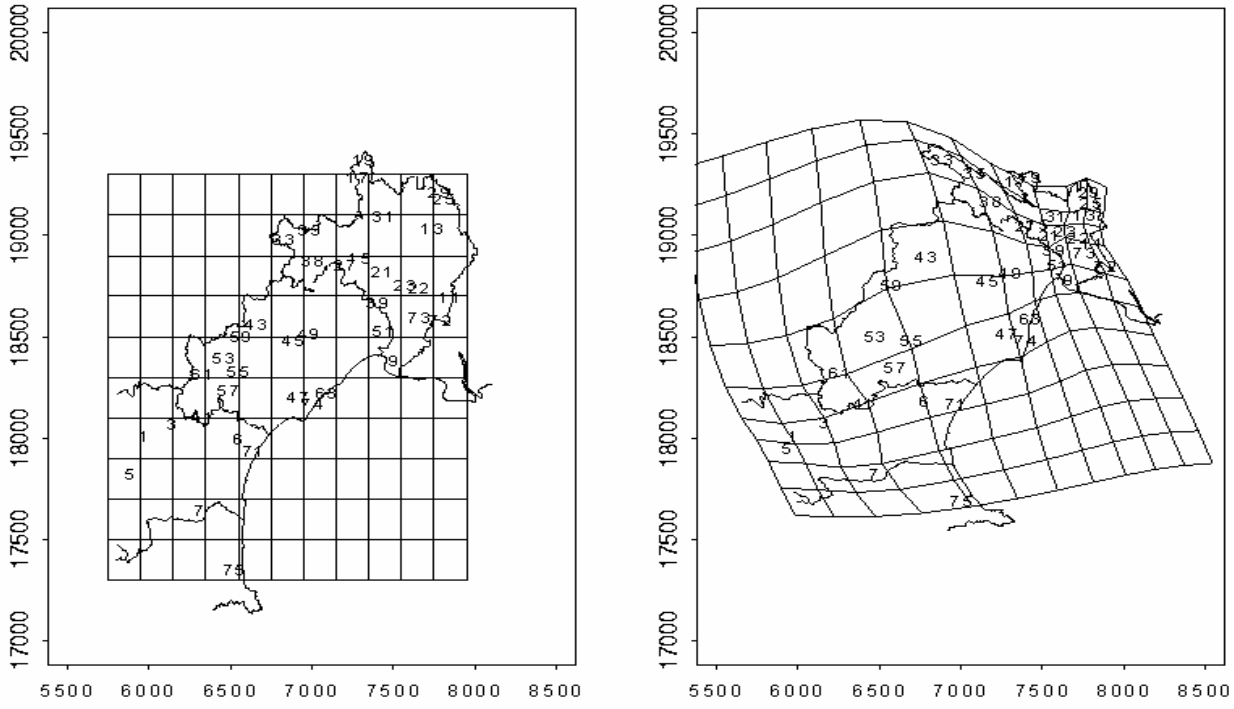
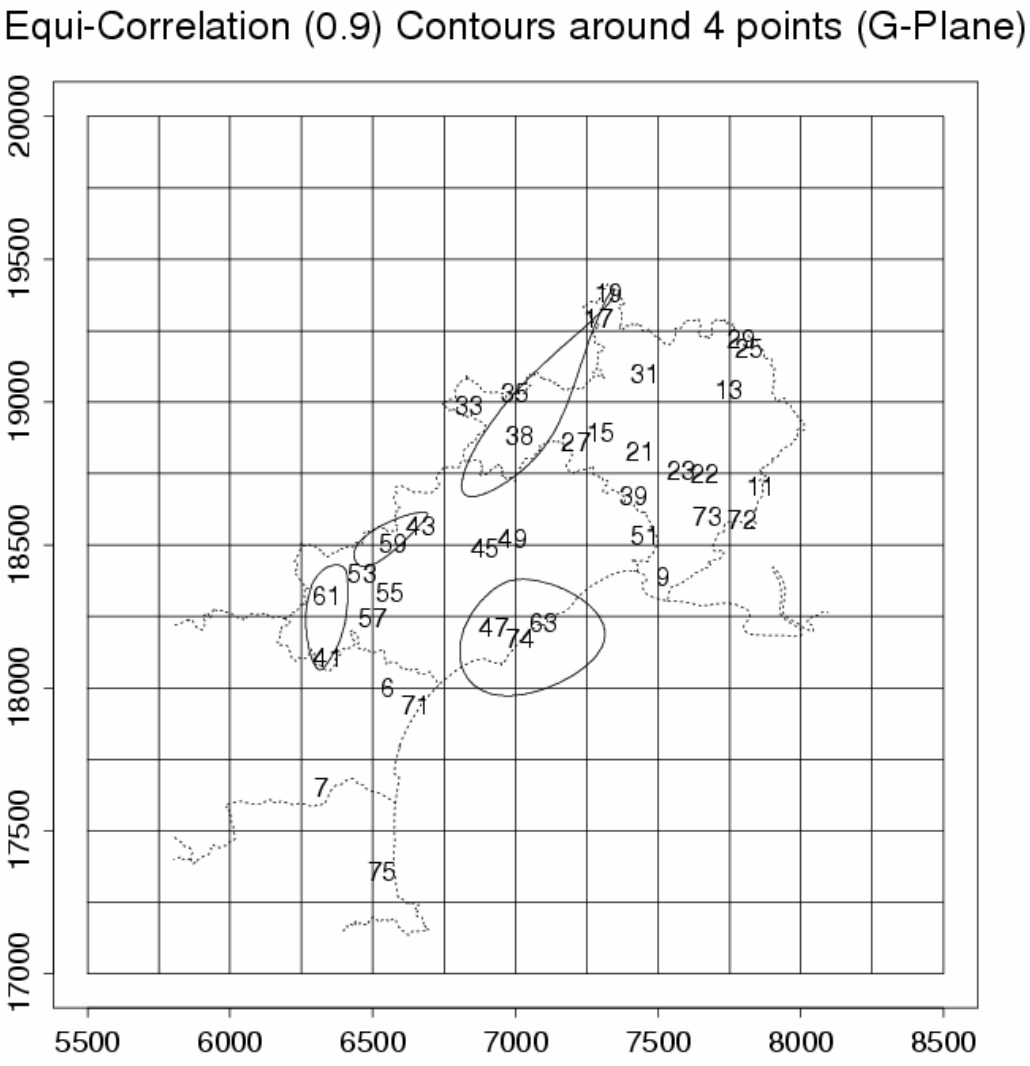


