Spatial Extremes Analyses in Climate Studies

- **P. Naveau**

  Laboratoire des Sciences du Climat et de l'Environnement

  IPSL, CNRS, France

- **D. Cooley**

  Applied Math Dept, Colorado University USA

- **P. Poncet**

  Université Paris X, France

naveau@lsce.cnrs-gif.fr

Outline of the Talk

1. Motivations

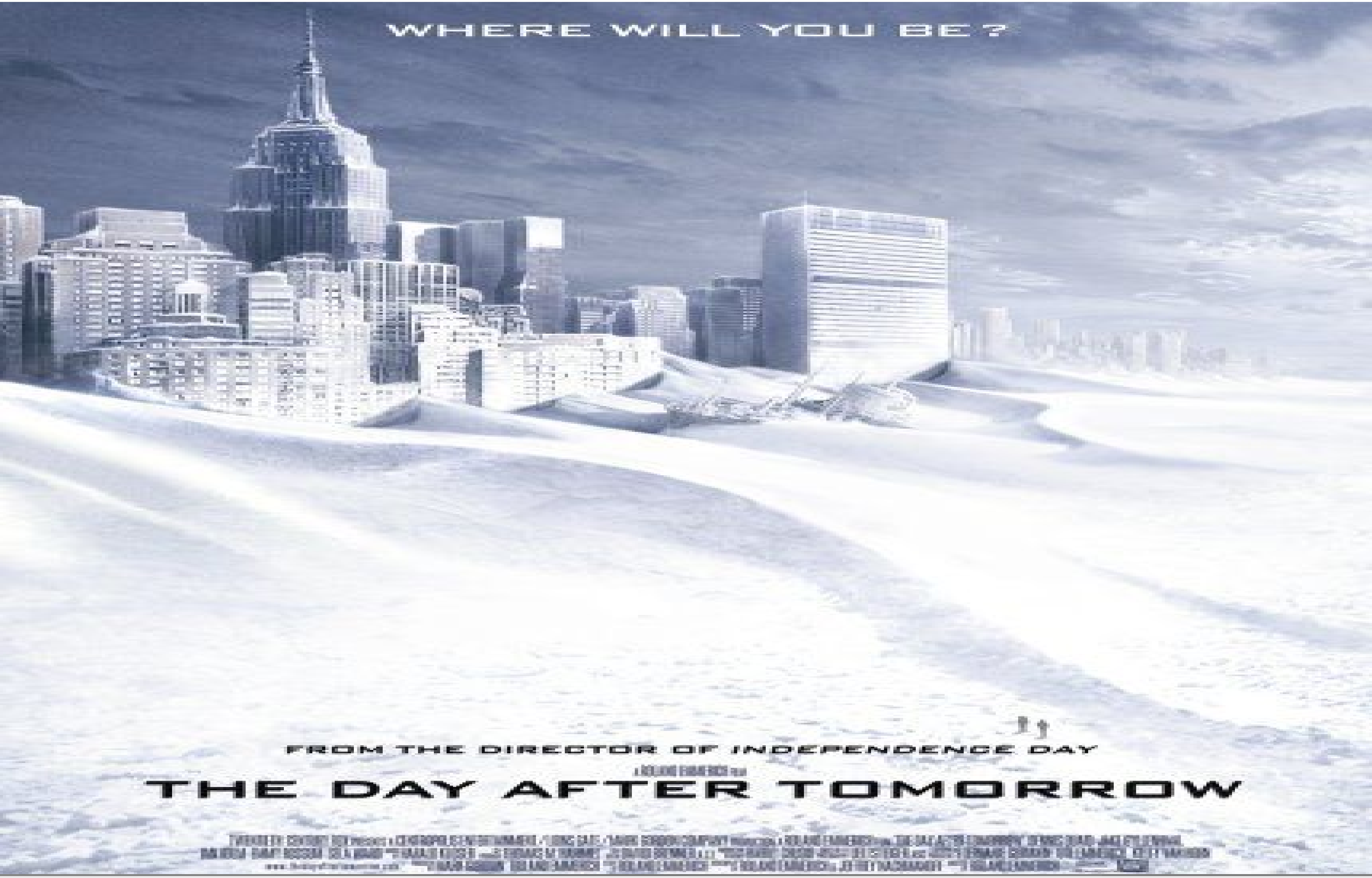
2. Maxima distribution

3. Extremal coefficient

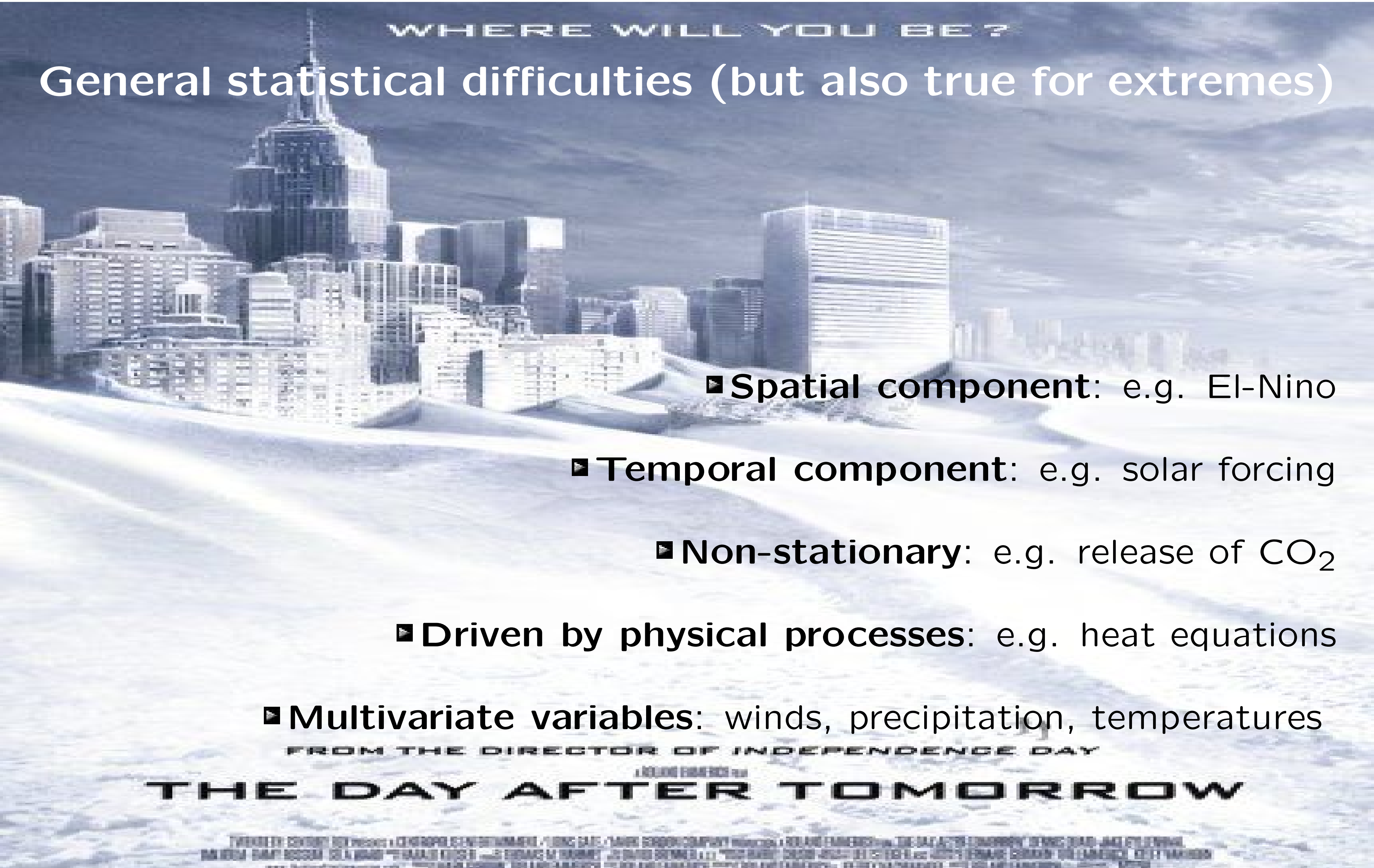
