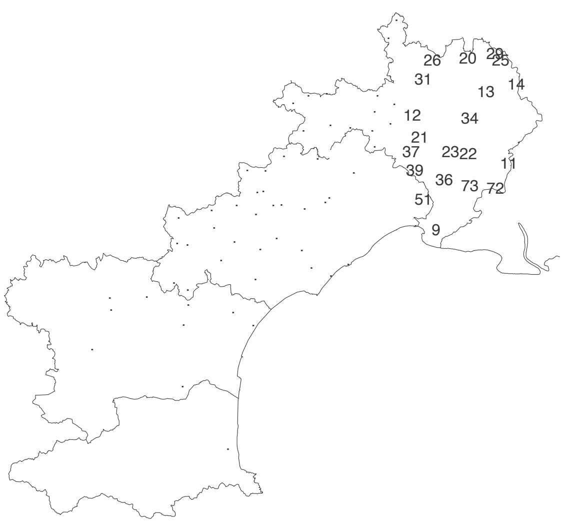
Application: French Precipitation Data, and a few Miscellaneous Comments/Directions

Wendy Meiring
University of California, Santa Barbara

(Joint work with Peter Guttorp, Paul Sampson, Pascal Monestiez)



Environmental and ecological data usually exhibit dependence in space and time

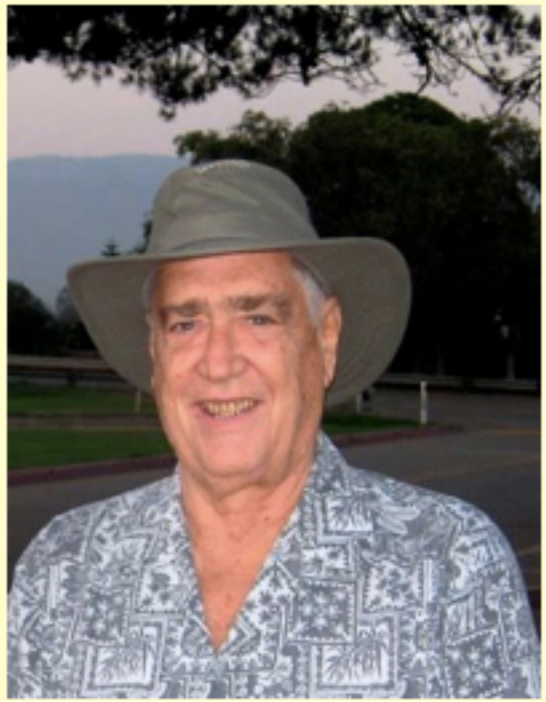
Aim: use this dependence to estimate/predict process of interest in space and time,

taking into account both small-scale space-time dependence, and large-scale space-time trends

Tobler's First Law of Geography

"Everything is related to everything else, but near things are more related than distant things."

Professor Waldo Tobler, UCSB Geography Dept

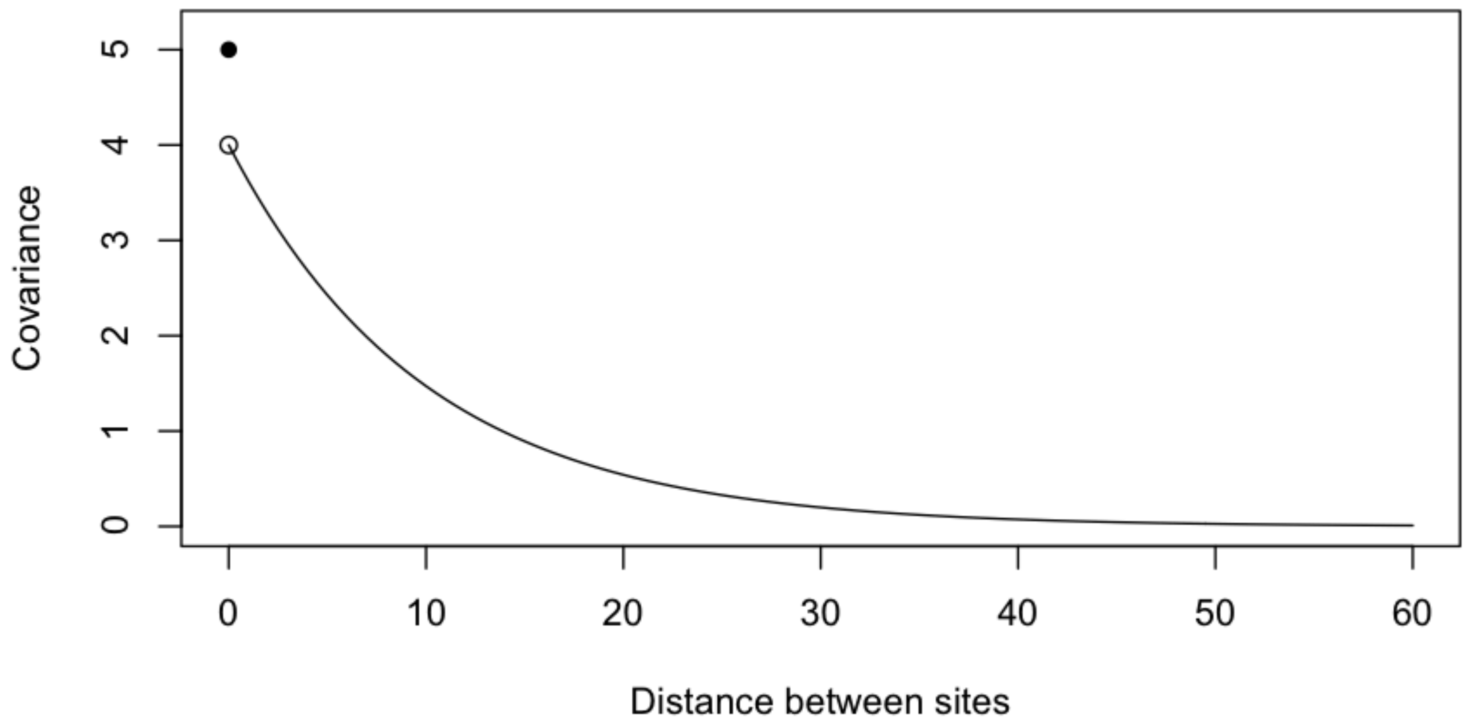


Source: ESRI online GIS dictionary
http://support.esri.com/es/knowledgebase/GISDictionary/term/Tobler's%20First%20Law%20of%20Geography

How can we model associations between observations at two spatial locations?

Frequently use spatial Covariance and Correlation (assuming finite variance field)

An example of an “isotropic” spatial covariance function: an exponential spatial covariance function with nugget



More generally - spatial variogram/structure function/spatial dispersion

Consider data $Z(x_i)$, N monitoring sites
Spatial dispersion between sites x_i and x_j

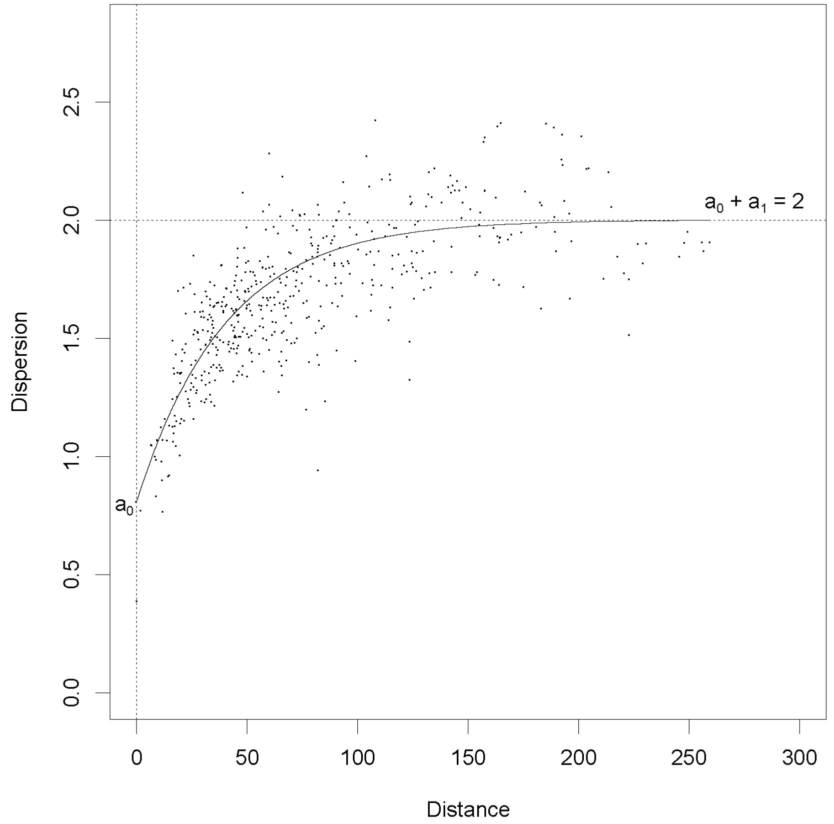
$$D(x_i, x_j) = \text{var}[Z(x_i) - Z(x_j)]$$

(Spatial variogram or structure function)

If spatial covariances exist, then standardizing by variance field gives

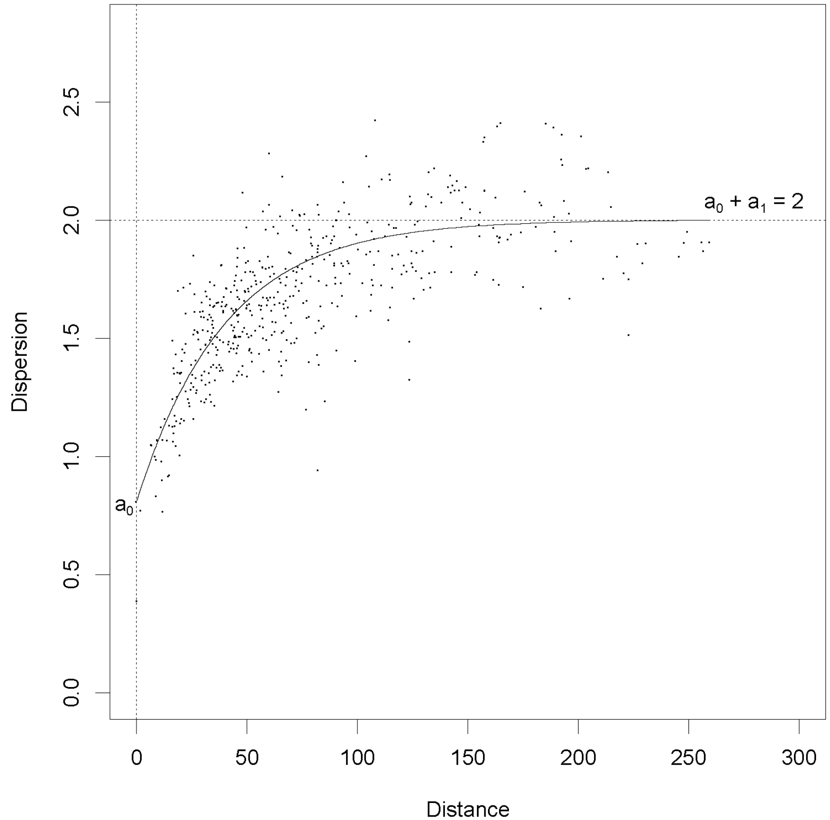
$$D(x_i, x_j) = 2(1 - \text{Corr}[Z(x_i), Z(x_j)])$$

An Exponential variogram (on a standardized scale) - an example of an isotropic model for spatial association



$$D(x_i, x_j) = \begin{cases} 0 & \text{if } h_{ij} = 0 \\ a_0 + a_1\{1 - \exp(-t_0\, h_{ij})\} & \text{if } h_{ij} > 0 \end{cases}$$

An Exponential variogram (on a standardized scale) - an example of an isotropic model for spatial association



Here spatial dependence only depends on the Euclidean distance between geographic locations.

$$D(x_i, x_j) = \begin{cases} 0 & \text{if } h_{ij} = 0 \\ a_0 + a_1 \{1 - \exp(-t_0 \, h_{ij})\} & \text{if } h_{ij} > 0 \end{cases}$$

Non-isotropic spatial covariance modeling

- A deformation approach. **Sampson-Guttorp approach**
- There are many other approaches – active development area. **Non-stationary covariance workshop**

My collaborations on deformation approach:

- Work with Peter Guttorp and Paul Sampson at U. Washington, Pascal Monestiez and Olivier Perrin in France.