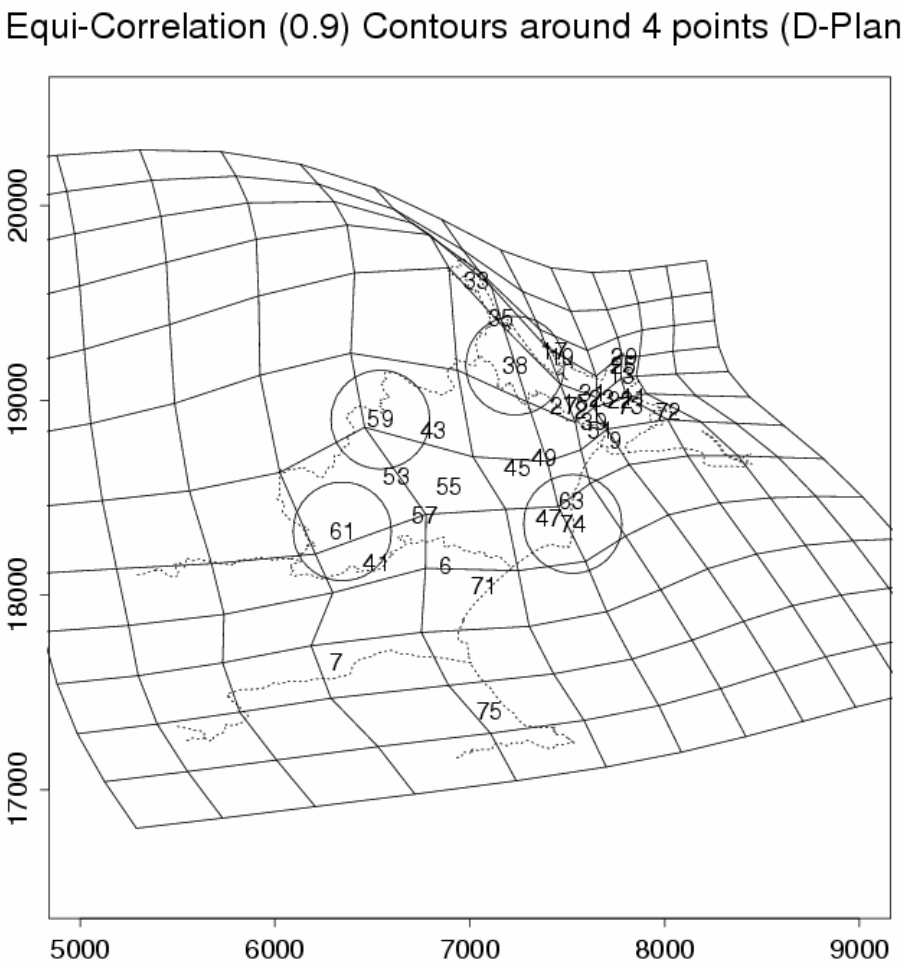
Fig. 7: Precipitation in Southern France -
an example of a non-linear deformation



G-plane Equicorrelation Contours



D-plane Equicorrelation Contours



Theoretical properties of the deformation model

Identifiability

Perrin and Meiring (1999): Let

$$D(x,y) = \gamma\left(\left\|f(x) - f(y)\right\|\right),\ (x,y) \in R^n \times R^n$$

If (1) f and f^{-1} are differentiable in R^n

(2) $\gamma(u)$ is differentiable for $u>0$

then (f,γ) is unique, up to a scaling for γ and a homothetic transformation for f (rotation, scaling, reflection)

A Bayesian implementation

Likelihood:

$$L(S \mid \Sigma) = (2\pi|\Sigma|)^{-(T-1)/2} \exp\left\{-\frac{T}{2} tr\Sigma^{-1}S\right\}$$

Nonlinear part: Bending energy Prior:

$$p(W) \propto \exp\left(-\frac{1}{2\tau}\sum_{i=1}^{2} W_i'\tilde{S}W_i\right)$$

Linear part:

- –fix two points in the G-D mapping
- –put a (proper) prior on the remaining two parameters

Posterior computed using Metropolis-Hastings

Likelihood: given observation vectors $Z_1,\ldots,Z_N$ of length T:

$$\mathcal{L}(\mu,\theta,f,\nu,\sigma_\varepsilon^2;Z_1,\ldots,Z_N)=\left[Z_1,\ldots,Z_N\mid\mu,\boldsymbol{\Sigma}\right]=\left[\bar{Z},\mathbf{S}\mid\mu,\boldsymbol{\Sigma}\right]=$$

$$=\left|2\pi\boldsymbol{\Sigma}\right|^{-T/2}\exp\left\{-\frac{(T-1)}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\mathbf{S}-\frac{T}{2}\left(\bar{Z}-\mu\right)'\boldsymbol{\Sigma}^{-1}\left(\bar{Z}-\mu\right)\right\}$$

with covariance matrix having elements

$$\sigma_{ij}=\begin{cases}\sqrt{\nu_i\nu_j}\,\rho_\theta\left(\left\|\xi_i-\xi_j\right\|\right) & i\neq j\\ \nu_j+\sigma_\varepsilon^2 & i=j\end{cases},1\le i,j\le N$$

Integrating out a flat prior on the (constant) mean

$$\left[\mu\right]\propto 1\Rightarrow$$

$$\left[\mathbf{S}\mid\boldsymbol{\Sigma}\right]=\int\left[\bar{Z},\mathbf{S}\mid\mu,\boldsymbol{\Sigma}\right]\left[\mu\right]\mathrm{d}\mu\propto\left|\boldsymbol{\Sigma}\right|^{-(T-1)/2}\exp\left\{-\frac{(T-1)}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\mathbf{S}\right\}$$

Posterior:

Log-normal variance field: $\left[\nu \middle| \tilde{\mu}, \tilde{\sigma}^2, \tilde{\theta}, \{c, \mathbf{A}, \mathbf{W}\}\right] \propto$

$$\frac{1}{\prod \nu(x_i)} \left|\tilde{\mathbf{\Sigma}}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\log\nu - \tilde{\mu}\cdot\mathbf{1})'\tilde{\mathbf{\Sigma}}^{-1}(\log\nu - \tilde{\mu}\cdot\mathbf{1})\right\}'$$

Full posterior is: $\left[\theta, f:\{c, \mathbf{A}, \mathbf{W}\}, \sigma_\varepsilon^2, \nu, \tilde{\mu}, \tilde{\sigma}^2, \tilde{\theta} \middle| \mathbf{S}\right] \propto$

$$\left[\mathbf{S}|\mathbf{\Sigma}\right]\cdot[\theta]\left[\sigma_\varepsilon^2\right]\left[\nu \middle| \tilde{\mu}, \tilde{\sigma}^2, \tilde{\theta}, c, \mathbf{A}, \mathbf{W}\right][c, \mathbf{A}][\mathbf{W}][\tilde{\mu}]\left[\tilde{\sigma}^2\right][\tilde{\theta}]$$

$[c, \mathbf{A}]$: diffuse normal prior on 2 free linear params (4 constr)

$$[\mathbf{W}] \propto \exp\left\{-\frac{1}{2\tau}(\mathbf{W}_1'\tilde{\mathbf{S}}\mathbf{W}_1 + \mathbf{W}_2'\tilde{\mathbf{S}}\mathbf{W}_2)\right\} I_{\{\mathbf{W}\in\mathbf{V}\times\mathbf{V}\}}, \quad \tilde{\mathbf{S}}_{ij} = \sigma(x_i - x_j)$$

the "bending energy" prior on space orthogonal to linear

Summary of prior distributions:

Deformation parameters

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} s1a & 0 \\ 0 & s2a \end{bmatrix} \right)$$

$$[\mathbf{W}] \propto \exp\left\{ -\tfrac{1}{2\tau}\left(W_1^T \tilde{\mathbf{S}} W_1 + W_2^T \tilde{\mathbf{S}} W_2 \right) \right\} \mathbf{I}_{\{\mathbf{W} \in \mathbf{V} \times \mathbf{V}\}}.$$