4. Geostatistics

5. Conclusions

Climate Studies



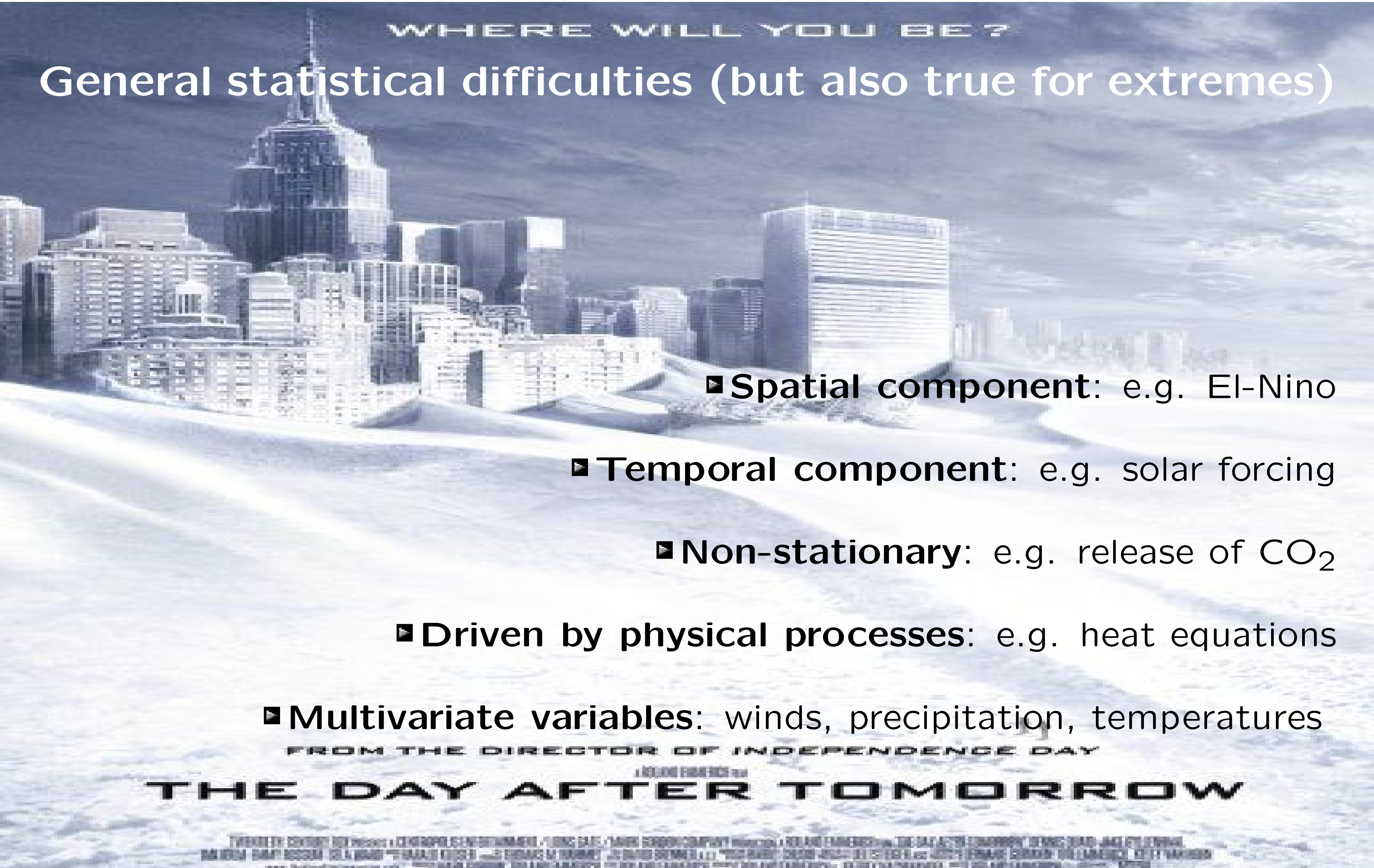
Climate Studies



General statistical difficulties (but also true for extremes)

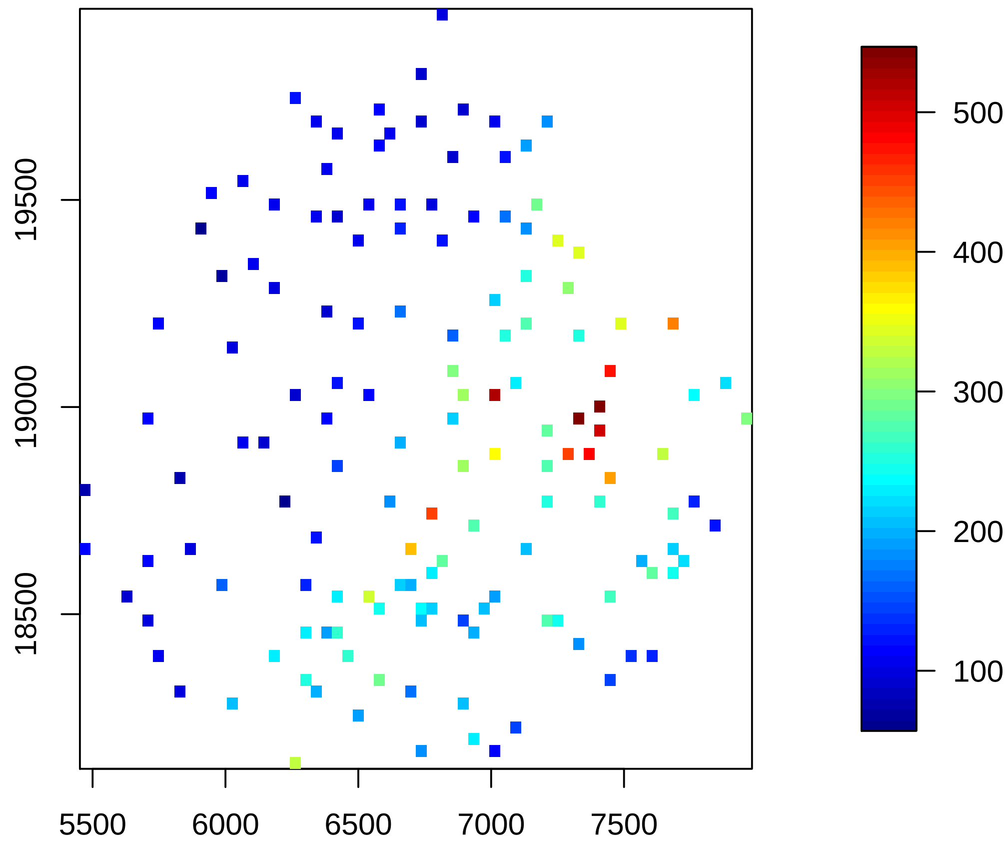
- **Spatial component**: e.g. El-Nino
- **Temporal component**: e.g. solar forcing
- **Non-stationary**: e.g. release of CO_2
- **Driven by physical processes**: e.g. heat equations
- **Multivariate variables**: winds, precipitation, temperatures