Terminology and definitions

1. Consider data $Z(x_i, t)$

   N monitoring sites

   T time points. Independent in time.

(For simplicity in this talk, assume independence in time. Replications.)

2. Spatial dispersion between sites x_i and x_j

$$D\left(x_i, x_j\right) = \text{var}\left[Z(x_i, t) - Z(x_j, t)\right]$$

3. Sample estimates d^*_{ij}.

$$d^*_{ij} = s_{ii} + s_{jj} - 2s_{ij}$$

   where s_{ij} = sample covariance.

   Standardize to correlations.

$$d_{ij} = 2\left(1 - r_{ij}\right)$$

s_{ii} may vary in space. r_{ij} is the sample correlation.

- Consider

$$Z(x,t) \;=\; \mu(x,t) + E_{\tau}(x) + E_{\epsilon}(x,t)$$

- Model

$$d_{ij} \;=\; \gamma_{\theta}\left(\left\| x_i^* \;-\; x_j^* \right\| \right) + e_{ij}$$

Variogram is modeled as isotropic as a function of distance in the deformed space (in D-space)

γ_{θ} isotropic variogram.

$$w:\; \underbrace{\mathbf{R}^p}_{\text{G-space}} \;\rightarrow\; \underbrace{\mathbf{R}^d}_{\text{D-space}}, \qquad x_i^* = w(x_i).$$

- Spatial dispersions are stationary and isotropic in D-space.

- Consider

$$Z(x,t) \; = \; \mu(x,t) + E_{\tau}(x) + E_{\epsilon}(x,t)$$

- Model

$$d_{ij} \; = \; \gamma_{\theta}\left(\left\| x_i^* \; - \; x_j^* \right\| \right) + e_{ij}$$

Variogram is modeled as isotropic as a function of distance in the deformed space (in D-space)

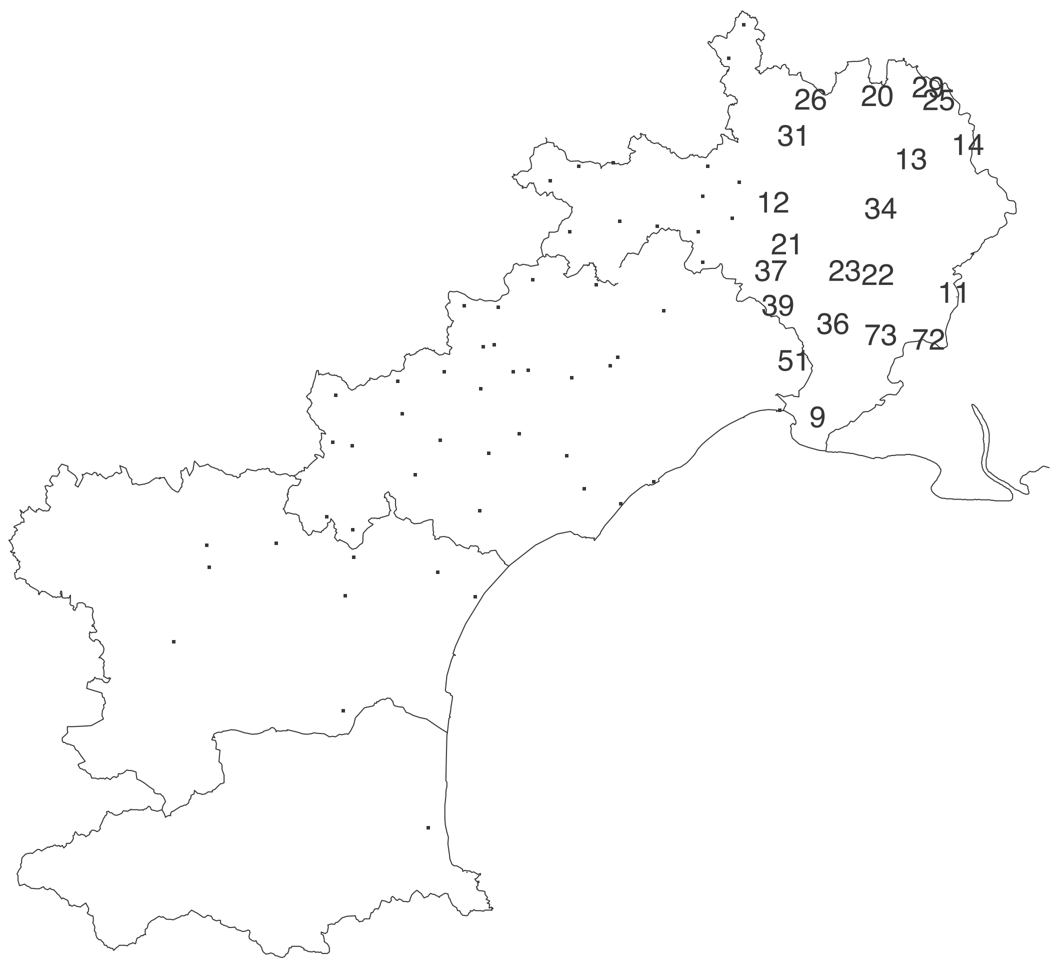
γ_θ isotropic variogram.

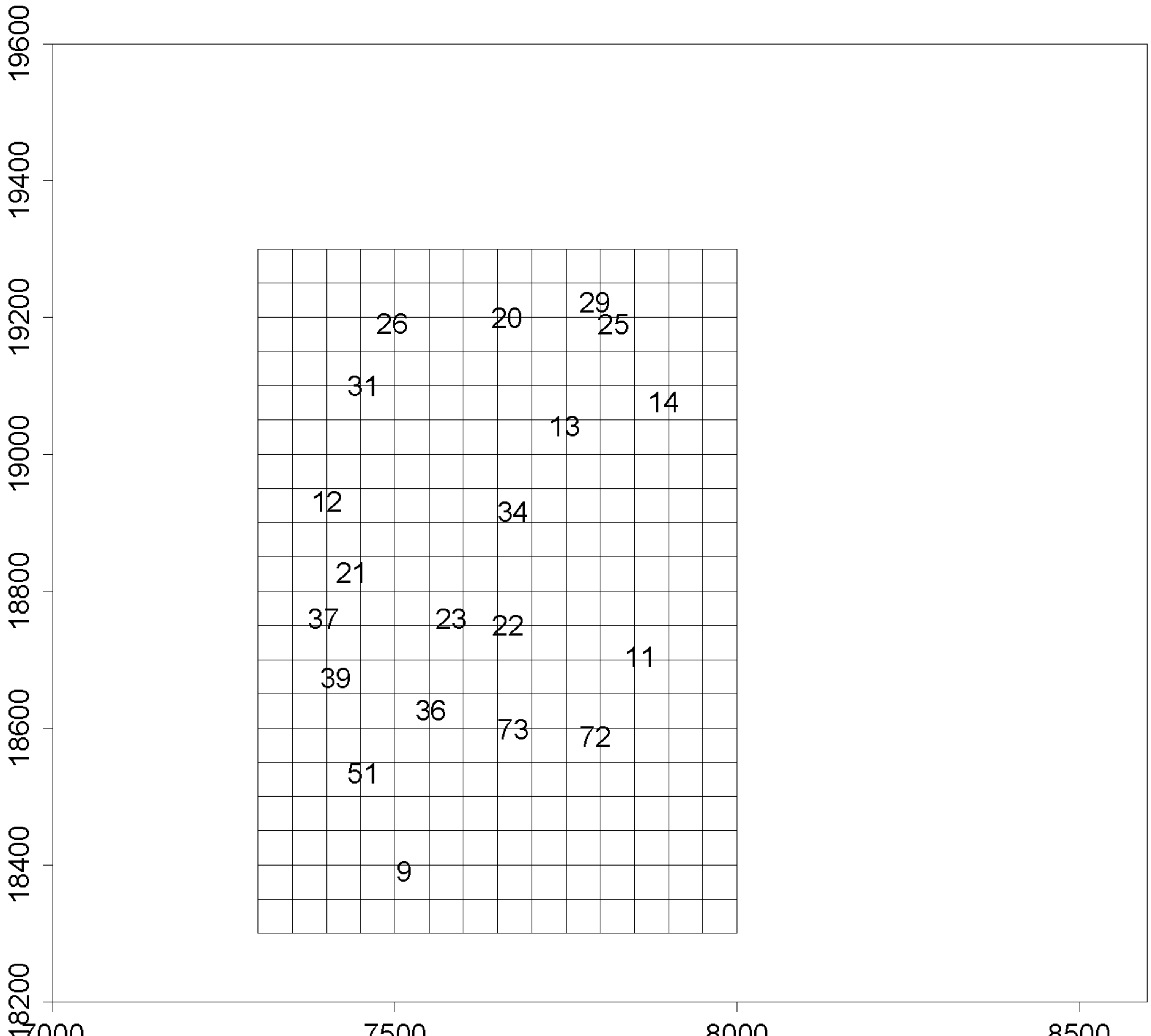
$$w: \underbrace{\mathbf{R}^p}_{\text{G-space}} \to \underbrace{\mathbf{R}^d}_{\text{D-space}}, \qquad x_i^* = w(x_i).$$

In this talk p=d=2

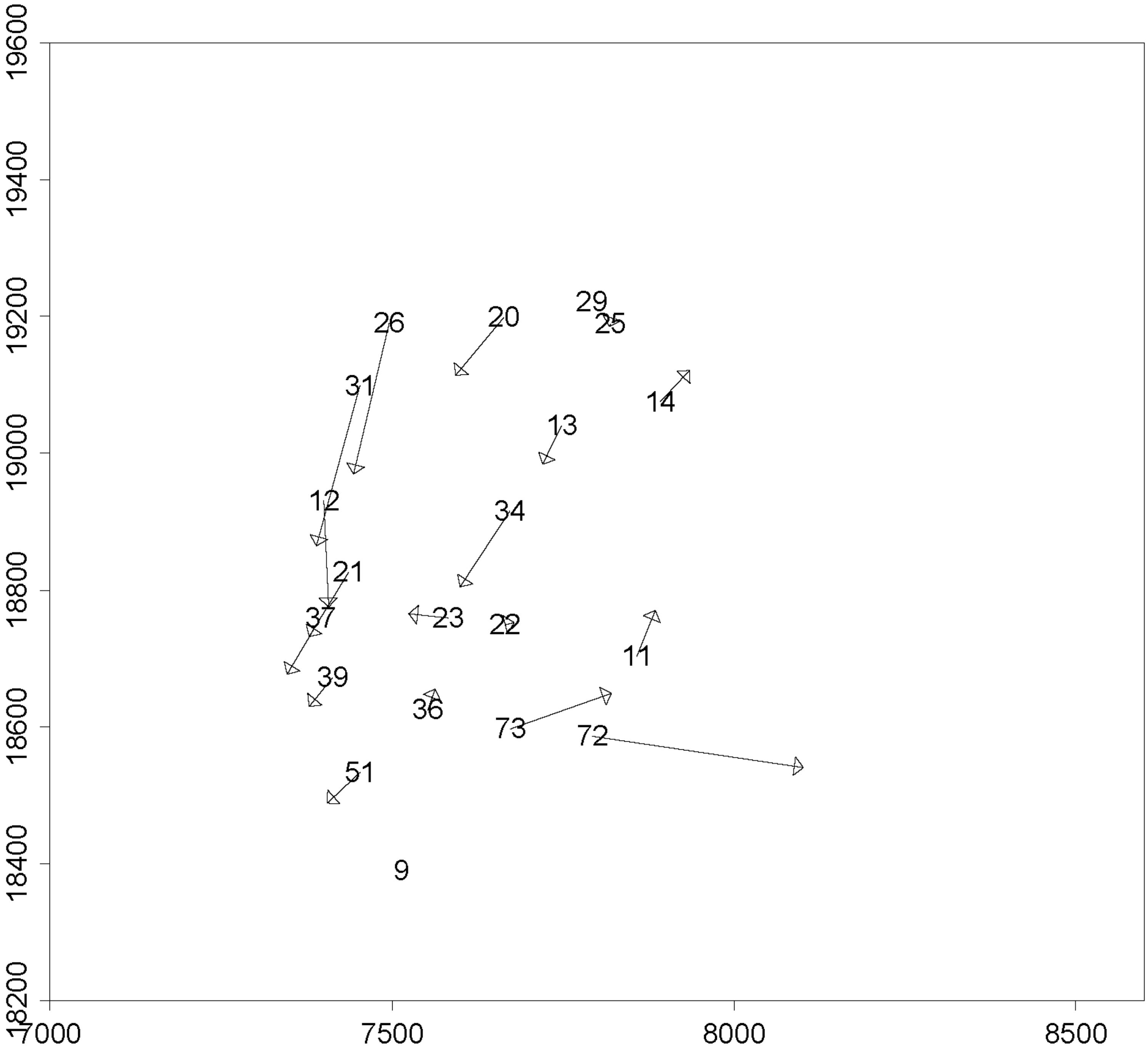
- Spatial dispersions are stationary and isotropic in D-space.