Correlation parameters

$\theta_1 \sim \exp(ptheta)$

$\theta_2 \sim U(0, ptheta2)$ **---** (if power exponential)

Variance parameters

$\sigma_\varepsilon^2 \sim \exp(pnugget)$

Summary or prior distributions (cont)

Variance parameters (cont)

$$[\nu] \propto \frac{1}{\prod_{i=1}^{N} \nu_i} \left|\tilde{\Sigma}\right|^{-1/2} \exp\left\{-\tfrac{1}{2}\left(\log(\nu) - \mu \cdot 1\right)^T \right) \tilde{\Sigma}^{-1} \left(\log(\nu) - \mu \cdot 1\right)^T \right)\}$$

$\tilde{\Sigma}$ is $N \times N$ with elements $\tilde{\sigma}_{ij} = \tilde{\sigma}^2 \rho_{\tilde{\theta}}\left(\left\|f(x_i) - f(x_j)\right\|\right)$,

$$\rho_{\tilde{\theta}}(d) = \exp\left(-\tilde{\theta} d\right)$$

$$\tilde{\sigma}^2 \sim \Gamma(palpha, pbeta)$$

$$\tilde{\theta} \sim \exp(pthetat)$$

$$\mu \sim N(pmu, psigma2)$$

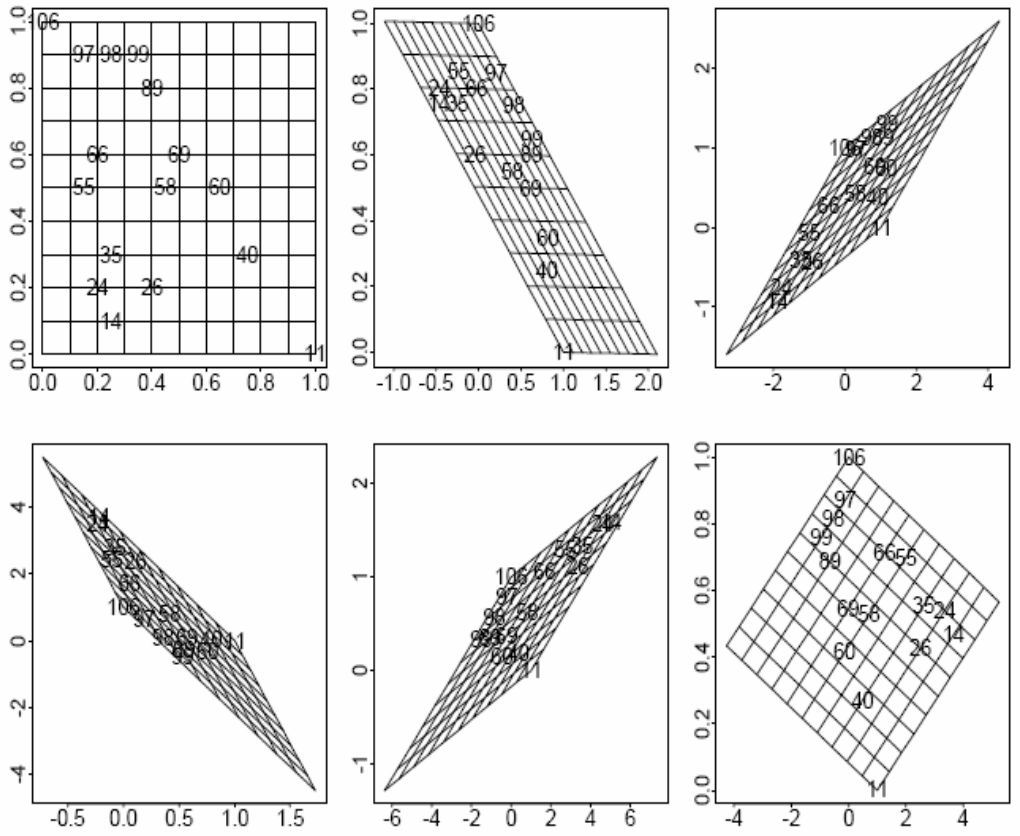


Figure 3.1: *The linear part of the spatial deformation of the unit square. The upper left panel shows the square itself. The other panels represent five random samples from the prior distribution. The two free linear components were sampled independently, with means 0 and 1 respectively and prior variances equal to 10.*

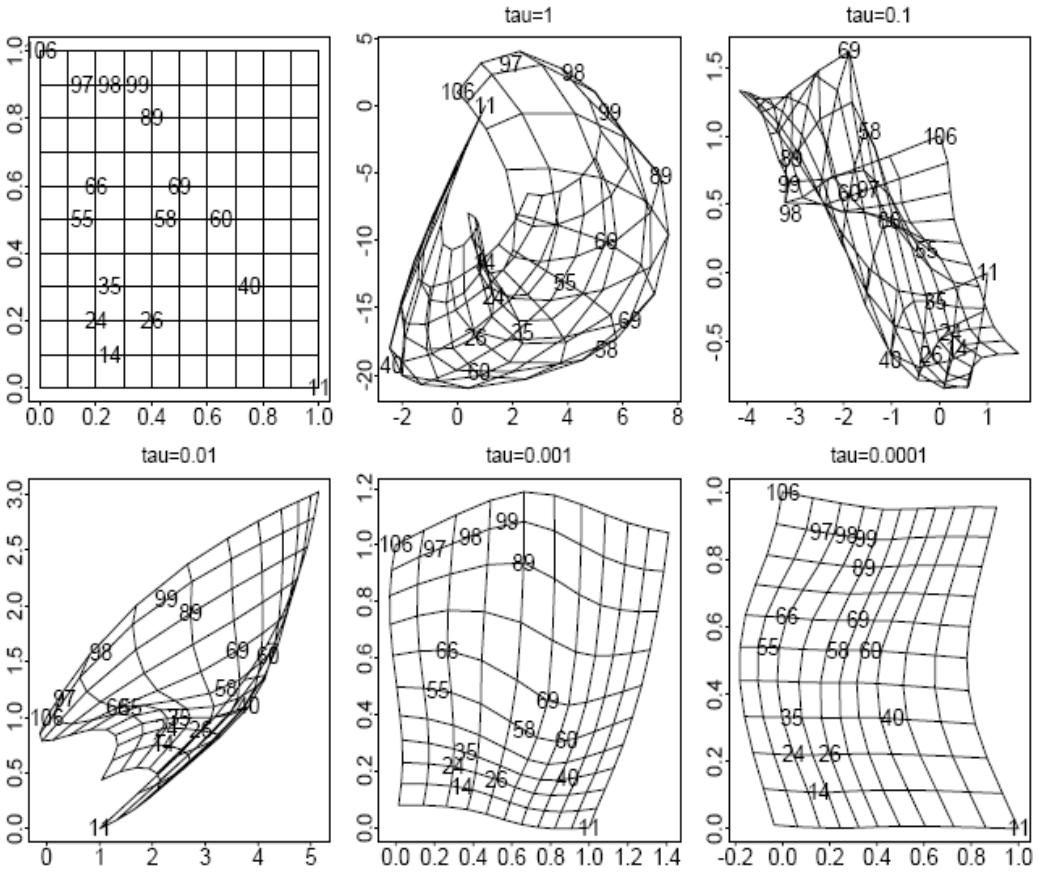


Figure 3.2: *The effect of τ on the non-linear part of the deformation of the unit square (samples from the prior distribution). The upper left panel shows the square itself. As τ gets smaller, the deformations sampled become smoother.*