Climate Studies



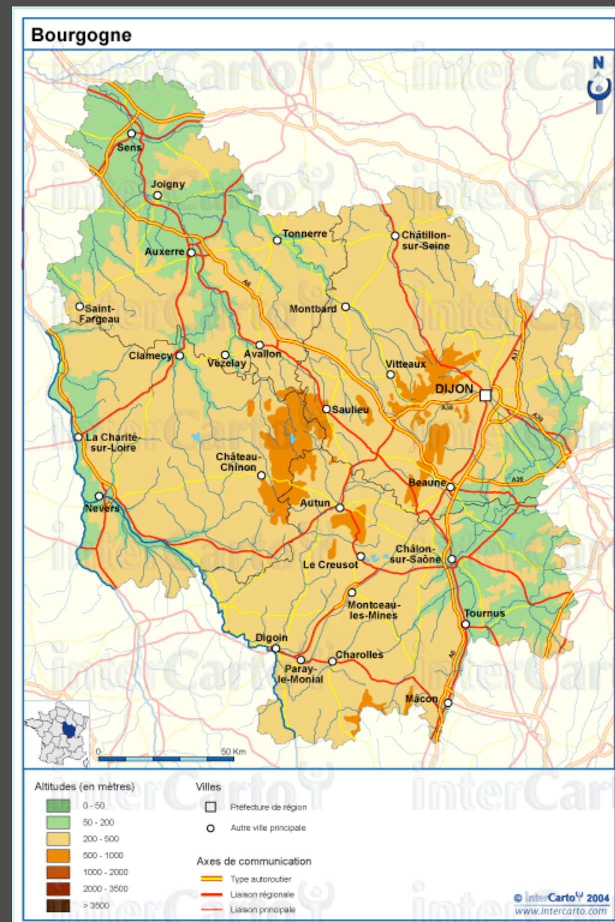
General statistical difficulties (but also true for extremes)

- **Spatial component**: e.g. El-Nino
- **Temporal component**: e.g. solar forcing
- **Non-stationary**: e.g. release of CO_2
- **Driven by physical processes**: e.g. heat equations
- **Multivariate variables**: winds, precipitation, temperatures



How to describe the **spatial dependence** as a function of the **distance** between two points?

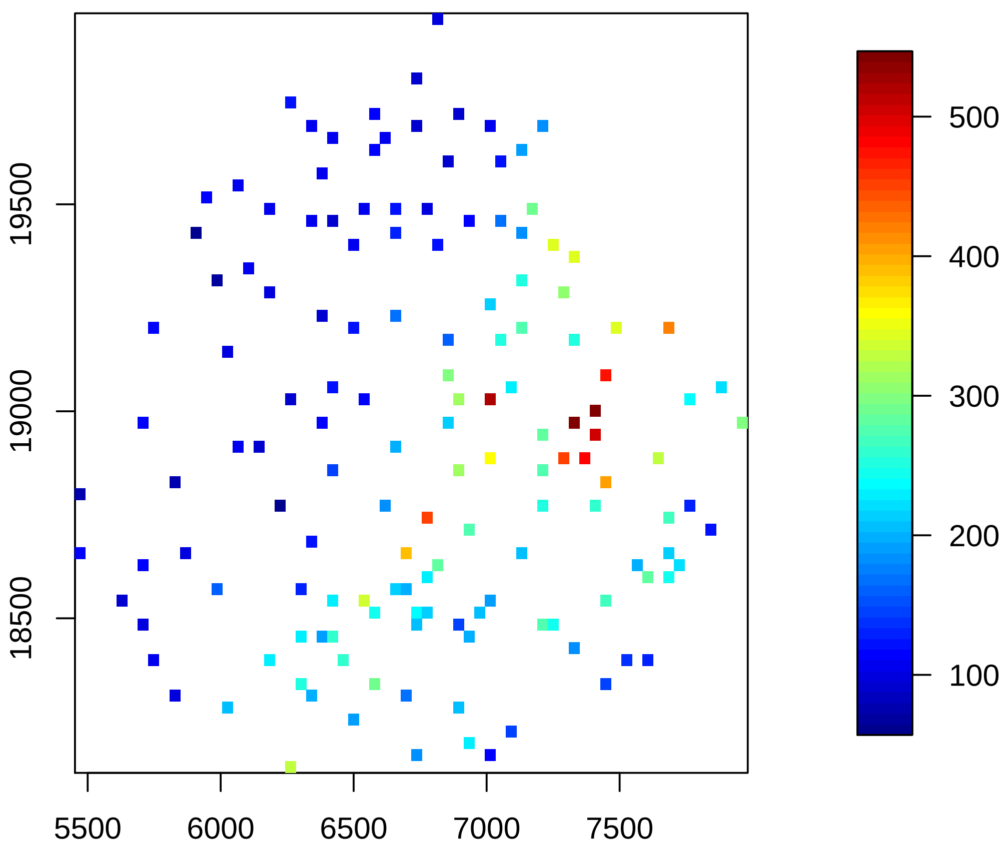
Our Data: Daily precipitation



Our Data: Daily precipitation

Dijon's mustard region!!

Our Data: Daily precipitation



- **Maxima** over 1970-2004
- Data **homogenized** by O. Mestre
- Côte d'Or, Bourgogne France
- 83 locations

A few approaches for modeling spatial extremes

- **Max-stable processes**: Adapting asymptotic results for multivariate extremes
Schlather & Tawn (2003), Naveau et al. (2005), de Haan & Pereira (2005)

- **Bayesian or latent models**: spatial structure **indirectly** modeled via the EVT parameters distribution
Coles & Tawn (1996), Cooley et al. (2005)

- **Linear filtering**: Auto-Regressive spatio-temporal heavy tailed processes,
Davis and Mikosch (2005)

- **Gaussian anamorphosis**: Transforming the field into a Gaussian one
Wackernagel (2003)

Assumptions

- Suppose we know the **marginal** distributions of maxima $M(x)$ with $M(x) =$ the maximum recorded at the location x from a **stationary** field.

- Without loss of generality, we assume that the margins follow an **unit Fréchet**

$$F(u) = \mathbb{P}[M(x) \leq u] = \exp(-1/u)$$

Assumptions

- Suppose we know the **marginal** distributions of maxima $M(x)$ with $M(x)$ = the maximum recorded at the location x from a **stationary** field.

- Without loss of generality, we assume that the margins follow an **unit Fréchet**

$$F(u) = \mathbb{P}[M(x) \leq u] = \exp(-1/u)$$

$$\mathbb{P}\left[M(x) < u_1, M(x+h) < u_2\right] = ??$$

Bivariate case $(M(x), M(x+h))$

A well-known non-parametric structure

$$\mathbb{P}[M(x) < u_1, M(x+h) < u_2] = \exp\left[-\int \max\left\{\frac{g(s,0)}{u_1}, \frac{g(s,h)}{u_2}\right\} \delta(ds)\right]$$

Bivariate case $(M(x), M(x+h))$

A well-known non-parametric structure

$$\mathbb{P}\left[M(x) < u_1, M(x+h) < u_2\right] = \exp\left[-\int \max\left\{\frac{g(s,0)}{u_1}, \frac{g(s,h)}{u_2}\right\}\delta(ds)\right]$$

Special case $u_1 = u_2 = u$

$$\begin{aligned}\mathbb{P}\left[M(x) < u, M(x+h) < u\right] &= \exp(-\theta(h)/u) \\ &= F(u)^{\theta(h)}, \text{ with } F(u) = e^{-1/u}\end{aligned}$$

$\theta(h)$ = Extremal coefficient

$$\mathbb{P}\left[M(x) < u, M(x+h) < u\right] = F(u)^{\theta(h)}$$

with $F(u) = \mathbb{P}\left[M(x) < u\right] = \mathbb{P}\left[M(x+h) < u\right]$