Network of Precipitation Monitoring Sites in France





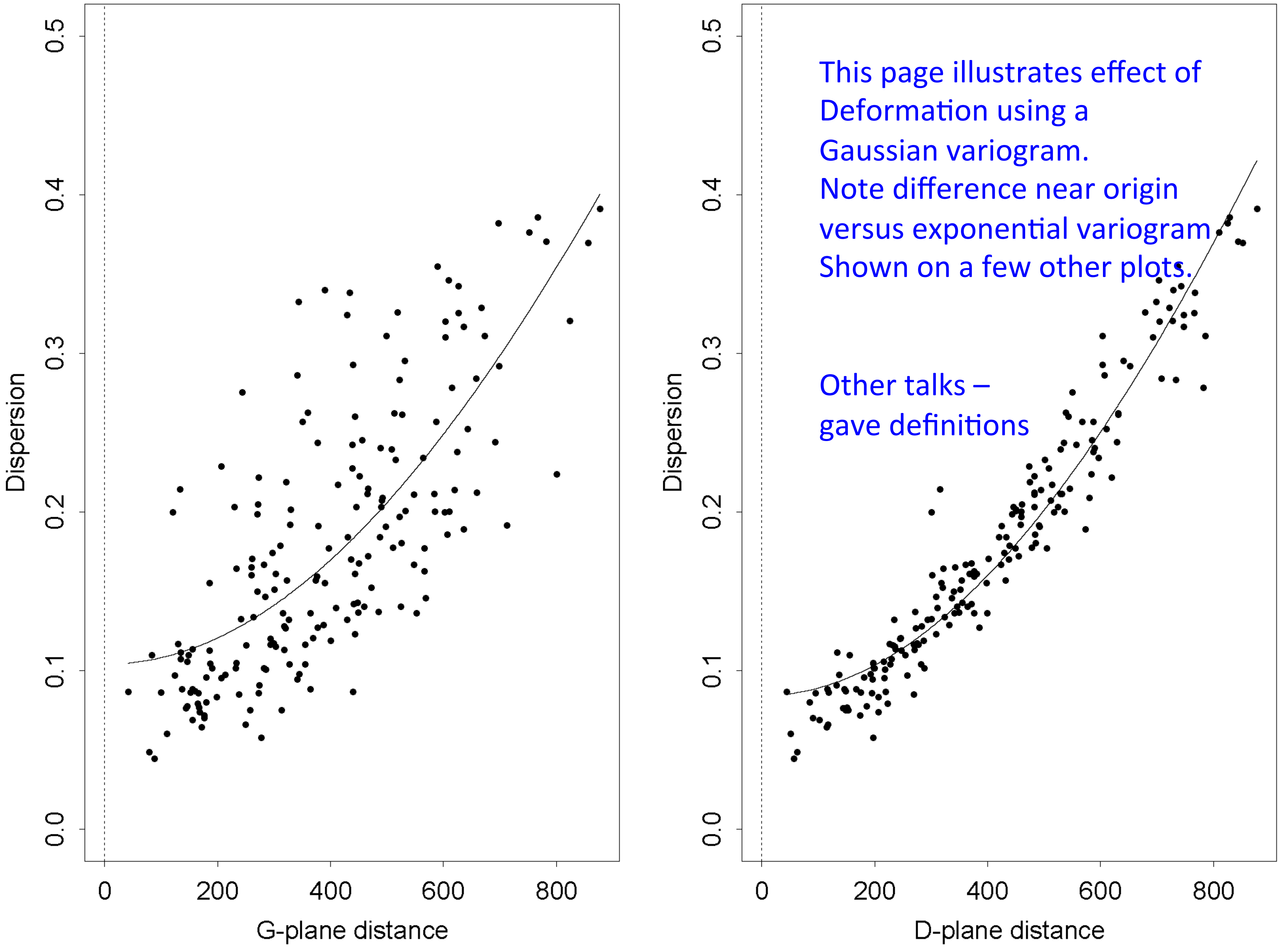
19600
19400
19200
19000
18800
18600
18400
18200
7000
7500
8000
8500
26
20
29
25
31
14
13
12
34
21
37
23
22
11
39
36
73
72
51
9



19600
19400
19200
19000
18800
18600
18400
18200
7000
7500
8000
8500
26
20
29
25
31
14
13
12
34
21
37
23
22
11
39
36
73
72
51
9

This page illustrates effect of Deformation using a Gaussian variogram.
Note difference near origin versus exponential variogram Shown on a few other plots.

Other talks – gave definitions



- Consider

$$Z(x,t) = \mu(x,t) + E_\tau(x) + E_\epsilon(x,t)$$

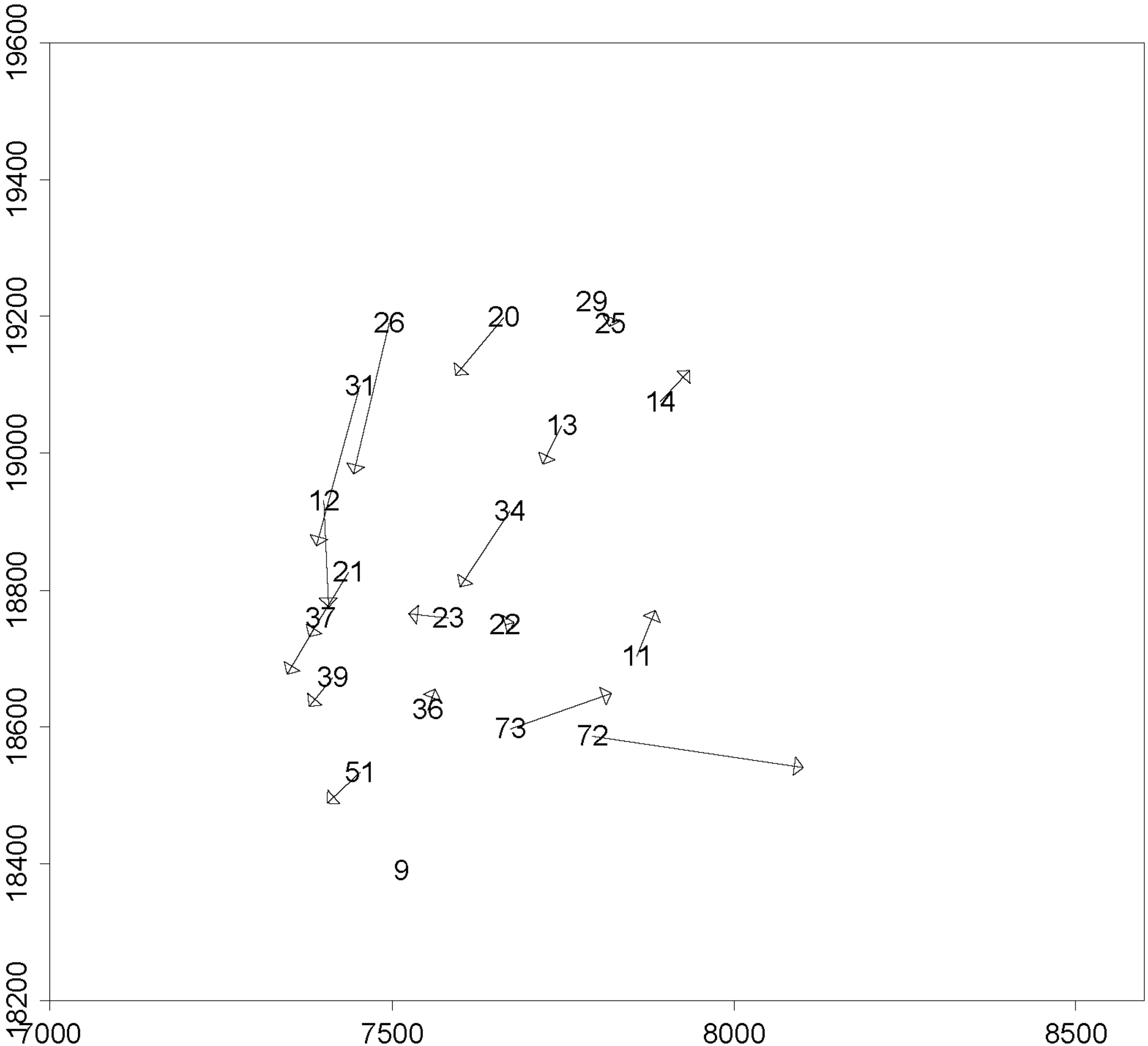
- Model

$$d_{ij} = \gamma_\theta\left(\left\| x_i^* - x_j^* \right\| \right) + e_{ij}$$

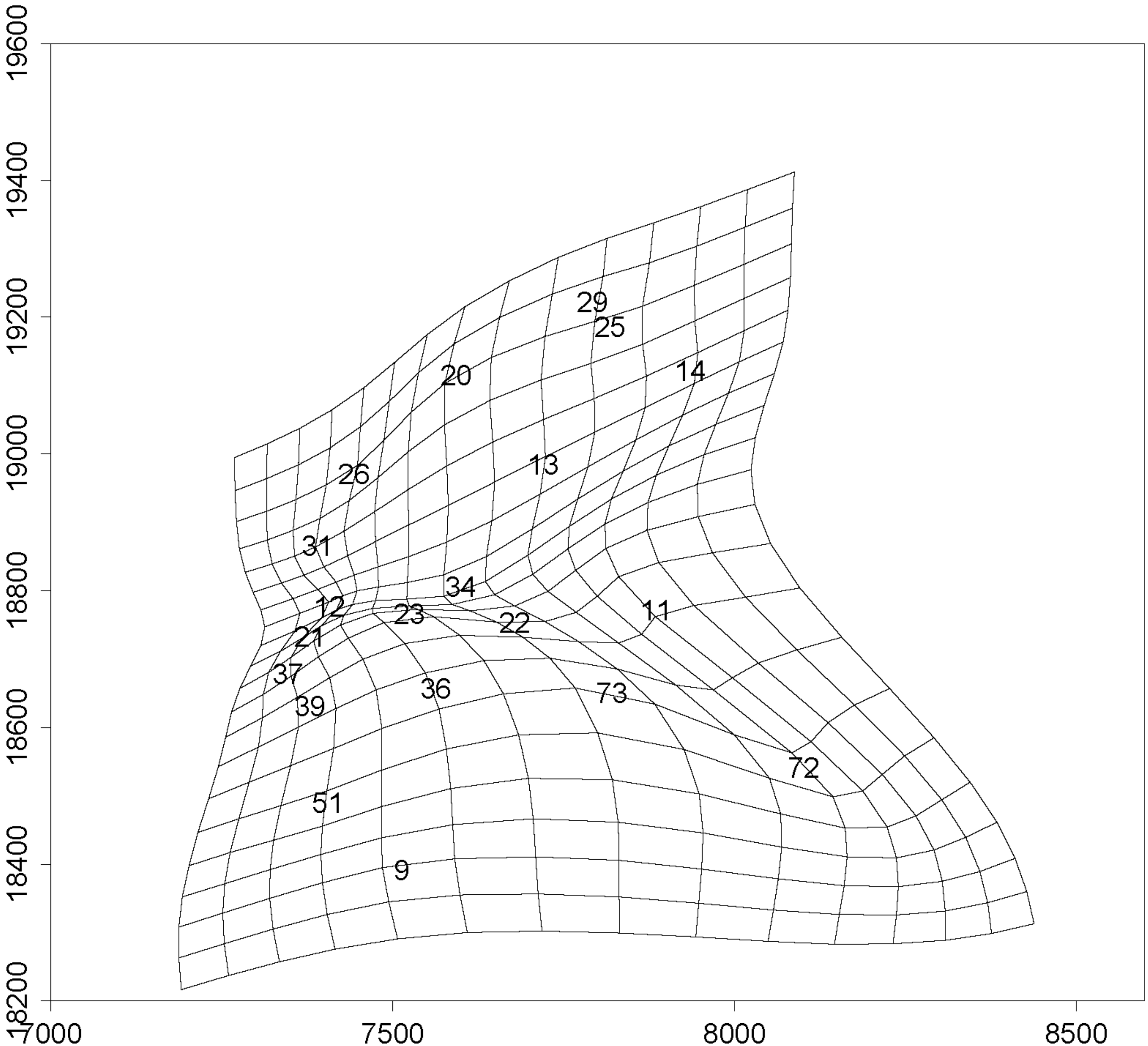
γ_θ isotropic variogram.

$$w: \underbrace{\mathbf{R}^p}_{\text{G-space}} \to \underbrace{\mathbf{R}^d}_{\text{D-space}}, \qquad x_i^* = w(x_i).$$

- Spatial dispersions are stationary and isotropic in D-space.