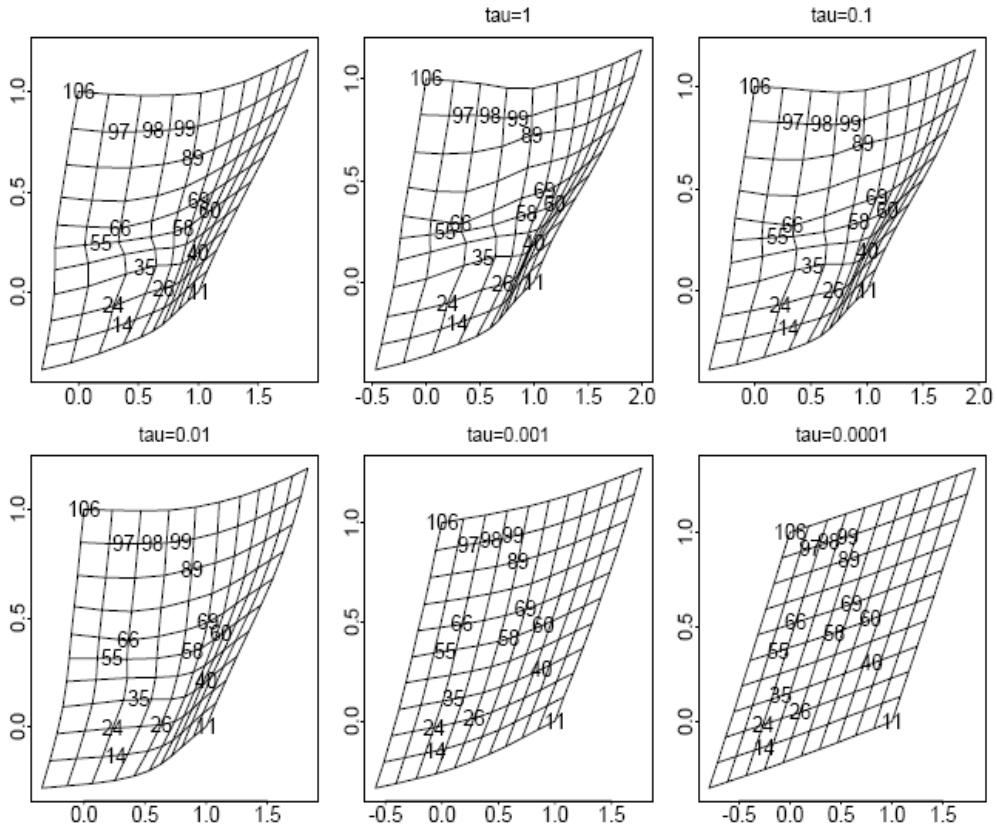


Figure 3.3: *The effect of τ on the posterior deformation of the unit square. The upper left panel shows the true deformation. As τ gets smaller, the estimated deformations become smoother. The two free linear components had prior means 0 and 1 respectively and variances equal to 10. The data consisted of 200 samples at the 16 observation locations shown on the graphs.*

Computation

Metropolis-Hastings algorithm for sampling from the highly multidimensional posterior.

Given estimates of D-plane locations, $f(x_i)$, the transformation is extrapolated to the whole domain using thin-plate splines. (Visualization and diagnostics.)

Predictive distributions for

(a) temporal variance at unobserved sites,

(b) the spatial covariance for pairs of observed and/or unobserved sites,

(c) the observation process at unobserved sites.

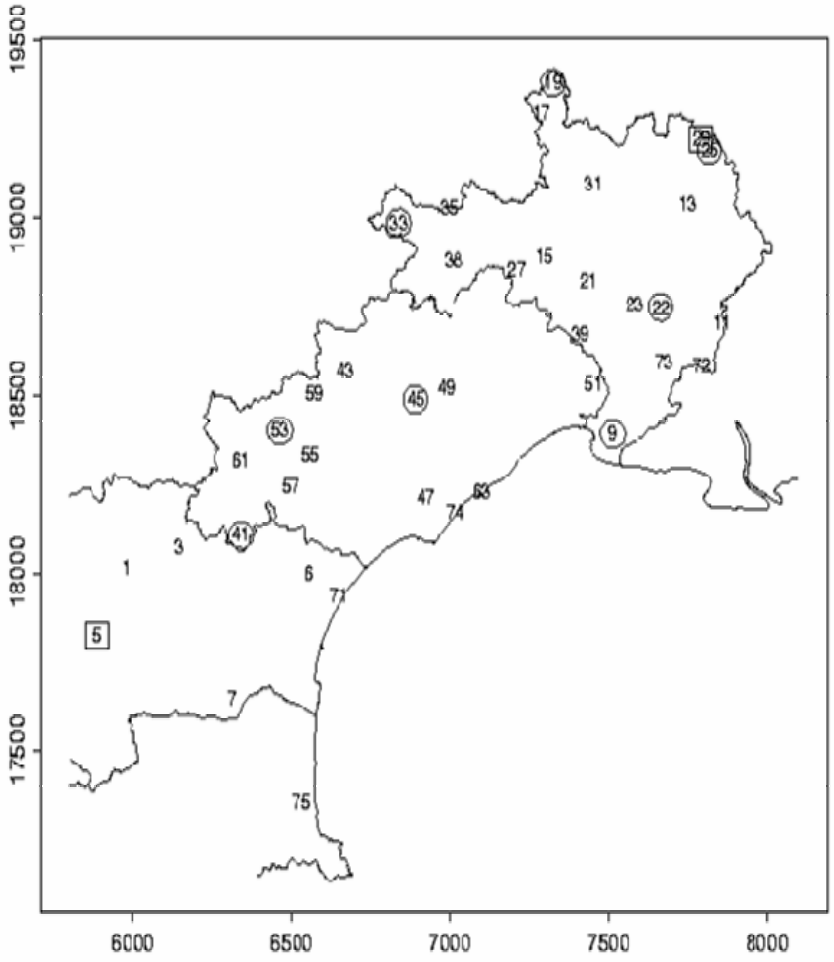
Application to Languedoc-Roussillon Precipitation Data

108 altitude-adjusted, 10-day aggregate precip records at 39 sites (Nov-Dec, 1975-1992).

Data log-transformed and site-specific means removed (for this analysis).