$\theta(h)$ = Extremal coefficient

$$\mathbb{P}\left[M(x) < u, M(x+h) < u\right] = F(u)^{\theta(h)}$$

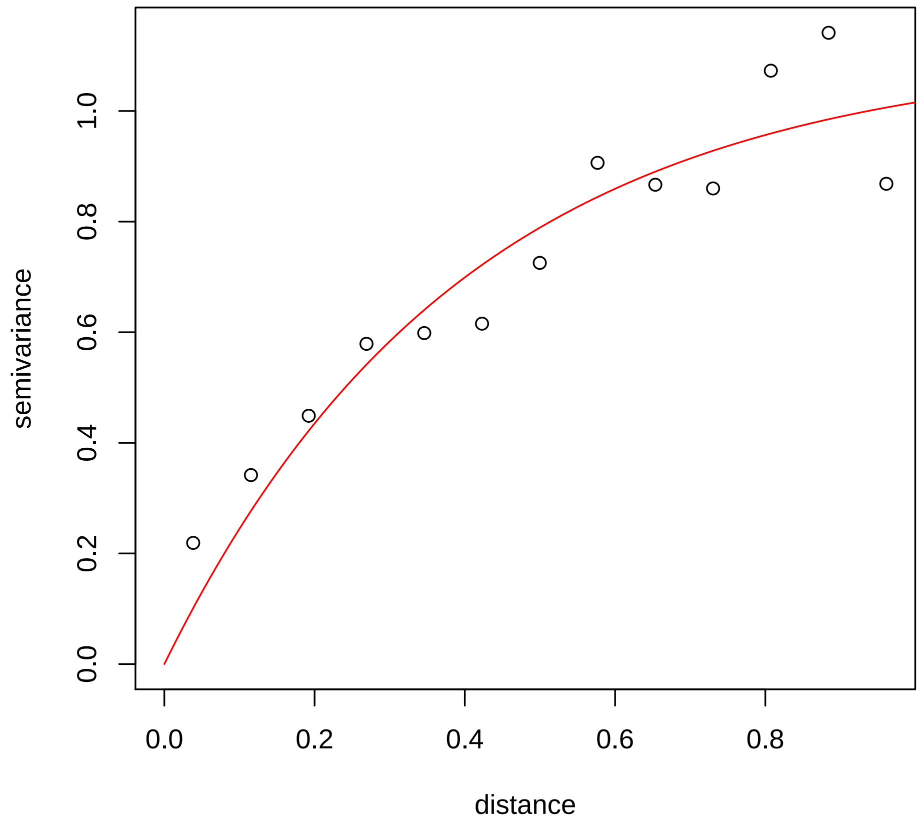
with $F(u) = \mathbb{P}\left[M(x) < u\right] = \mathbb{P}\left[M(x+h) < u\right]$

Interpretation

- Independence $\Rightarrow \theta(h) = 2$
- $M(x) = M(x+h) \Rightarrow \theta(h) = 1$
- Do not completely characterize the **full** bivariate dependence structure

Geostatistics: Variograms

$\gamma(h) = \frac{1}{2}\mathbb{E}|Z(x+h) - Z(x)|^2$ if $\{Z(x)\}$ stationary field s.t. $\mathbb{E}|Z(x)|^2 < \infty$



- Finite if **light** tails
- Capture **all** spatial structure if $\{Z(x)\}$ **Gaussian** fields
- but not well adapted for extremes

A Different Variogram

$$\frac{1}{2}\mathbb{E}\,|M(x+h)-M(x)|^2 \quad =???$$

where $\{M(x)\}$ stationary **max-stable** field with unit-Fréchet margins

A Different Variogram

$$\frac{1}{2}\mathbb{E}\left|M(x+h)-M(x)\right|^2 \quad =???$$

where $\{M(x)\}$ stationary **max-stable** field with unit-Fréchet margins

- $M(x)$ unit-Fréchet $\Rightarrow \mathbb{E}M(x) = \infty$
- $\mathbb{E}\left|M(x+h)-M(x)\right|^2$ not finite

A Different Variogram

$$|F(M(x+h)) - F(M(x))|$$

with $F(u) = \exp(-1/u)$

A Different Variogram

$$\nu(h) = \frac{1}{2}\mathbb{E}\left|F(M(x+h)) - F(M(x))\right|$$

with $F(u) = \exp(-1/u)$

- Defined for light & heavy tails
- Called a Madogram
- Nice links with extreme value theory

A Different Variogram

$$\nu(h) = \frac{1}{2}\mathbb{E}\left|F(M(x+h)) - F(M(x))\right|$$

Why does it work?

A Different Variogram

$$\nu(h) = \frac{1}{2}\mathbb{E}\,|F(M(x+h)) - F(M(x))|$$

Why does it work?

$$\frac{1}{2}|a-b| = \max(a,b) - \frac{1}{2}(a+b)$$

A Different Variogram

$$\nu(h) = \frac{1}{2}\mathbb{E}\left|F(M(x+h)) - F(M(x))\right|$$

Why does it work?

$$\frac{1}{2}|a-b| = \max(a,b) - \frac{1}{2}(a+b)$$

- $a = F(M(x+h))$ and $b = F(M(x))$
- $\mathbb{E}a = \mathbb{E}b = 1/2$

$$\mathbb{E}\max(a,b) = \mathbb{E}F(\underbrace{\max(M(x+h), M(x))}_{\text{max-stable}}) = \frac{\theta(h)}{1+\theta(h)}$$

Madogram $\nu(h) \Rightarrow$ Extremal coeff $\theta(h)$

$$\theta(h) = \frac{1 + 2\nu(h)}{1 - 2\nu(h)}$$

- The madogram $\nu(h)$ gives the extremal coefficient $\theta(h)$

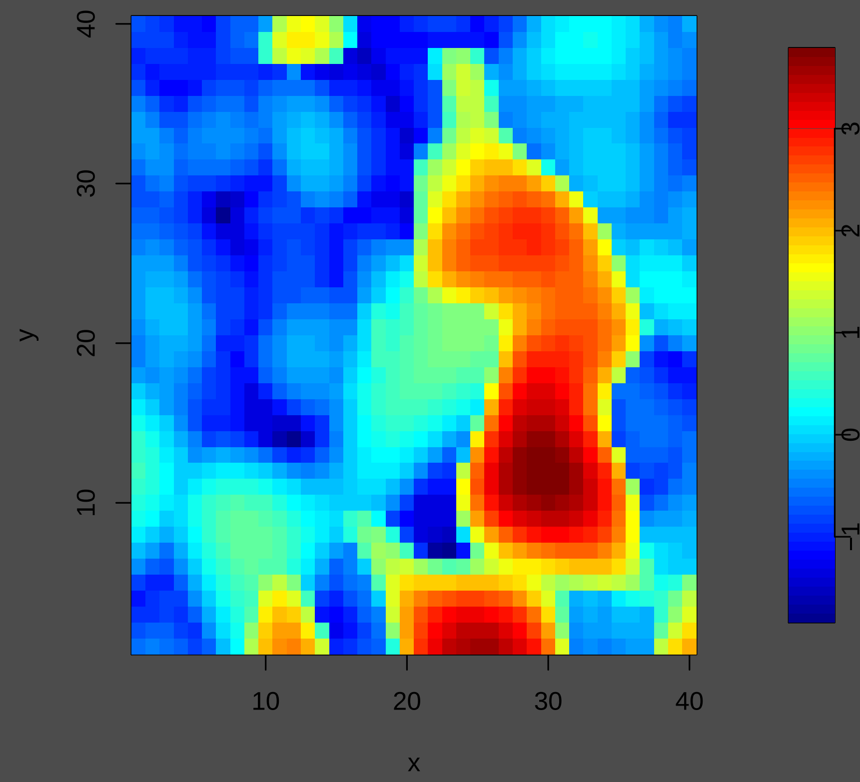
Madogram $\nu(h)$ ⇒ Extremal coeff $\theta(h)$

$$\theta(h) = \frac{1 + 2\nu(h)}{1 - 2\nu(h)}$$

- The madogram $\nu(h)$ gives the extremal coefficient $\theta(h)$
- The madogram $\nu(h) = \frac{1}{2}\mathbb{E}\,|F(M(x+h)) - F(M(x))|$ is easy to estimate:

$$\widehat{\nu}(h) = \frac{1}{N_h} \sum_{||x_i - x_j|| = h} \left| F(M(x_i) - F(M(x_j)) \right|$$

Schlather's models (2003)