19600
19400
19200
19000
18800
18600
18400
18200
7000
7500
8000
8500
26
20
29
25
31
14
13
12
34
21
37
23
22
11
39
36
73
72
51
9


19600
19400
19200
19000
18800
18600
18400
18200
7000
7500
8000
8500
29
25
20
14
26
13
31
34
12
23
11
22
21
37
36
73
39
72
51
9

$\mathbf{R}^2 \to \mathbf{R}^2$ deformations.

$$w(x,y) = (f_1(x,y), f_2(x,y))$$

f_1: $\mathbf{R}^2 \to \mathbf{R}$ thin-plate spline mapping.
f_2: $\mathbf{R}^2 \to \mathbf{R}$ thin-plate spline mapping.

$$f_i(x,y) \;=\; a + bx + cy + \sum_{j=1}^{N} \eta_j\, h_j^2 \log\left(h_j^2\right)$$

where

$$h_j \;=\; \left\|\left(x_j, y_j\right) - (x\ , y)\right\|$$

Wahba (1991).
Spline models for observational data.

Estimation decisions, Including penalized weighted least squares optimization Criterion.

Other approaches: variants on Penalized likelihood (Paul Sampson's talk)

- Choose form of variogram γ_θ.
- Transformation map and variogram parameters chosen to minimize

$$\sum_{j=2}^{N} \sum_{i=1}^{j-1} \left[\frac{d_{ij} - \widehat{d_{ij}}}{\widehat{d_{ij}}} \right]^2 + \lambda \, \mathrm{BEP},$$

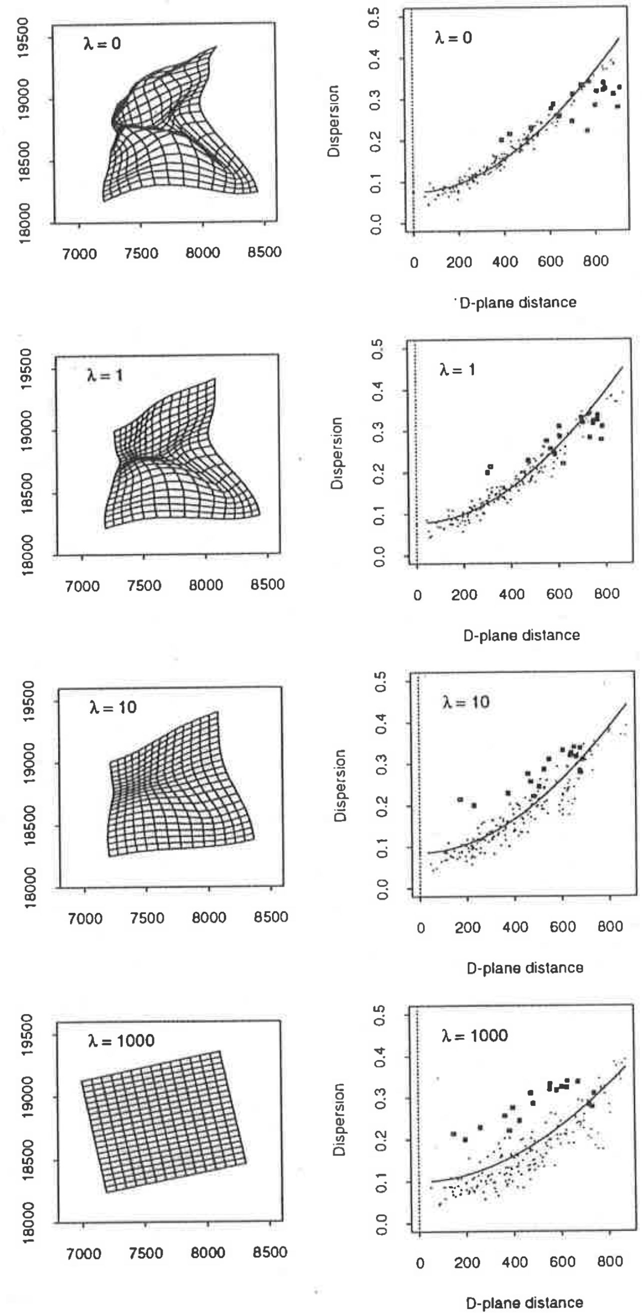
where $\widehat{d_{ij}} = \gamma_{\hat{\theta}}\left(\left\| x_i^* - x_j^* \right\| \right)$, and BEP is the sum of the bending energies for the two thin-plate spline mappings.

- Choice of λ.

  For $\lambda = 0$, overfitting, folding.

  As $\lambda \to \infty$, homogeneous model.

  Cross-validation (Monestiez et al. 1993).


λ = 0
λ = 1
λ = 10
λ = 1000
7000 7500 8000 8500
18000 18500 19000 19500
Dispersion
0.0 0.1 0.2 0.3 0.4 0.5
0 200 400 600 800
D-plane distance

$\mathbf{R}^2 \to \mathbf{R}^2$ deformations.

$$w(x,y) = (f_1(x,y), f_2(x,y))$$

f_i: $\mathbf{R}^2 \to \mathbf{R}$ thin-plate spline mapping,
$i = 1, 2.$

$$\mathbf{u} = (f_1(x_1), \ldots, f_1(x_N)),$$

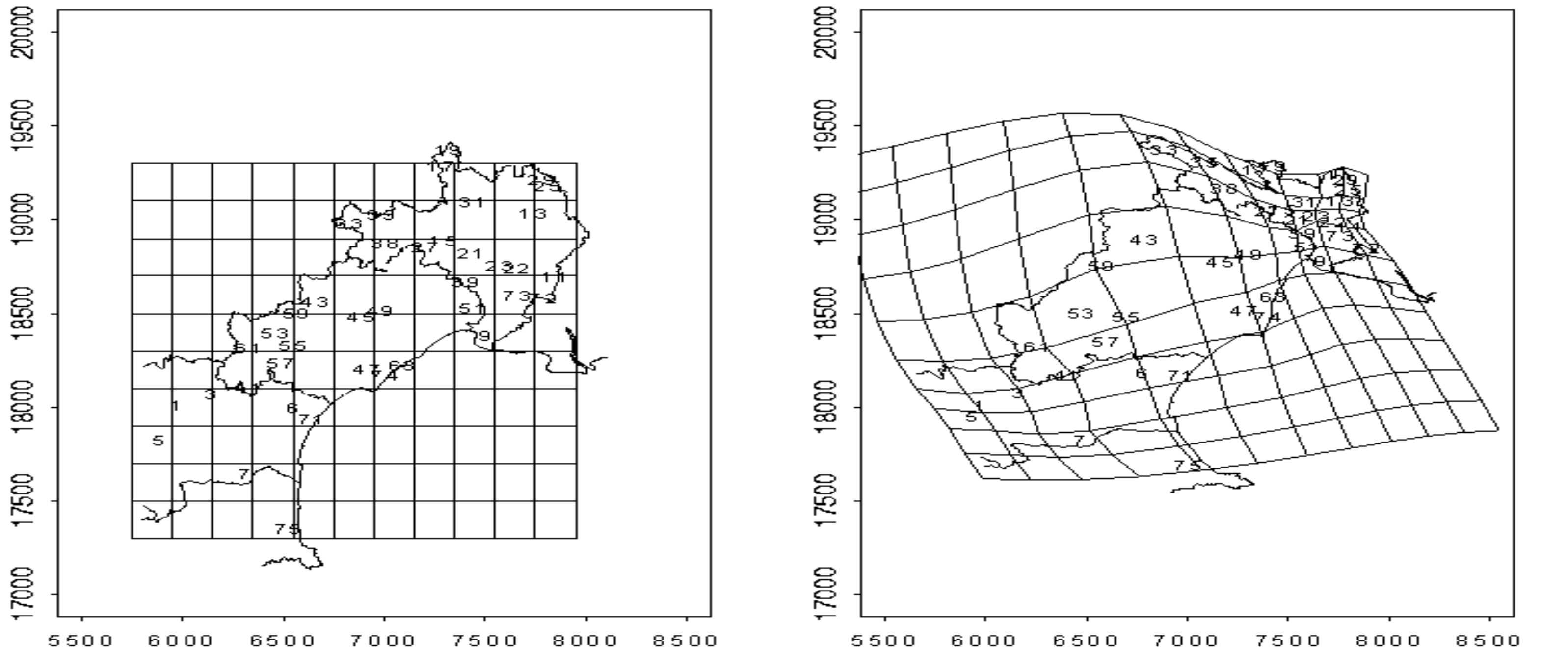
$$\mathbf{v} = (f_2(x_1), \ldots, f_2(x_N))$$

$$\text{BEP} = \mathbf{u}'\mathbf{B}\mathbf{u} + \mathbf{v}'\mathbf{B}\mathbf{v}$$

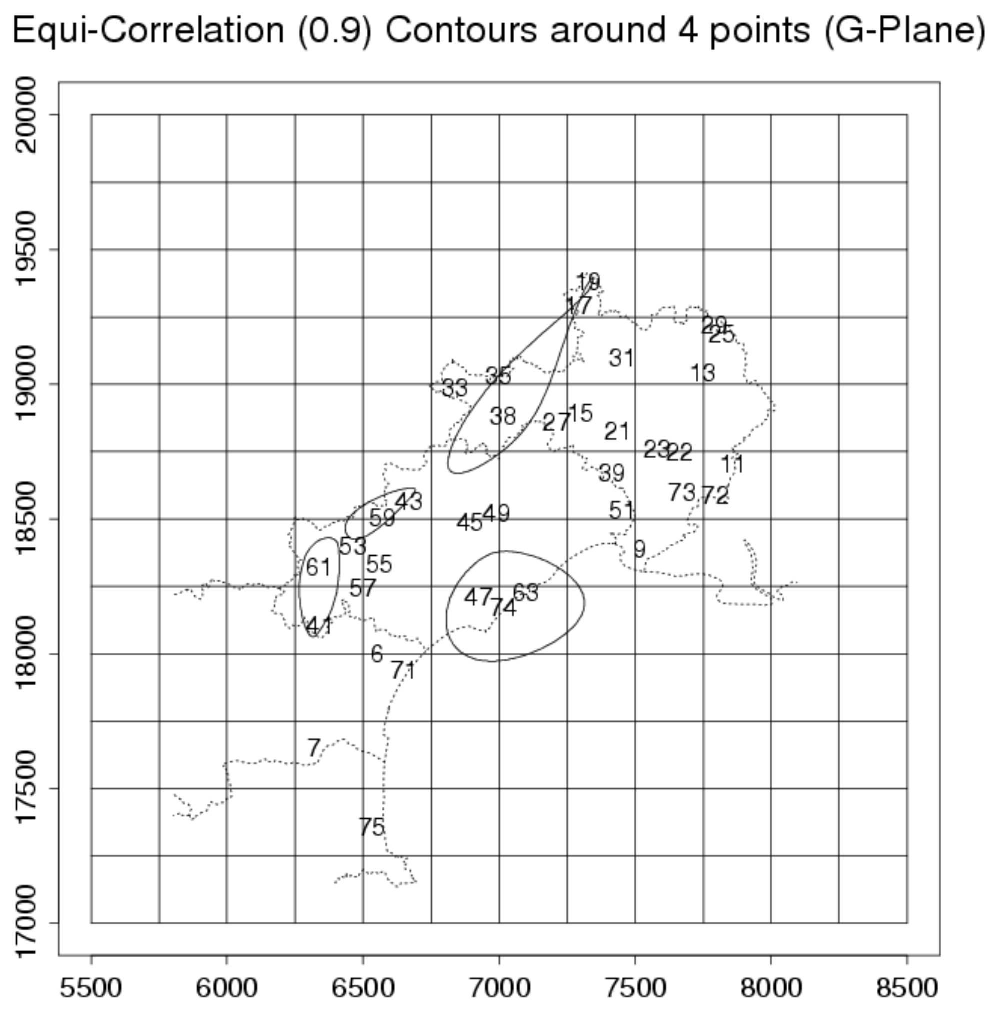
B depends only on geographic locations of monitoring sites.