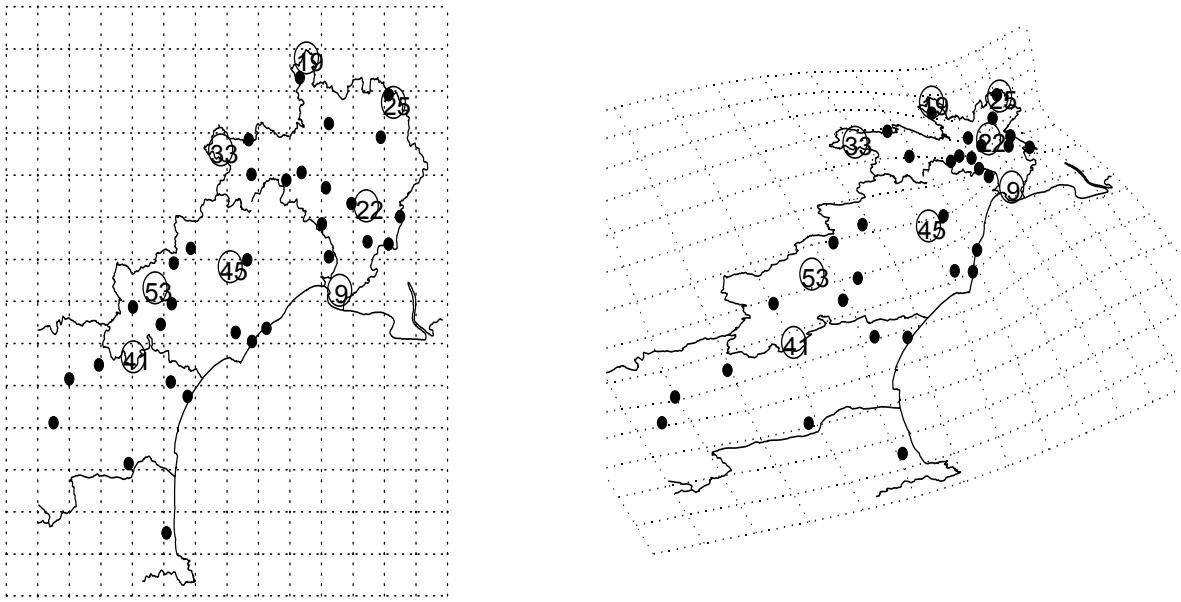
Estimated deformation is non-linear: correlation stronger in the NE region, weaker in the SW.

Languedoc-Roussillon Precipitation



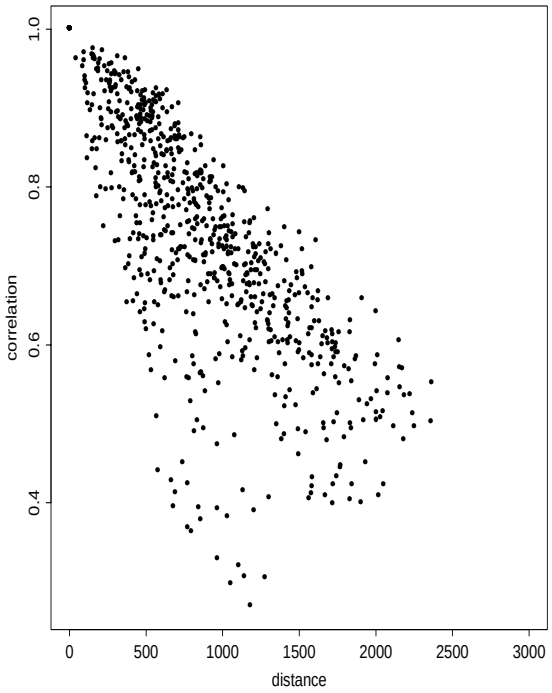
Estimated deformation of Languedoc-Roussillon region

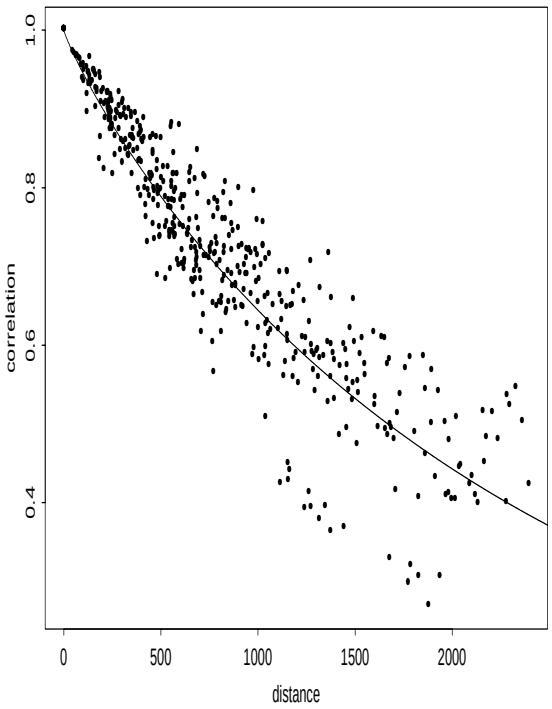
(a) (b)



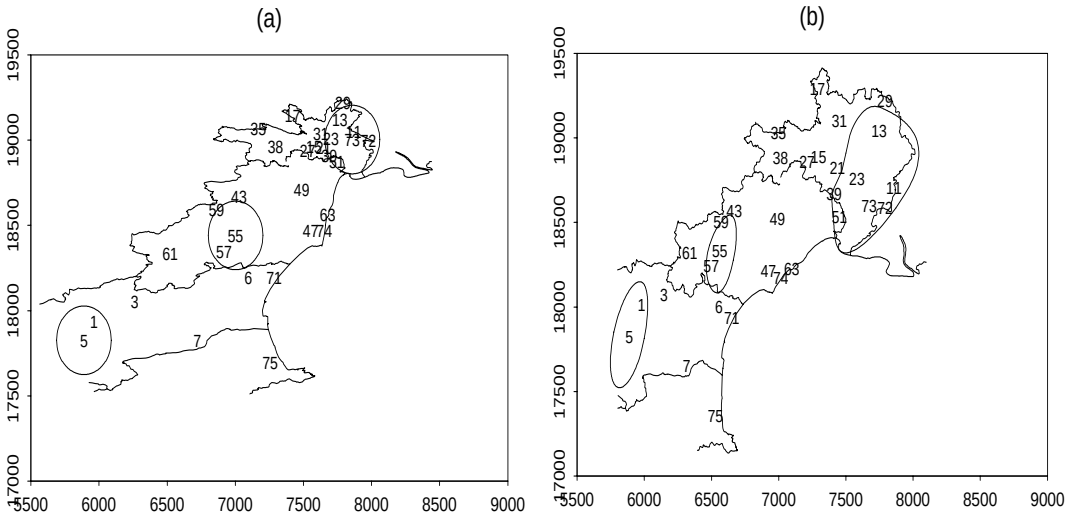
Circled monitoring sites are reserved for model validation