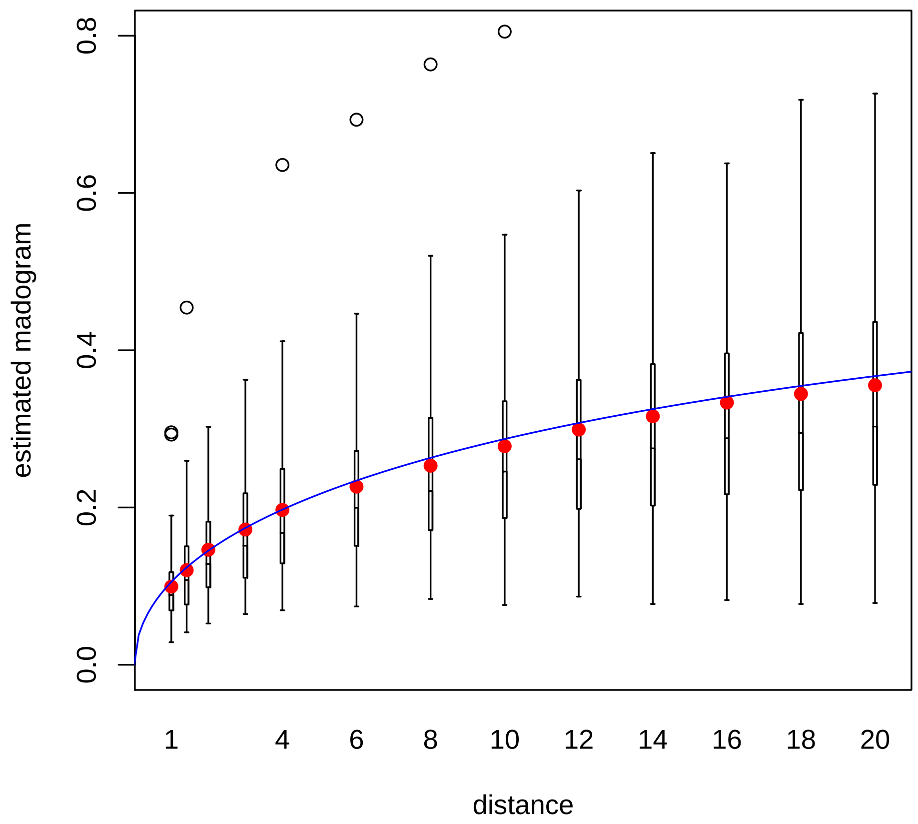


$$\theta(h) = 1 + \sqrt{1 - \frac{1}{2}(\exp(-h/40) + 1)}$$

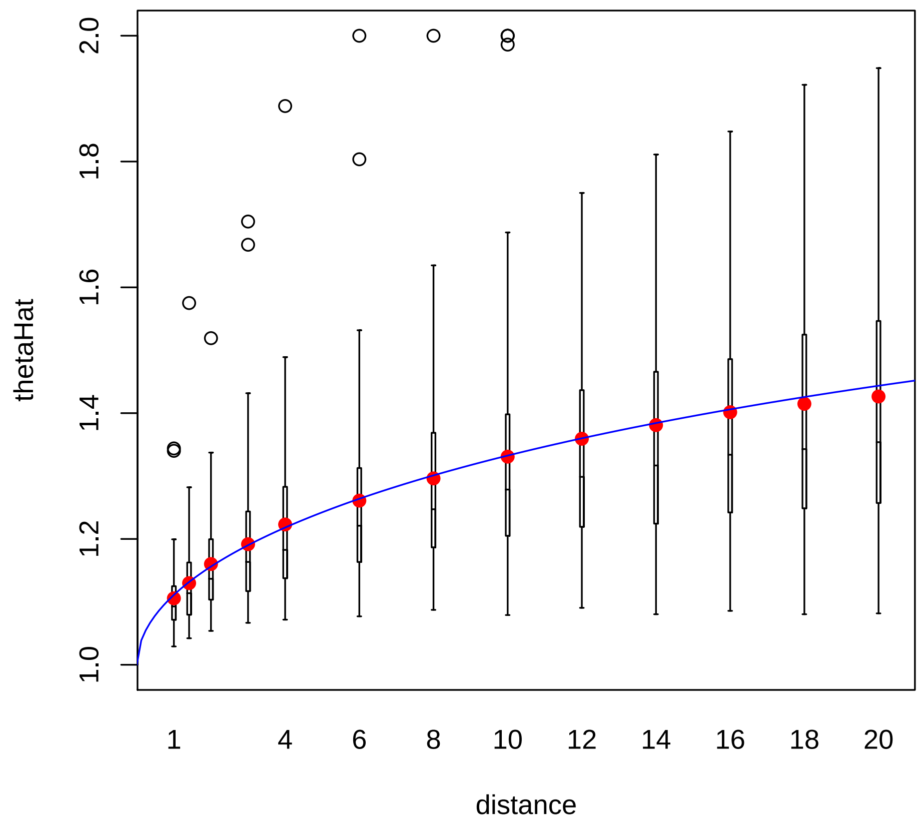
Madogram $\nu(h) \Rightarrow$ Extremal coeff $\theta(h)$