$$\mathbf{u}'\mathbf{B}\mathbf{u} \propto \int_{\Re^2} \left[\left(\frac{\partial^2 f_1}{\partial x^2} \right) + 2 \left(\frac{\partial^2 f_1}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f_1}{\partial y^2} \right) \right] dx\, dy$$

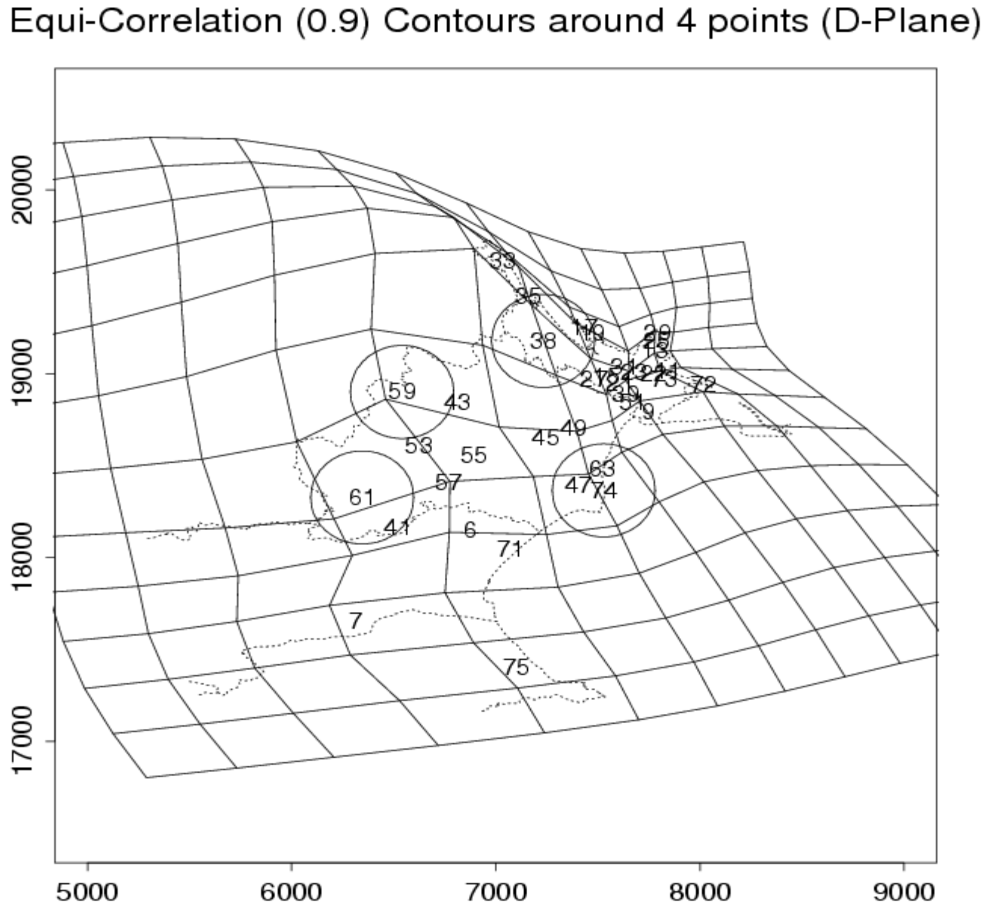
Fig. 7: Precipitation in Southern France -
an example of a non-linear deformation



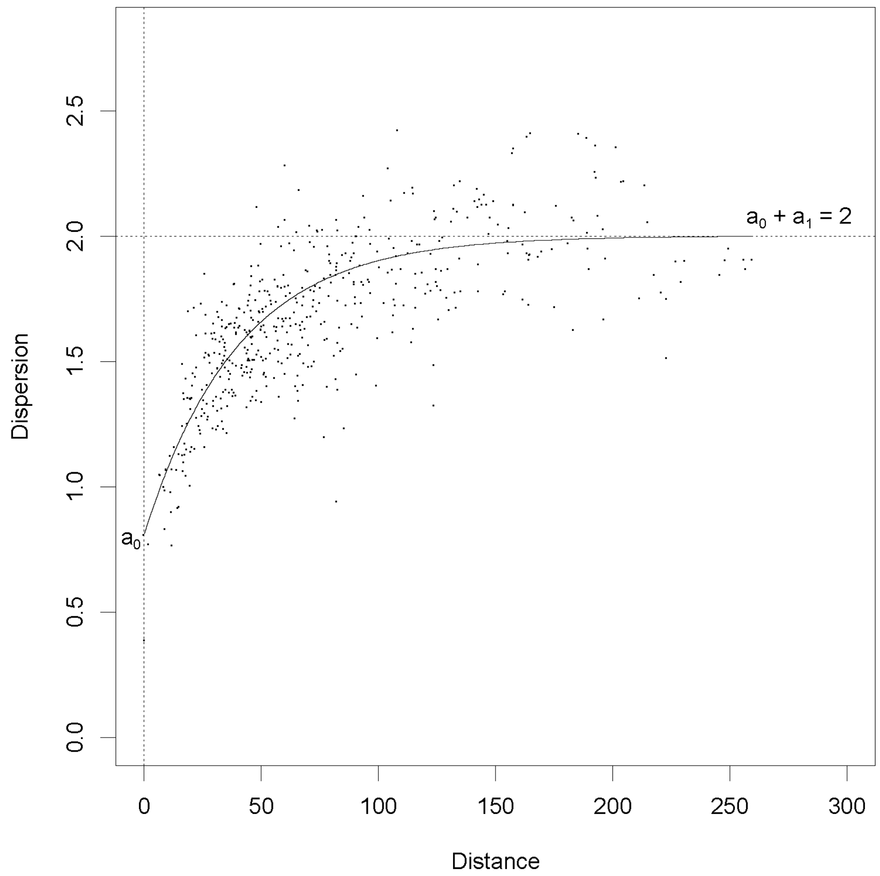
G-plane Equicorrelation Contours



D-plane Equicorrelation Contours



Exponential variogram - an example of an isotropic model for spatial association



$$D(x_i, x_j) = \begin{cases} 0 & \text{if } h_{ij} = 0 \\ a_0 + a_1 \{1 - \exp(-t_0\ h_{ij})\} & \text{if } h_{ij} > 0 \end{cases}$$

Cross-validation choice of λ

Omit each monitoring site x_i in turn.

For each λ, $\hat{w}_{i\lambda}$ and $\gamma_{\hat{\theta}_{i\lambda}}$ found.

λ chosen to minimize

$$\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}-\{i\}}\left(\frac{d_{ij}-\gamma_{\hat{\theta}_{i\lambda}}\left(\left\|\hat{w}_{i\lambda}(x_i)-\hat{w}_{i\lambda}(x_j)\right\|\right)}{\gamma_{\hat{\theta}_{i\lambda}}\left(\left\|\hat{w}_{i\lambda}(x_i)-\hat{w}_{i\lambda}(x_j)\right\|\right)}\right)^2$$

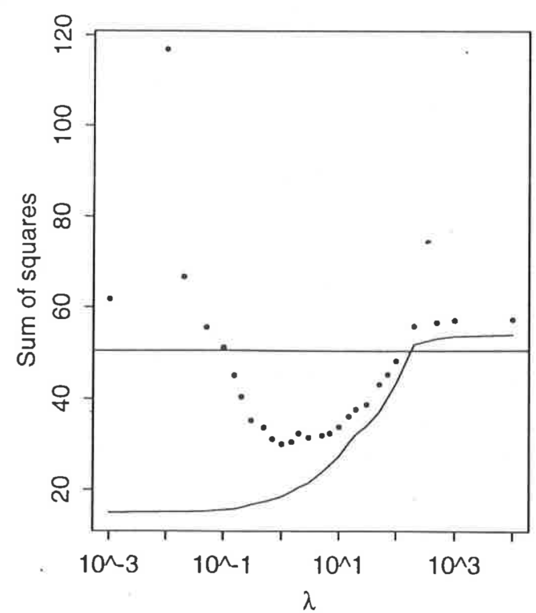


Figure based on a Simulation study.

Estimation decisions, Including penalized weighted least squares optimization Criterion.

Other approaches: variants on Penalized likelihood (Paul Sampson's talk)

EnviroStat – currently Uses penalized least squares

- Choose form of variogram γ_θ.
- Transformation map and variogram parameters chosen to minimize

$$\sum_{j=2}^{N} \sum_{i=1}^{j-1} \left[\frac{d_{ij} - \widehat{d_{ij}}}{\widehat{d_{ij}}} \right]^2 + \lambda \text{ BEP},$$

where $\widehat{d_{ij}} = \gamma_{\hat{\theta}}\left(\left\| x_i^* - x_j^* \right\| \right)$, and BEP is the sum of the bending energies for the two thin-plate spline mappings.

- Choice of λ.

For $\lambda = 0$, overfitting, folding.

As $\lambda \to \infty$, homogeneous model.

Cross-validation (Monestiez et al. 1993).

Optimization

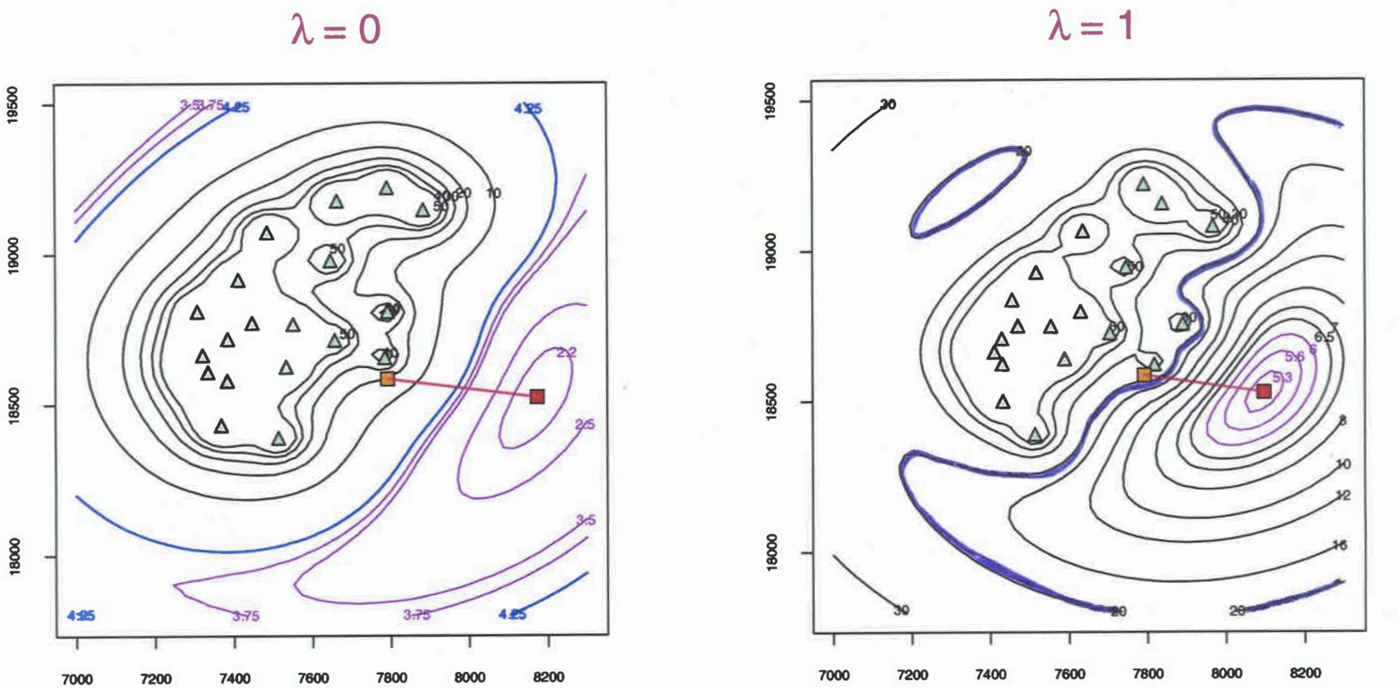
- **Alternating algorithm:**
  - Parameter estimation for D-space variogram
  - Estimation of D-space locations of monitoring sites