Correlation vs Distance in G-plane and D-plane

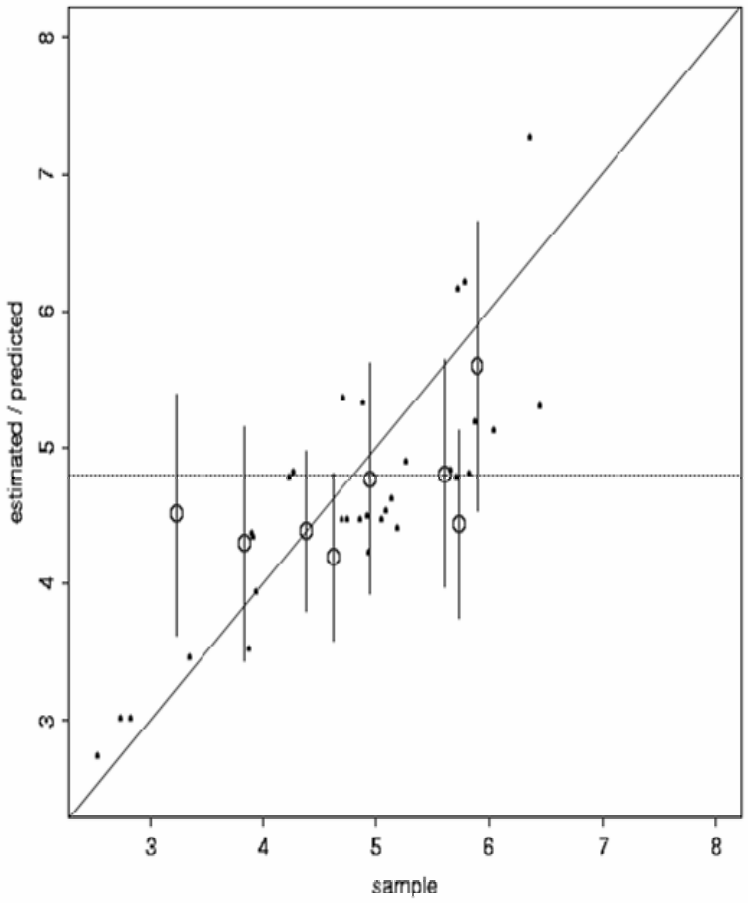




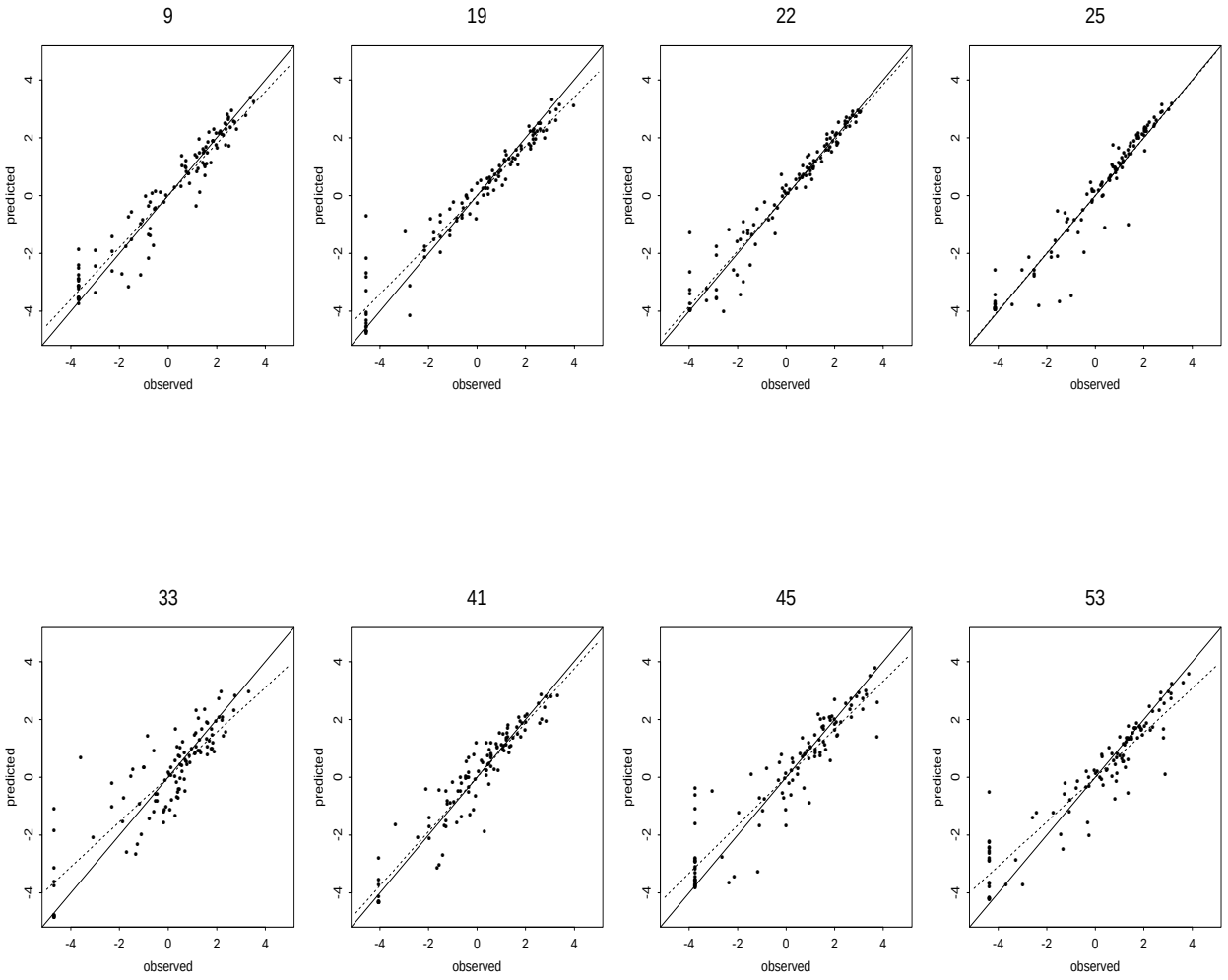
Equi-correlation (0.9) contours D-plane (a) and G-plane (b)



Estimated (•) and predicted (◦) variances, $v(x_0)$, vs. observed

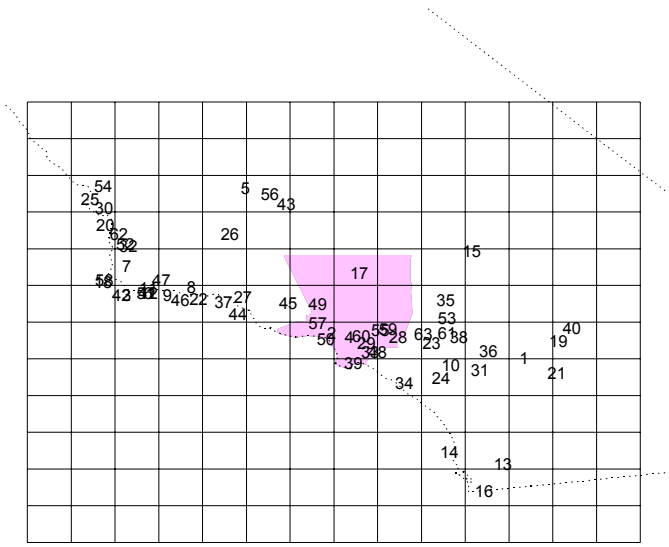


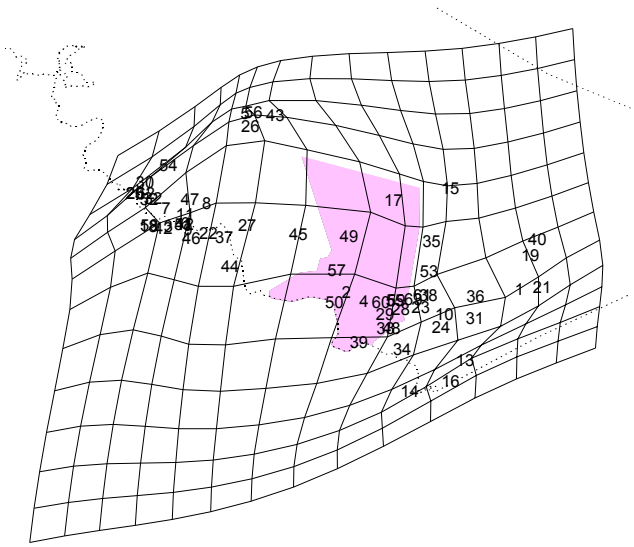
Assessment of (10-day aggregate) precipitation predictions at validation sites