Schlather's fields

Madogram

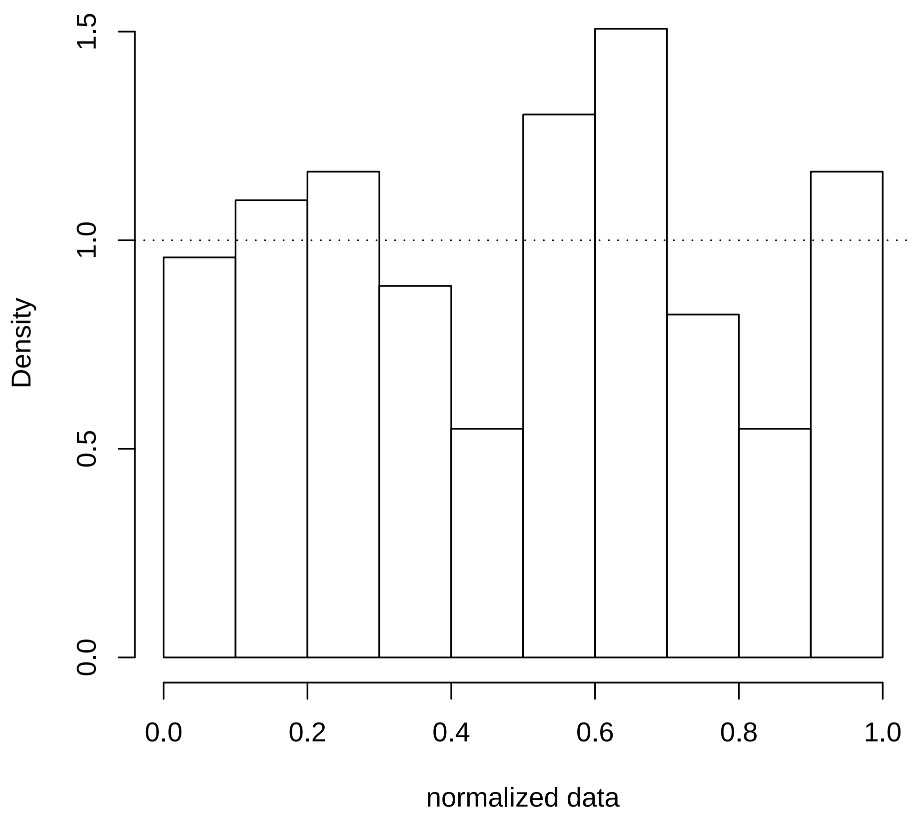


Extremal coeff

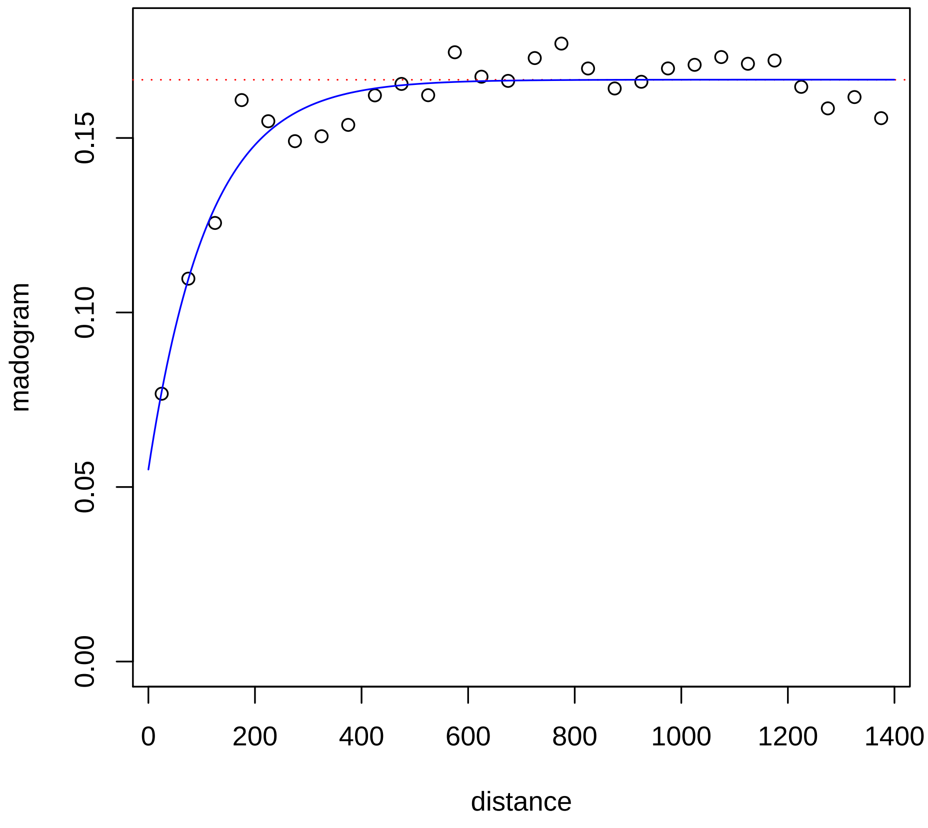


Precipitation

Histogram



Madogram



Building valid $\theta(h)$

Proposition A

Any extremal coefficient function $\theta(h)$ is such that $2 - \theta(h)$ is positive semi-definite.

Building valid $\theta(h)$

Proposition A

Any extremal coefficient function $\theta(h)$ is such that $2-\theta(h)$ is positive semi-definite.

Proposition B

Any extremal coefficient function $\theta(h)$ satisfies the following inequalities

$$\begin{aligned}
\theta(h+k) &\leq \theta(h)\theta(k), \\
\theta(h+k)^\tau &\leq \theta(h)^\tau + \theta(k)^\tau - 1, \text{ for all } 0 \leq \tau \leq 1, \\
\theta(h+k)^\tau &\geq \theta(h)^\tau + \theta(k)^\tau - 1, \text{ for all } \tau \leq 0.
\end{aligned}$$

Complete bivariate structure

Special case $u_1 = u_2 = u$

$$\mathbb{P}\left[M(x) < u, M(x+h) < u\right] \;=\; \exp(-\theta(h)/u)$$

Complete bivariate structure

Special case $u_1 = u_2 = u$

$$\mathbb{P}\left[M(x) < u, M(x+h) < u\right] \;=\; \exp(-\theta(h)/u)$$

General case

$$\mathbb{P}\left[M(x) < u_1, M(x+h) < u_2\right] = \exp(-\theta_\lambda(h)/u_1), \text{ with } \lambda = \frac{u_1}{u_2}$$

Complete bivariate structure

Special case $u_1 = u_2 = u$

$$\theta(h) = \frac{1 + 2\nu(h)}{1 - 2\nu(h)}$$

Complete bivariate structure

Special case $u_1 = u_2 = u$

$$\theta(h) = \frac{1 + 2\nu(h)}{1 - 2\nu(h)}$$

General case

$$\theta_\lambda(h) = \frac{c_\lambda + \nu_\lambda(h)}{1 - \nu_\lambda(h) - c_\lambda}, \text{ with } c_\lambda = \frac{1 + 3\lambda}{4(1 + \lambda)}$$

and

$$\nu_\lambda(h) = \frac{1}{2}\mathbb{E}\,|F(M(x+h)) - F(\lambda M(x))|$$

Complete bivariate structure

Estimators of $\nu_\lambda(h) = \frac{1}{2}\mathbb{E}\,|F(M(x+h)) - F(\lambda M(x))|$

- The "naive" estimator is

$$\widehat{\nu}_\lambda(h) = \frac{1}{N_h} \sum_{||x_i - x_j|| = h} \left| F(M(x_i) - F(\lambda M(x_j)) \right|$$

but it does not satisfy $\widehat{\theta}_0(h) = \widehat{\theta}_\infty(h) = 1$

Complete bivariate structure

Estimators of $\nu_\lambda(h) = \frac{1}{2}\mathbb{E}\,|F(M(x+h)) - F(\lambda M(x))|$

- The "naive" estimator is

$$\hat{\nu}_\lambda(h) = \frac{1}{N_h} \sum_{||x_i - x_j|| = h} \left| F(M(x_i) - F(\lambda M(x_j)) \right|$$

but it does not satisfy $\hat{\theta}_0(h) = \hat{\theta}_\infty(h) = 1$

- An unbiased estimator satisfying the above condition is

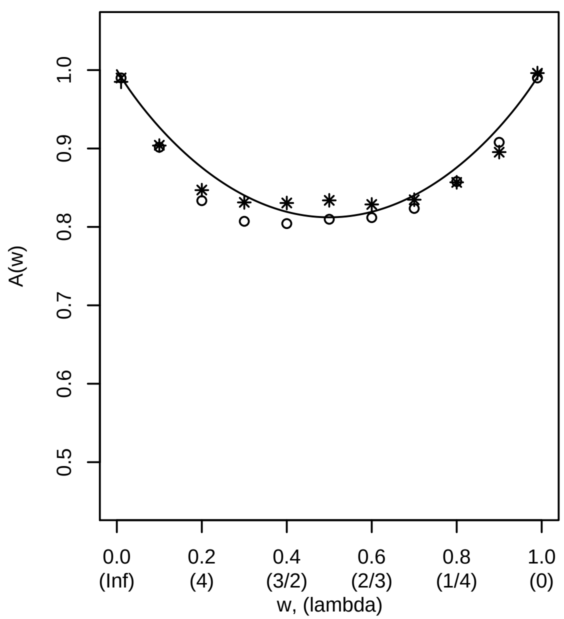
$$\hat{\hat{\nu}}_\lambda(h) = \hat{\nu}_\lambda(h) - (1 - \omega(\lambda))\frac{1}{n}\sum \overline{F}(\lambda M(x_j)) - \omega(\lambda)\frac{1}{n}\sum F(\lambda M(x_j)) + \frac{1}{4}$$

with $\omega(\lambda) = \lambda/(1+\lambda)$

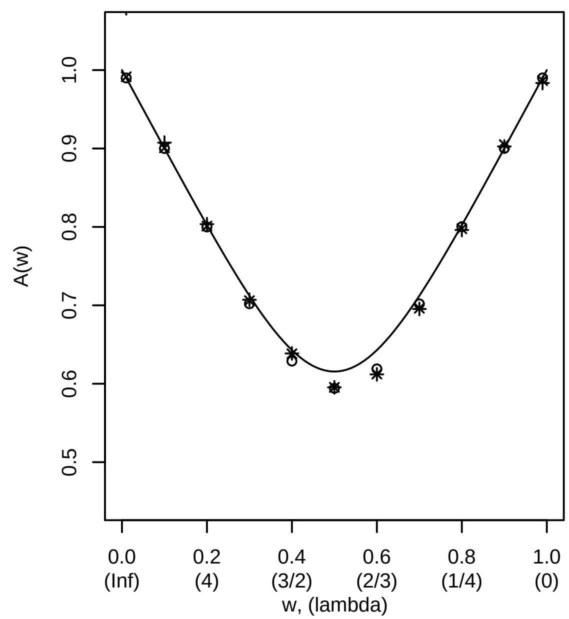
$$G(u_1, u_2) = \exp\left[-\left(\frac{1}{u_1} + \frac{1}{u_2}\right) A\left(\frac{u_1}{u_1 + u_2}\right)\right]$$

with $A(w) = (1 - t_1)w + (1 - t_2)(1 - w) + [(t_1 w)^{(1/\alpha)} + (t_2(1 - w))^{(1/\alpha)}]^{\alpha}$