  N sites $\Longrightarrow$ **2N-4 dimensions**

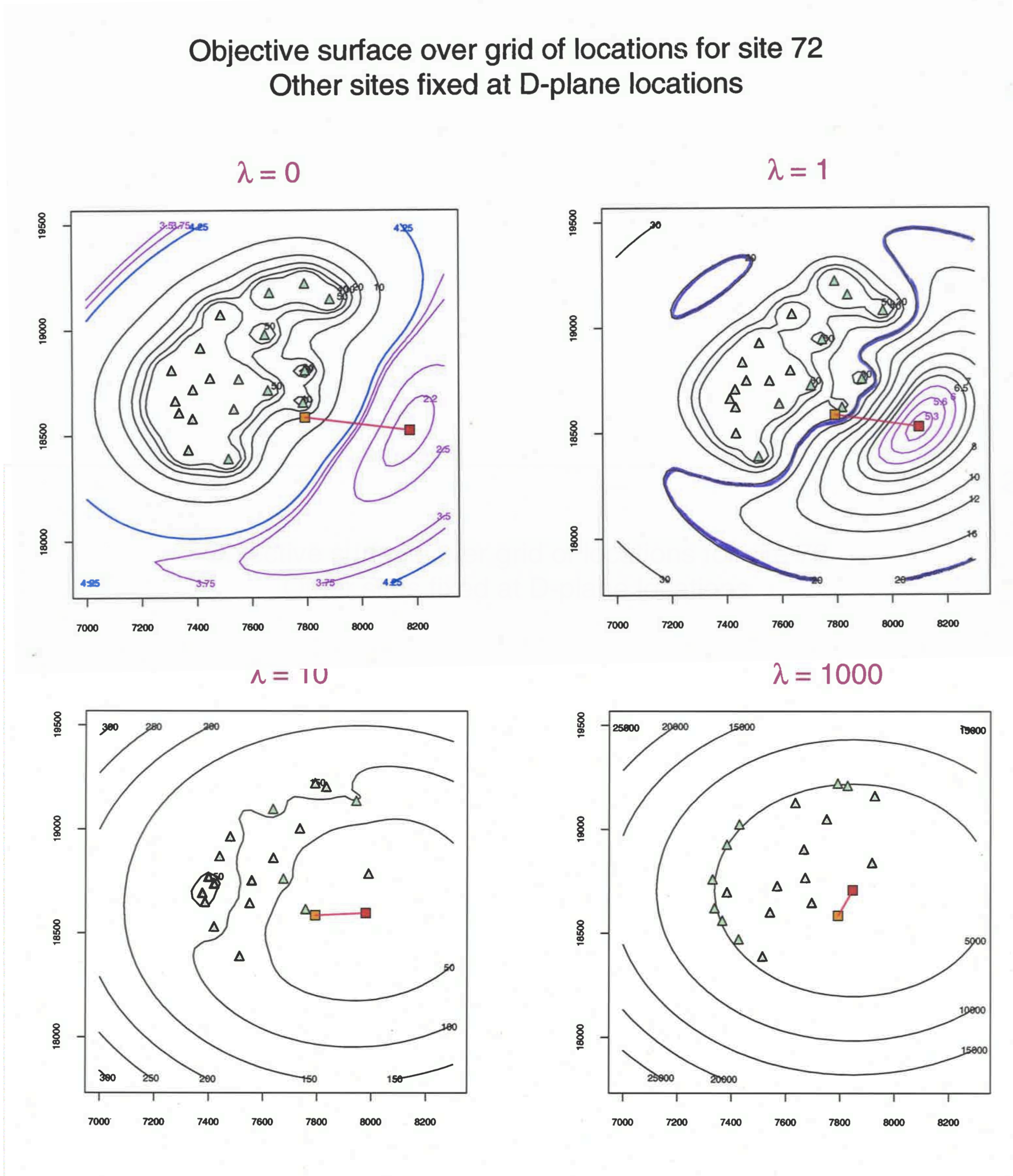
- **Scaling of coordinates**
  - Calculation of BEP
  - **Similar scales for optimization over variogram parameters**

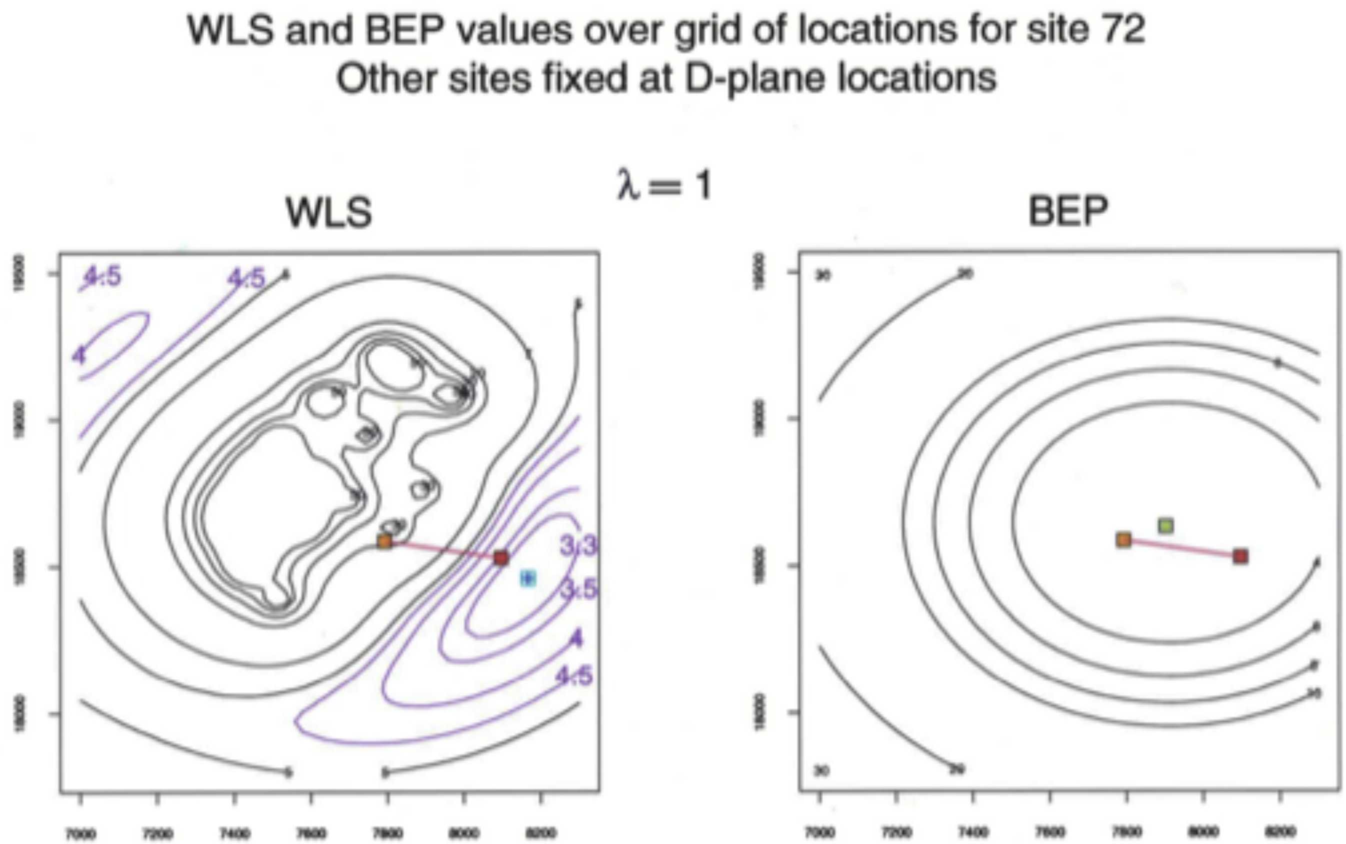
- **Starting values**
  - Locations (start at geographic locations)
  - Variogram parameters

Objective surface over grid of locations for site 72
Other sites fixed at D-plane locations



Complexity of the
Optimization
Surface decreases
As λ increases





Orange - G-plane location

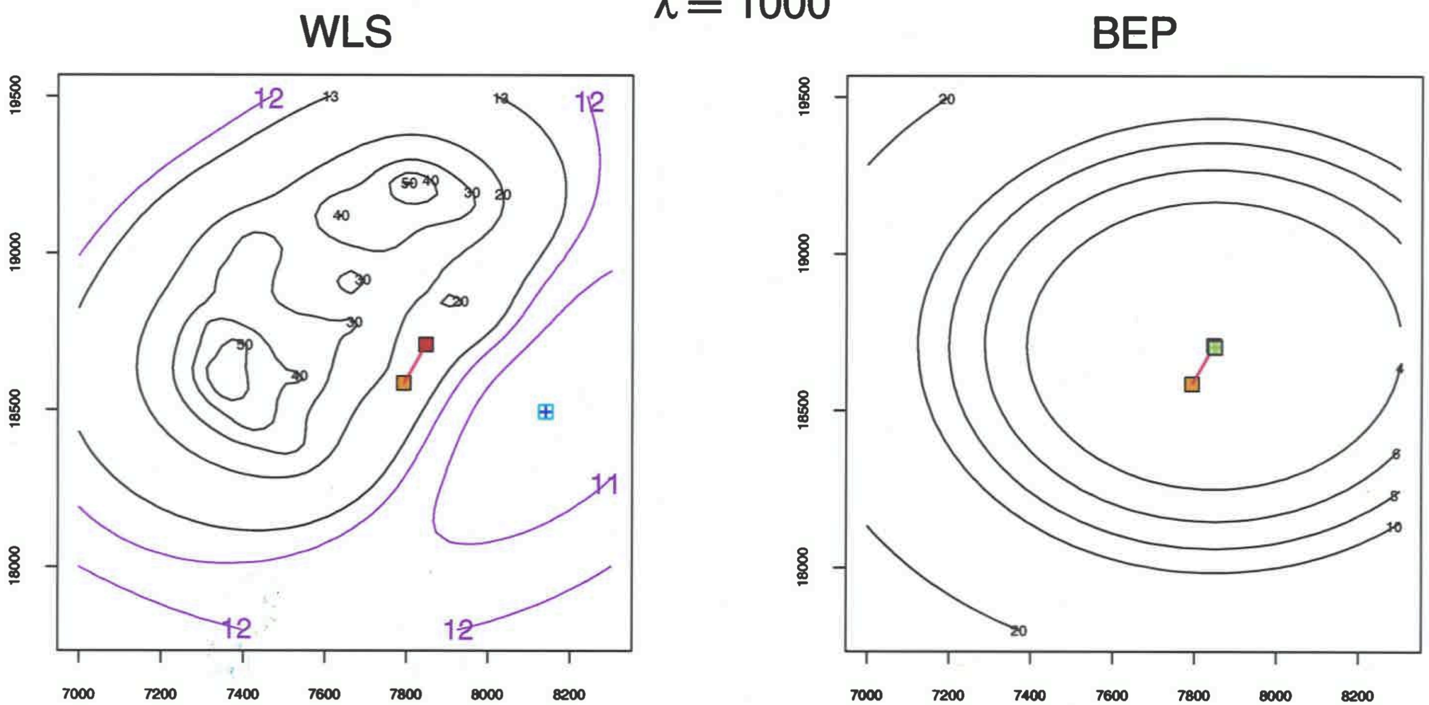
Red - D-plane location

Blue – minimizes WLS term (left)

Green – minimizes BEP term (right)

WLS and BEP values over grid of locations for site 72
Other sites fixed at D-plane locations