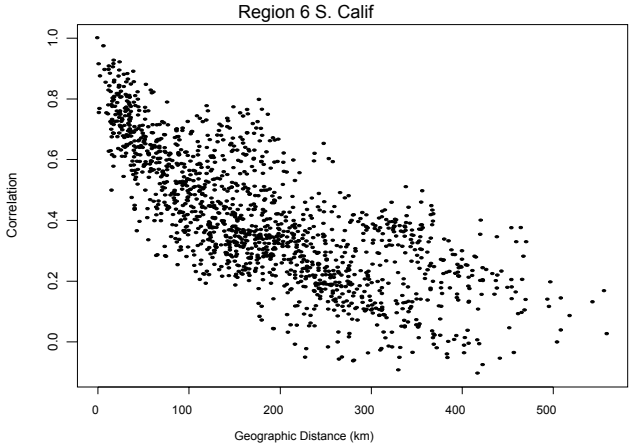


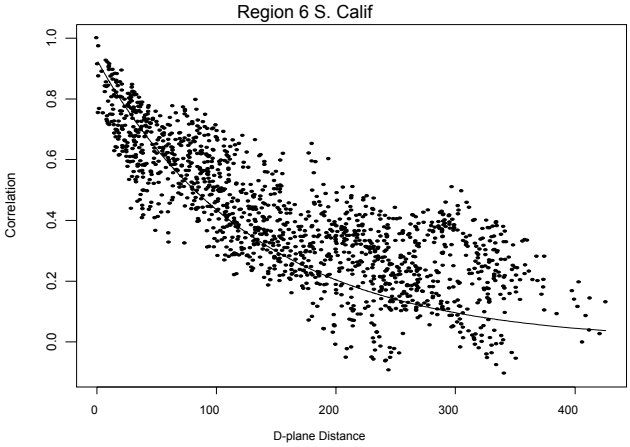
California ozone

63 Region 6 monitoring sites and their representation in a
deformed coordinate system reflecting spatial covariance
Thu Oct 30 00:12:36 PST 2003



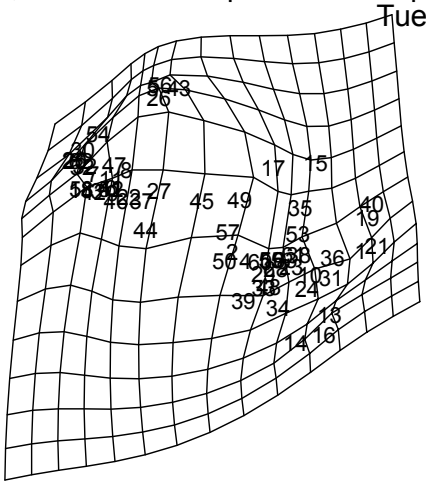


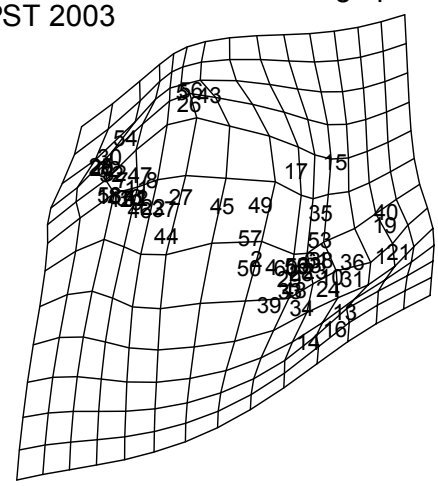

Region 6 S. Calif
Correlation
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
500
Geographic Distance (km)

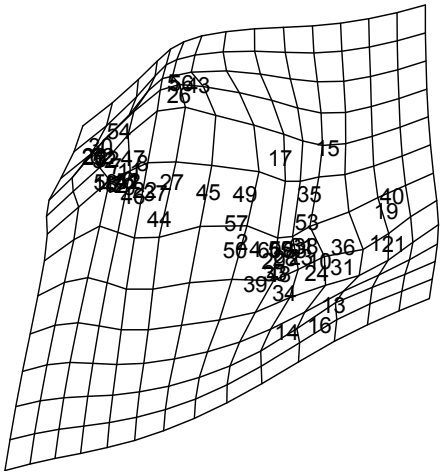

Region 6 S. Calif
Correlation
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
D-plane Distance

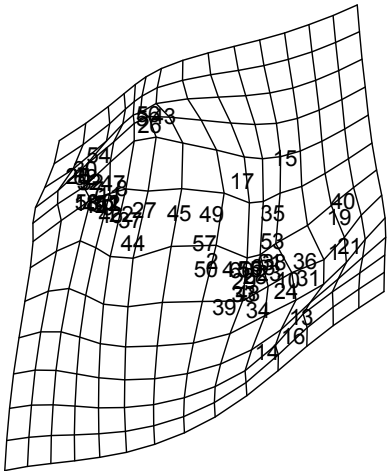
Posterior samples

N=63, S. Calif: 4 samples from the posterior distribution of deformations reflecting spatial covariance
Tue Oct 28 22:18:29 PST 2003









Other approaches:

- Haas, 1990, Moving window kriging
- Nott & Dunsmuir, 2002, Biometrika—computationally convenient, but …
- Higdon & Swall, 1998, 2000, Gaussian moving averages or "process convolution" model
- Fuentes, 2002, Kernel averaging of orthogonal, locally stationary processes.
- Kim, Mallock & Holmes, 2005, Piecewise Gaussian modeling
- Pintore & Holmes, 2005, Fourier and Karhunen-Loeve expansions

Gaussian moving averages

Higdon (1998), Swall (2000):

Let ξ be a Brownian motion without drift, and $X(s) = \int_{R^2} b(s-u)d\xi(u)$. This is a Gaussian process with correlogram

$$\rho(d) = \int_{R^2} b(u)b(u-d)du.$$

Account for nonstationarity by letting the kernel b vary with location:

$$\rho(s_1,s_2) = \int_{R^2} b_{s_1}(u)b_{s_2}(u)du$$

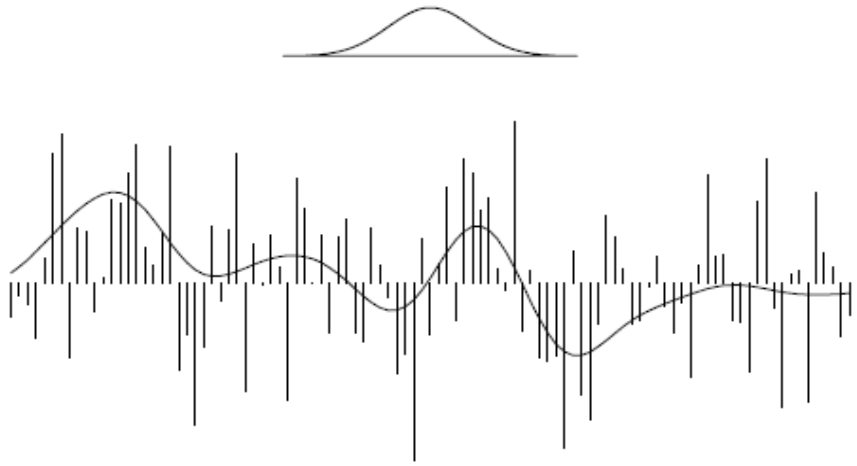


Figure 3.1: A smooth, stationary random field obtained by taking a convolution of i.i.d. normal realizations, using the normal kernel pictured at top.