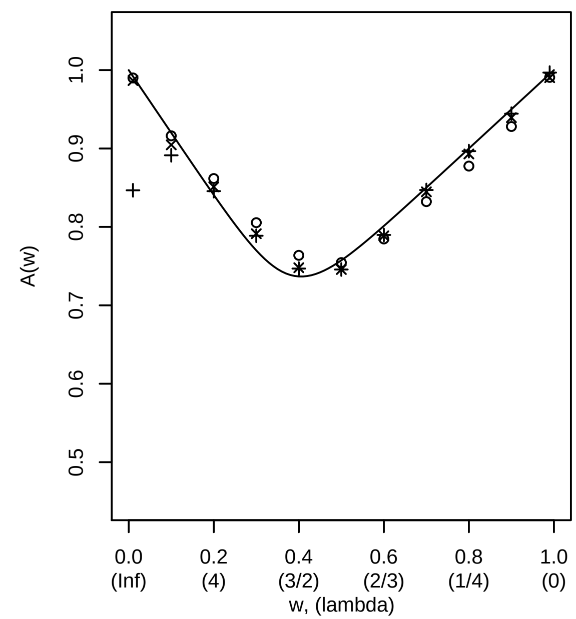
$\alpha = .7, t_1 = t_2 = 1$



$\alpha = .3, t_1 = t_2 = 1$



$\alpha = .2, t_1 = 0.8, t_2 = 0.5$

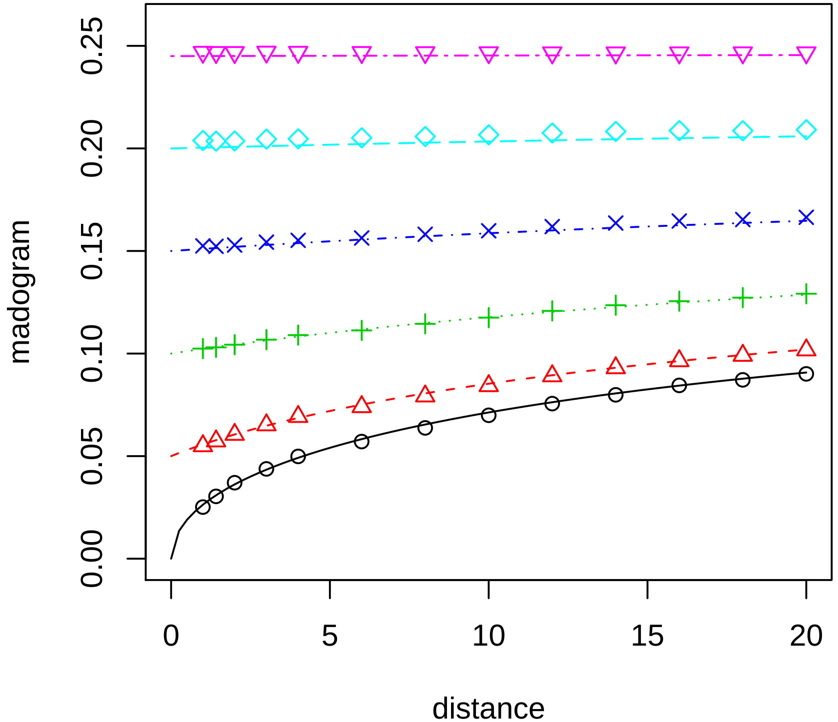


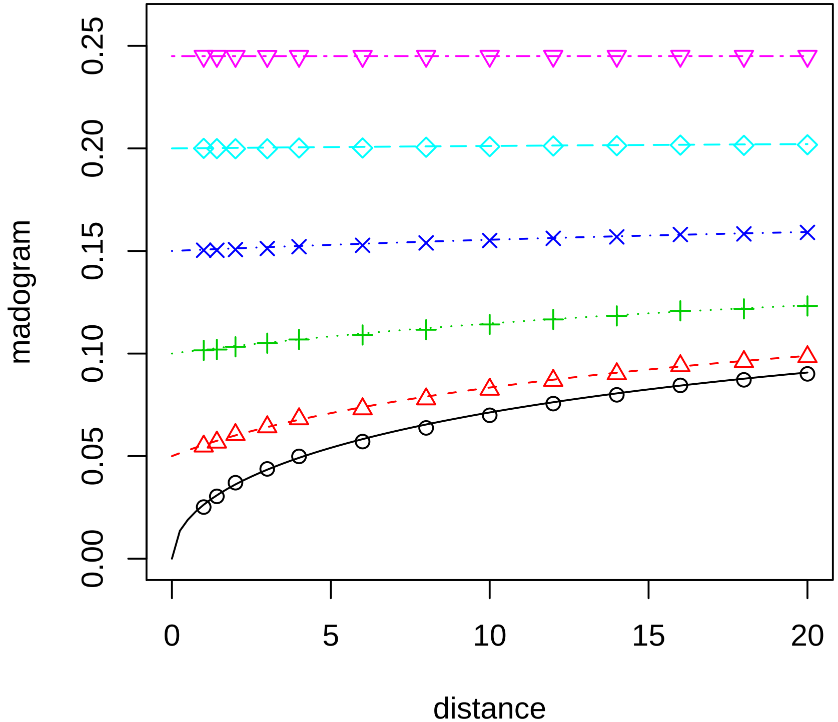
$\lambda-$Madogram for spatial fields

Schlather's fields

$\lambda = 1$, 2/3 ,..., 1/4 ,..., 1/99

$\lambda = 1$, 3/2 ,..., 4 ,..., 99





Take-home messages

- **Madogram** $\nu(h) \Rightarrow$ **Extremal coeff** $\theta(h)$
- **Extremal coeff** $\theta(h) \Rightarrow$ Info about the spatial dependence
- λ-madogram captures bivariate structure

Take-home messages

- **Madogram** $\nu(h) \Rightarrow$ **Extremal coeff** $\theta(h)$
- **Extremal coeff** $\theta(h) \Rightarrow$ Info about the spatial dependence
- λ-madogram captures bivariate structure

Future research

- Derive limiting results for unknown marginals
- Find "madograms" for exceedances
- Develop spatial interpolation methods for maxima
- Derive statistical schemes for downscaling of extremes

Acknowledgement

NSF-GMC (ATM-0327936)