$\lambda = 1000$



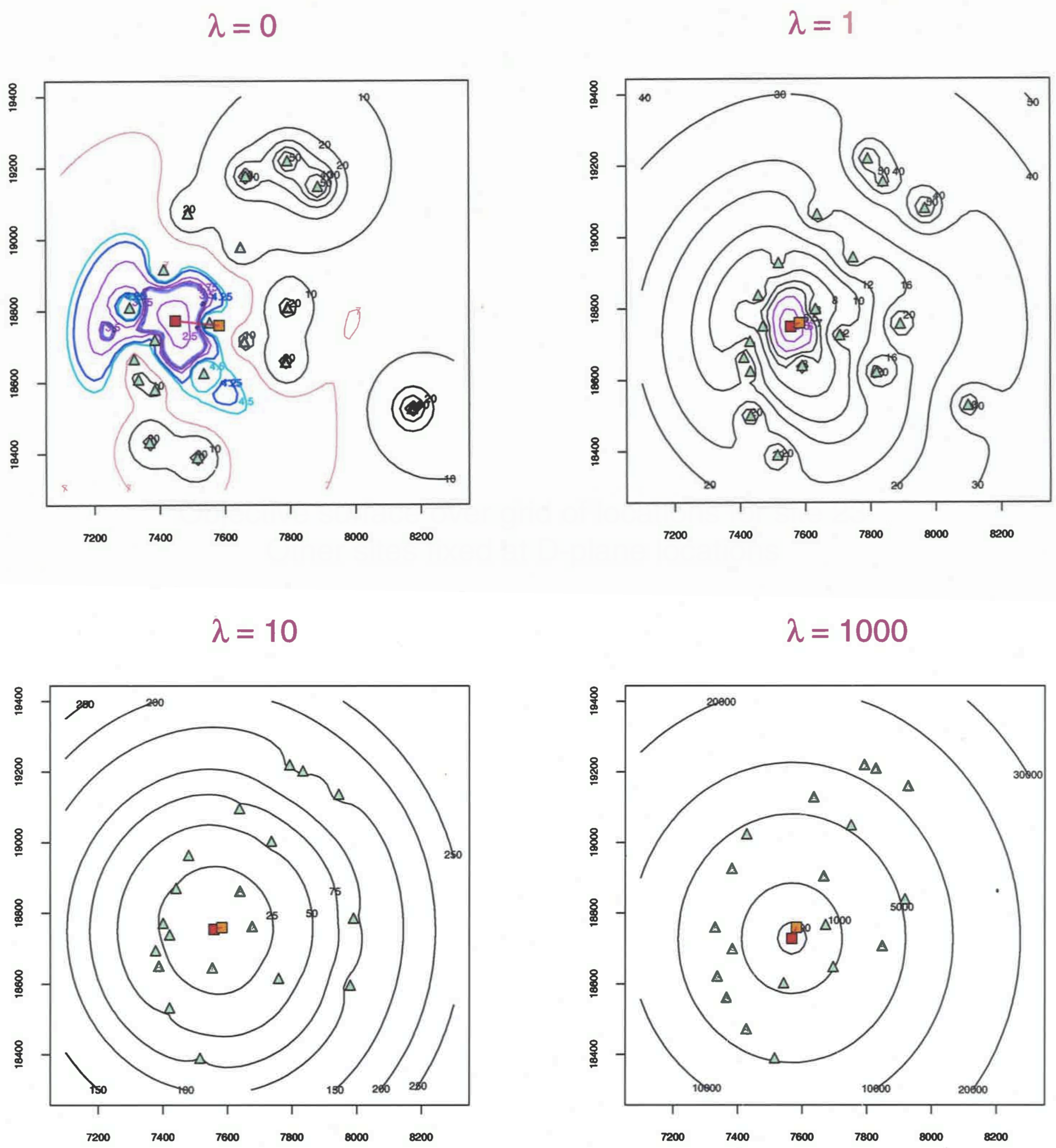
Orange - G-plane location

Red - D-plane location

Blue – minimizes WLS term (left)

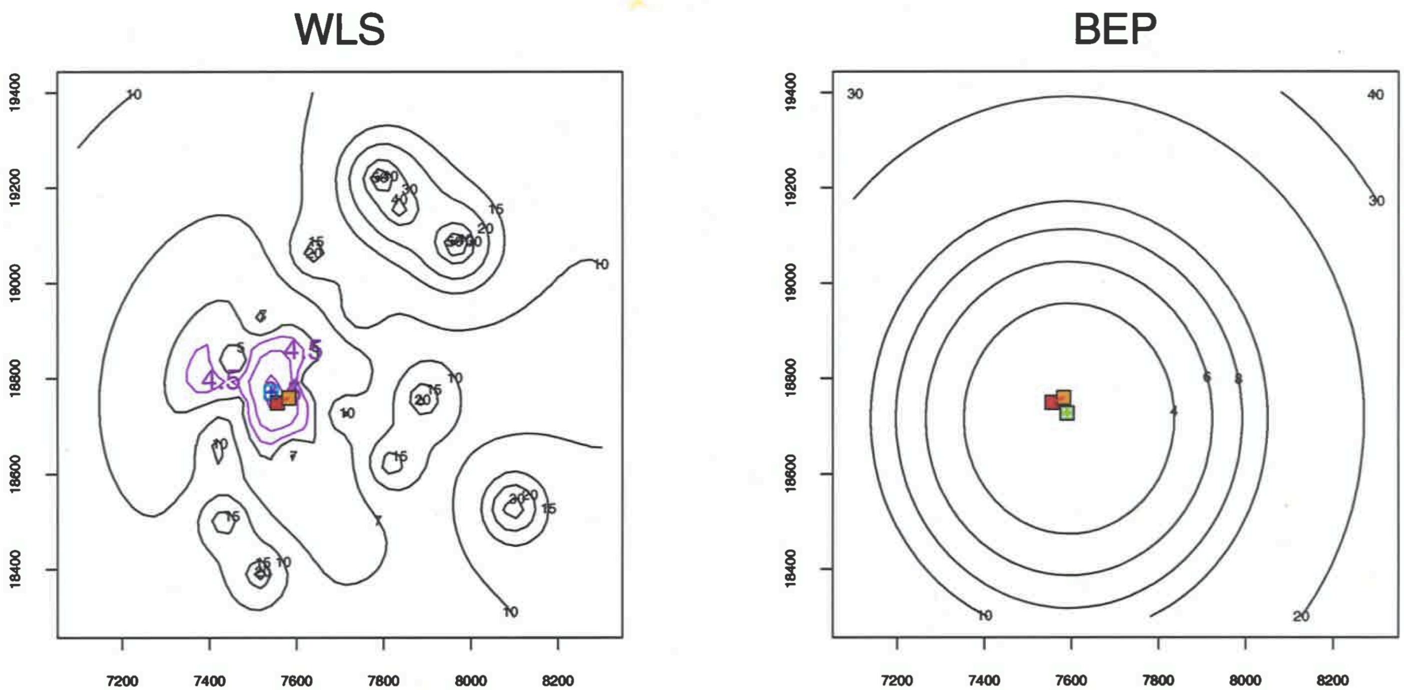
Green – minimizes BEP term (right)

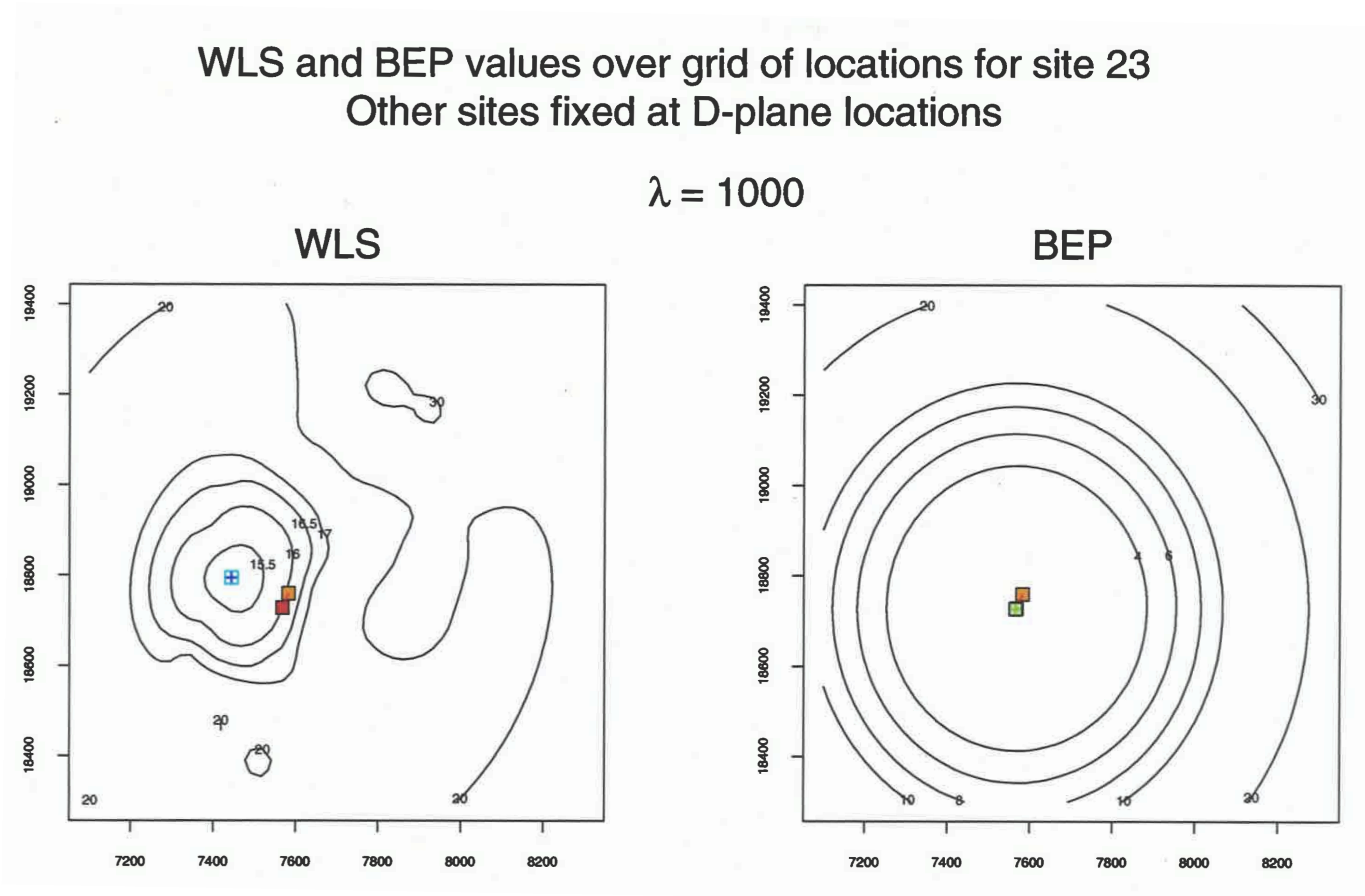
Objective surface over grid of locations for site 23
Other sites fixed at D-plane locations



WLS and BEP values over grid of locations for site 23
Other sites fixed at D-plane locations

$\lambda = 1$





Complexity of the optimization surface decreases as λ increases

Universal Kriging:
(here using s as spatial location instead of x)
Suppose

$$
\begin{aligned}
Z(\boldsymbol{s}) &= \sum_{j=1}^{p+1} f_{j-1}(\boldsymbol{s})\beta_{j-1} + \delta(\boldsymbol{s}), \qquad \boldsymbol{s} \in D. \\
\text{i.e.,} \quad \boldsymbol{Z} &= F\boldsymbol{\beta} + \boldsymbol{\delta}
\end{aligned}
$$

where

$$
\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}(\boldsymbol{s}_1) \\ \vdots \\ \boldsymbol{Z}(\boldsymbol{s}_n) \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \text{and}
$$

$$
F = \begin{bmatrix} f_0(\boldsymbol{s}_1) & f_1(\boldsymbol{s}_1) & \dots & f_p(\boldsymbol{s}_1) \\ f_0(\boldsymbol{s}_2) & f_1(\boldsymbol{s}_2) & \dots & f_p(\boldsymbol{s}_2) \\ & & \vdots & \\ f_0(\boldsymbol{s}_n) & f_1(\boldsymbol{s}_n) & \dots & f_p(\boldsymbol{s}_n) \end{bmatrix}_{n\times(p+1)} .
$$

Known functions $\{f_0(\boldsymbol{s}), \dots, f_p(\boldsymbol{s})\}$.
Unknown parameters: $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)' \in \Re^{p+1}$.
$\delta(.)$ is a spatial residual process.