Details

For Gaussian kernels, one can show that the nonstationary covariance takes the simple form

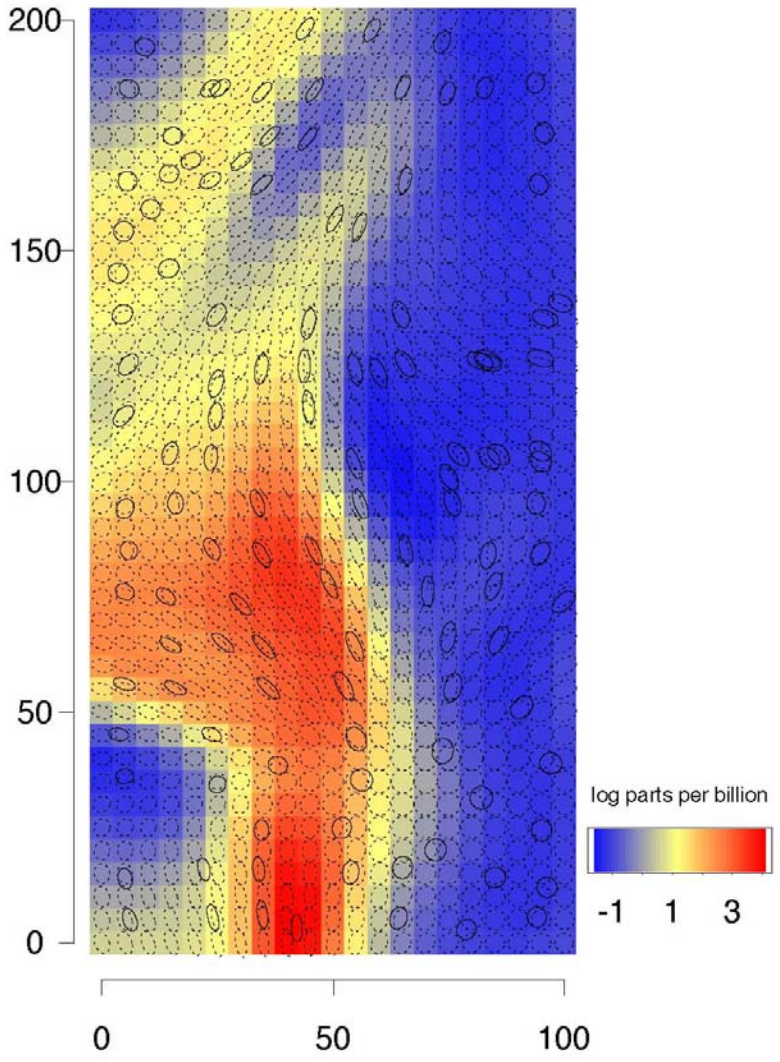
$$C^{NS}(x_i, x_j) = \sigma^2 |\Sigma_i|^{1/4} |\Sigma_i|^{1/4} \left| \frac{\Sigma_i + \Sigma_j}{2} \right|^{-1/2} \exp(-Q_{ij})$$

where

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

And where $\Sigma_i = \Sigma(x_i)$, the kernel matrix, is the covariance matrix of the Gaussian kernel centered at x_i.

Swall & Higdon. Process convolution approach, Posterior mean and covariance kernel ellipses.



Paciorek & Schervish thm:

If an isotropic correlation function, $R^s(\tau)$, is positive definite on R^p for every p = 1; 2; . . . , then the function

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

is a nonstationary correlation function.

The authors use a Matern correlation function. The challenge is specifying and estimating a field of smoothly varying kernels as a Gaussian process.

Kernel averaging

Fuentes (2000): Introduce orthogonal local stationary processes $Z_k(s)$, k=1,...,K, defined on disjoint subregions S_k and construct

$$Z(s) = \sum_{k=1}^{K} w_k(s) Z_k(s)$$

where $w_k(s)$ is a weight function related to dist(s,S_k). Then

$$\rho(s_1, s_2) = \sum_{k=1}^{K} w_k(s_1) w_k(s_2) \rho_k(s_1 - s_2)$$

A continuous version has

$$Z(s) = \int w(x - s) Z_{\theta(s)}(x) ds$$

Some recent atmospheric science literature and proposals for spatio-temporal covariance models

Desroziers, 1997, A coordinate change for data assimilation in spherical geometry of frontal structures. ***Monthly Weather Review*****.**

The main impact of this transformation in the framework of data assimilation is that it enables the use of anisotropic forecast correlations that are flow dependent.

Riishojgaard, 1998. A direct way of specifying flow-dependent background correlations for meteorological analysis systems. ***Tellus*****.**

Weaver and Courtier, 2001. Correlation modelling on the sphere using a generalized diffusion equation. ***Quar. J. Royal Met Soc.***

Generalization to account for anisotropic correlations are also possible by stretching and/or rotating the computational coordinates via a 'diffusion' tensor.

Some recent atmospheric science literature and proposals for spatio-temporal covariance models (cont)

Wu et al., 2002. 3-D variational analysis with spatially inhomogeneous covariances. *Monthly Weather Review*.

Purser et al., 2003, Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Monthly Weather Review.*

Fu et al., 2004, Ocean data assimilation with background error covariance derived from OGCM outputs. *Advances in Atmospheric Sciences*.

Incorporating covariates

- Carroll and Cressie, 1997. geomorphic site attributes in correlation model for snow water equivalent in river basins.

$$c(s_1, s_2) = \exp(-B\|s_1 - s_2\|) - CX_c - DX_d - EX_e - FX_f$$

Where X's represent differences between the two sites in elevation, slope, tree cover, aspect,

Alternative: deform R^2 into subspace of R^6.

- Riishojgaard, 1998, "flow-dependent" correlation structures for meteorological analysis systems. For $z(s)$, a realization of a random field in R^d,

$$c(s_1, s_2) = \varphi_d(\|s_1 - s_2\|) \cdot \varphi_1(|z(s_1) - z(s_2)|)$$

an embedding and deformation of the geographic coordinate space R^d into R^{d+1}, with a separable, stationary correlation model fitted in new coordinate space.

Covariance models for dynamic error structures in the context of data assimilation

- Cox and Isham, 1988: with v a velocity vector in R^2, a physical model for rainfall leads to space-time covariance function

$$c(s_1, s_2; t_1, t_2) = E_{\mathbf{V}} G\left(\left\| (s_2 - s_1) - \mathbf{V}(t_2 - t_1) \right\| \right)$$

where $G(r)$ denotes area of intersection of two disks of unit radius with centers a distance r apart.

There are variants in the meteorological and hydrological literature: depending on tangent line in a barotropic model; using geostrophic or semigeostropic coordinates; or working in a Lagrangian reference frame for convective rainstorms. These yield interesting, anisotropic and nonstationary correlation models (cf. Desroziers, 1997). They suggest interesting space-time extensions of current deformation approach and statistical model fitting questions.