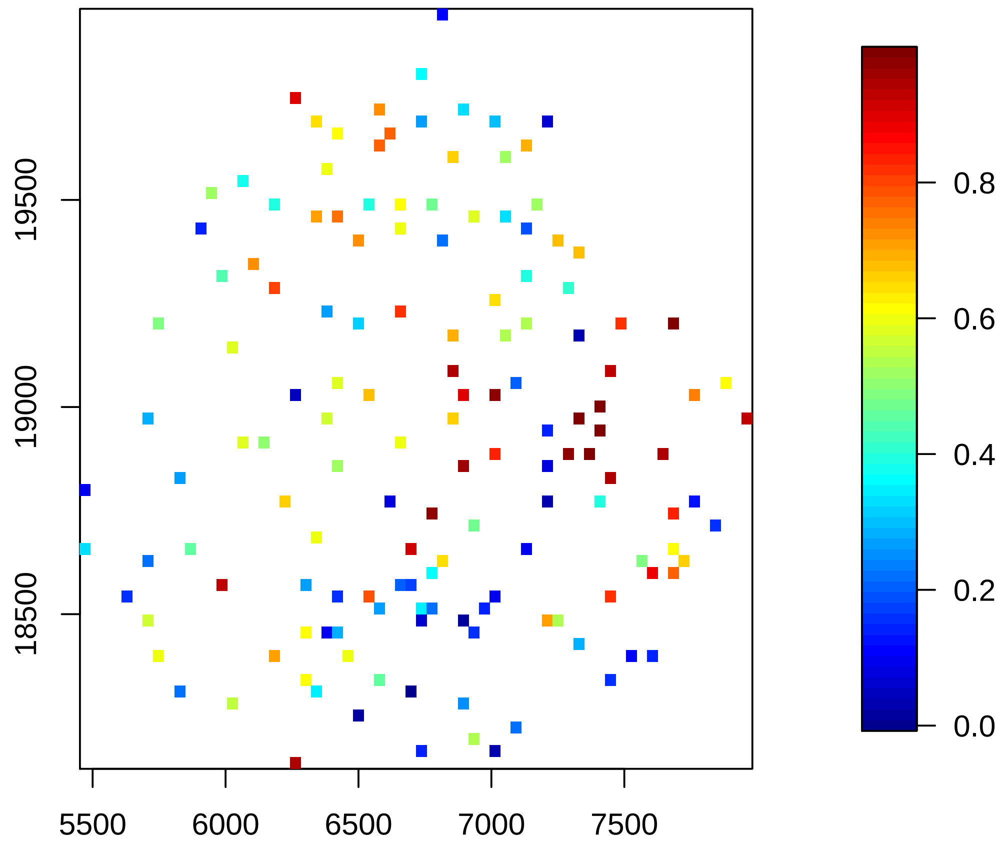
E2C2: **E**xtreme **E**vents, **C**ause and **C**onsequences: Post-docs

To learn more about this topic ...

http://amath.colorado.edu/faculty/naveau/

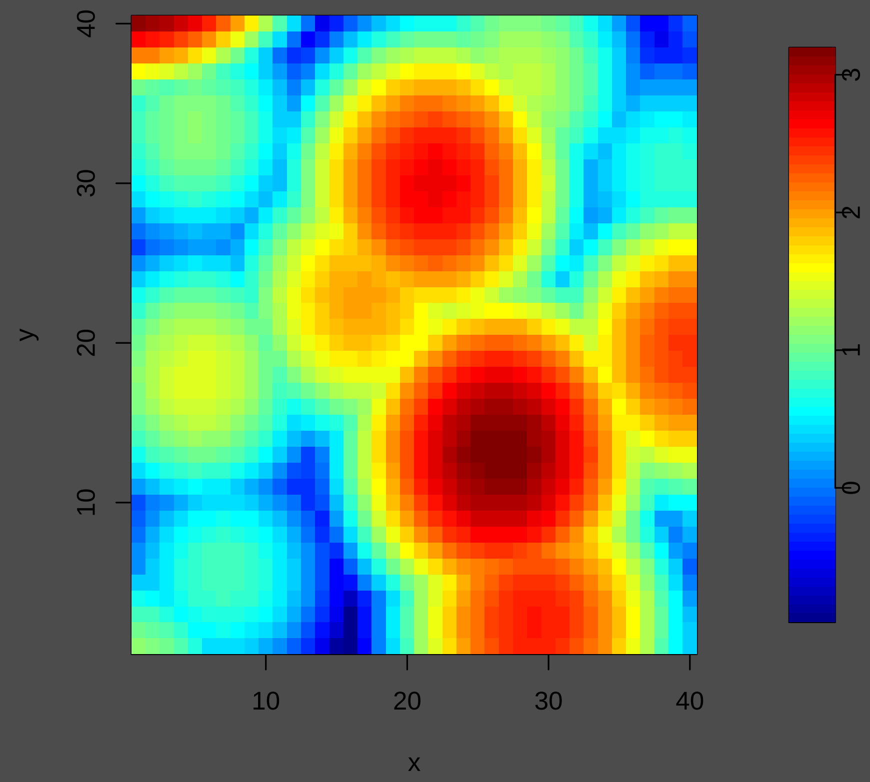
- Naveau, P., Poncet, P., and Cooley, D. (2005). First-order variograms for extreme bivariate random vectors (submitted).
- Schlather, M. and Tawn, J. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference.
- Davis, R. and Resnick, S. (1993). Prediction of stationary max-stable processes, Ann. of Applied Prob 3, 497–525
- Capeera, P., Fougéres A.L., and Genest, C (1997). A non-parametric estimation procedure for bivariate extreme value copulas.
- Smith, R. (1990). Max-stable processes and spatial extremes.

“Better to have the approximate solution to the correct problem than the exact solution to the wrong problem” -J. Tukey



Our "renormalized" random field

Smith's models (2003)

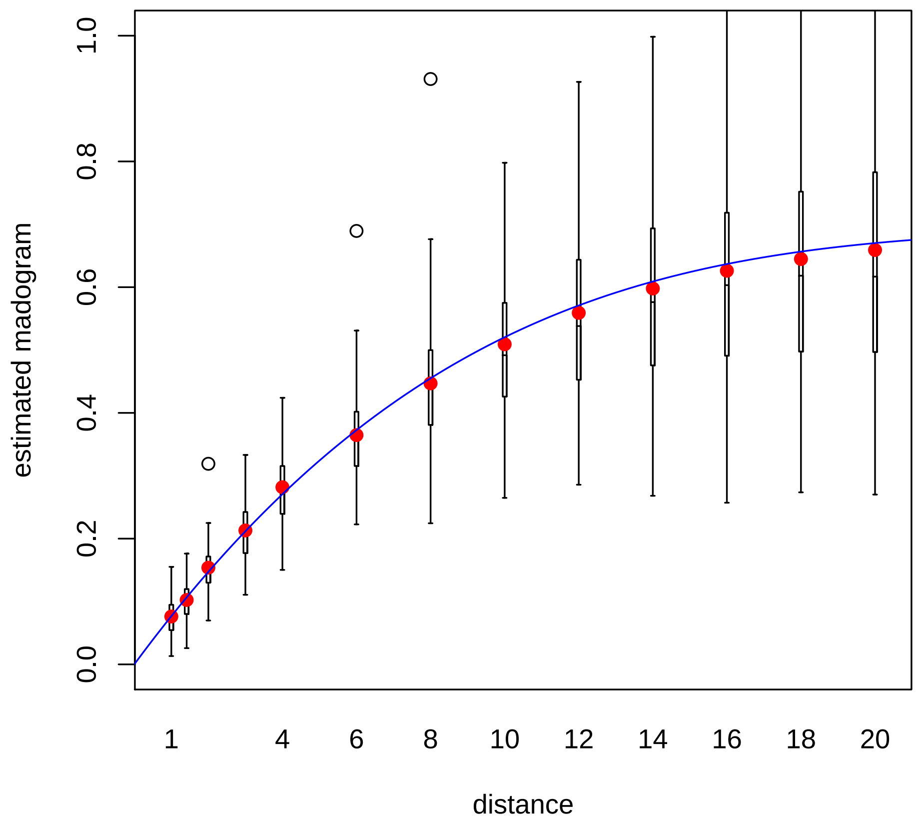


$$\theta(h) = 2\Phi\left(\sqrt{h^T\Sigma^{-1}h}/2\right)$$

Madogram $\nu(h) \Rightarrow$ Extremal coeff $\theta(h)$

Smith's fields

Madogram



Extremal coeff