Aim: Predict $Z(\boldsymbol{s}_0) = \boldsymbol{f}'\boldsymbol{\beta} + \delta(\boldsymbol{s}_0)$ where
$\boldsymbol{f} = (f_0(\boldsymbol{s}_0), \dots, f_p(\boldsymbol{s}_0))'$.

Special case: Second order polynomial trend in $\Re^2$:

For $\boldsymbol{s} = (x, y)$,

$$\mu(\boldsymbol{s}) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2$$

$$\begin{aligned} f_0(\boldsymbol{s}) &= 1 \\ f_1(\boldsymbol{s}) &= x \\ f_2(\boldsymbol{s}) &= y \\ f_3(\boldsymbol{s}) &= x^2 \\ f_4(\boldsymbol{s}) &= xy \\ f_5(\boldsymbol{s}) &= y^2 \end{aligned}$$

Minimize the mean square prediction error, subject to the uniform unbiasedness constraint.

mse

$$\begin{aligned}
&= E\left[\left(p\left(Z,\boldsymbol{s}_0\right)-Z\left(\boldsymbol{s}_0\right)\right)^2\right] \\
&= \mathrm{Var}\left(p\left(Z,\boldsymbol{s}_0\right)-Z\left(\boldsymbol{s}_0\right)\right)+\left[E\left(p\left(Z,\boldsymbol{s}_0\right)-Z\left(\boldsymbol{s}_0\right)\right)\right]^2. \\
&= \mathrm{Var}\left[\sum_{i=1}^{n}\lambda_i Z\left(\boldsymbol{s}_i\right)-Z\left(\boldsymbol{s}_0\right)\right] \\
&\quad +\left[E\left(\sum_{i=1}^{n}\lambda_i Z\left(\boldsymbol{s}_i\right)-Z\left(\boldsymbol{s}_0\right)\right)\right]^2
\end{aligned}$$

Minimizing the mse subject to the unbiasedness constraint is equivalent to minimizing

$$\begin{aligned}
Q &= E\left[\left(p\left(Z,\boldsymbol{s}_0\right)-Z\left(\boldsymbol{s}_0\right)\right)^2\right] \\
&\quad -2\sum_{j=1}^{p+1} m_{j-1}\left\{\sum_{i=1}^{n}\lambda_i f_{j-1}\left(\boldsymbol{s}_i\right)-f_{j-1}\left(\boldsymbol{s}_0\right)\right\}
\end{aligned}$$

with respect to $\lambda_1,\ldots,\lambda_n$, and $m_0,\ldots,m_p$

$$\begin{aligned}
\boldsymbol{\lambda}' &= \left\{\boldsymbol{c}+F(F'\Sigma^{-1}F)^{-1}\left(\boldsymbol{f}-F'\Sigma^{-1}\boldsymbol{c}\right)\right\}'\Sigma^{-1} \\
\boldsymbol{m}' &= \left(\boldsymbol{f}-F'\Sigma^{-1}\boldsymbol{c}\right)'\left(F'\Sigma^{-1}F\right)^{-1}
\end{aligned}$$

Universal kriging

Suppose $$Z(x) = \sum_{j=1}^{p+1} f_{j-1}(x)\beta_{j-1} + \delta(x),$$

for x in the region of interest D.

Assume $\{f_0(x), f_1(x), \ldots, f_p(x)\}$ are known functions of x, or of a related quantity at site x.

Suppose $\{\beta_0, \ldots, \beta_p\}$ are unknown,

$E[\delta(x)] = 0$ and $\delta(.)$ has known variogram $2\gamma(.)$.

Then there is a unique estimator of the form

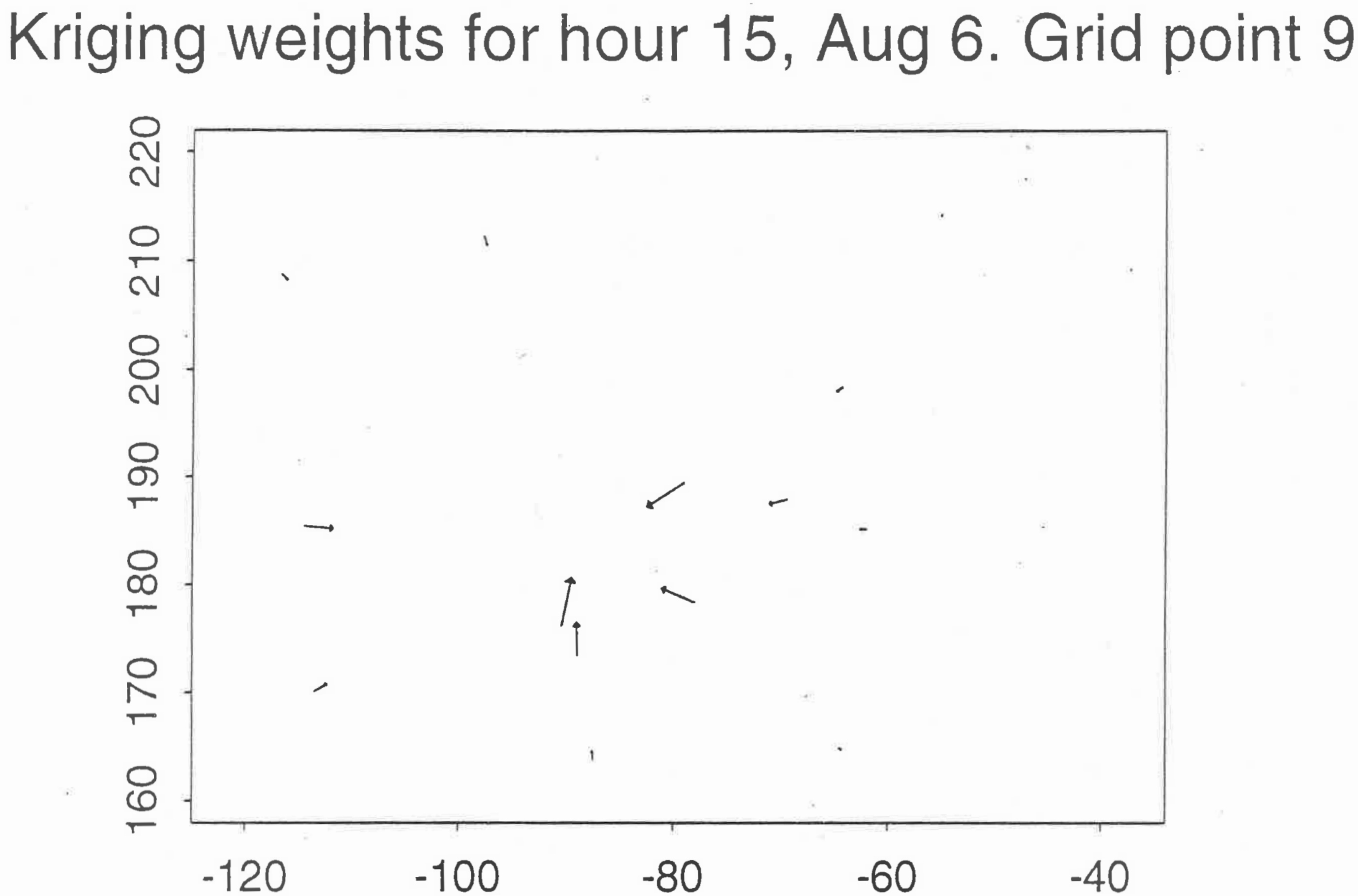
$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i)$$

that minimizes the mean-square prediction error subject to an unbiasedness constraint on $\{\lambda_1, \ldots, \lambda_n\}$.

Last minute addition given other talks

Special case: $p = 0$ and $f_0(s) = 1$.

Ordinary Kriging Weights Illustration – although for different data



Spatial prediction via universal kriging

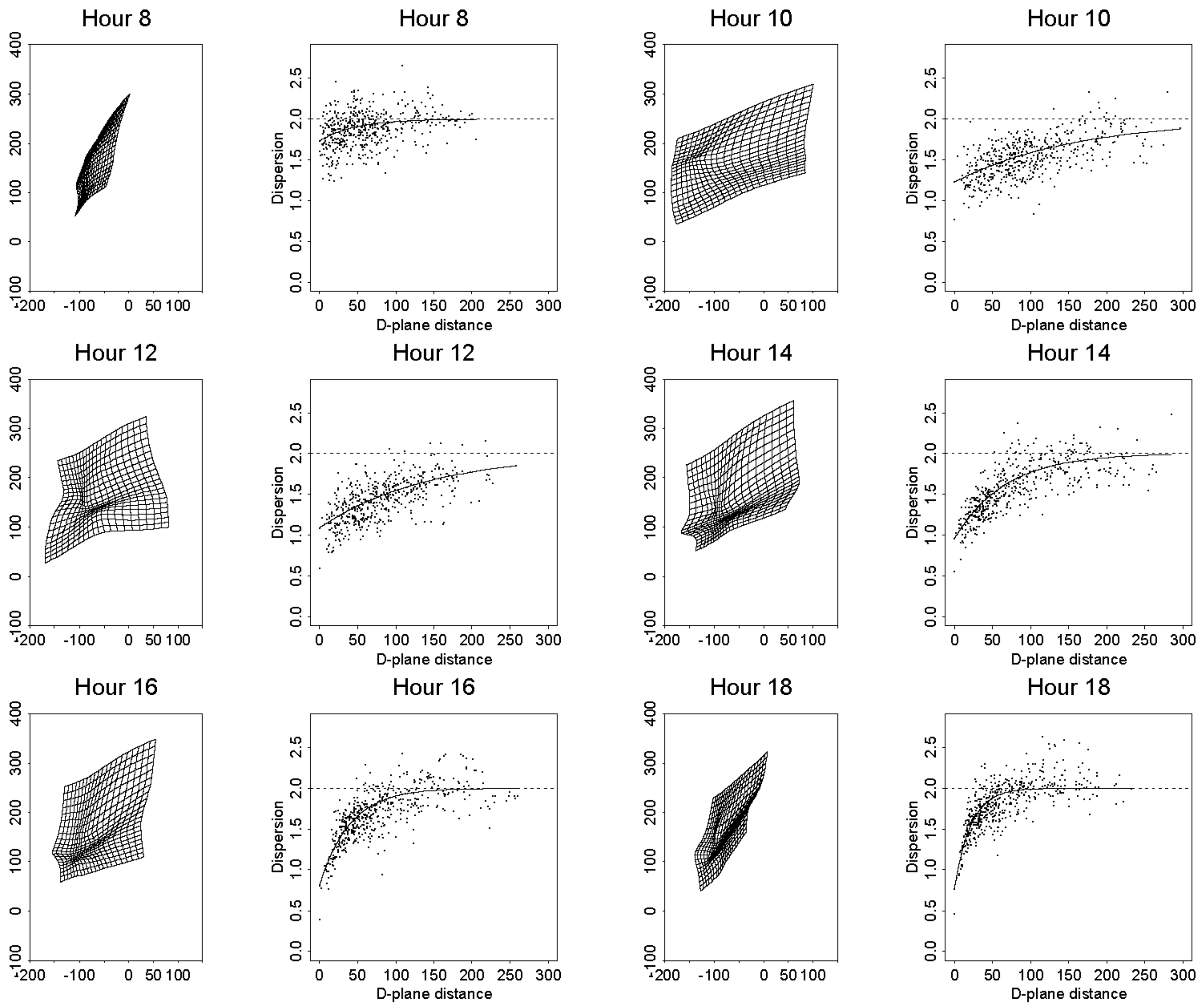
- Optimality results assume that the form and parameters of the spatial correlation function are known.

- **In practice, both the form and parameters are unknown.**

  - use Bayesian Kriging;

    or a bootstrap approach.

Space and Time Often Are Non-Separable. Ozone Example
Spatial Structure in Pre-whitened Residuals – varies by hour of day



Points for discussion/brain-storming

Extensions to space-time correlation.

Diagnostics for data that are different in their space-time second order structure from their neighbors – quality control (eg. Ozone eg. Guttorp et al., 1994)

Step 1 - Heterogeneous estimation approach can help identify and understand nature of anisotropy in spatial (or space-time) correlation, including changes in time such as seasonal.

Always motivates further study:
Is structure real?
Are there scientific explanations for these anisotropies?
Alexandra – including covariates in covariances.