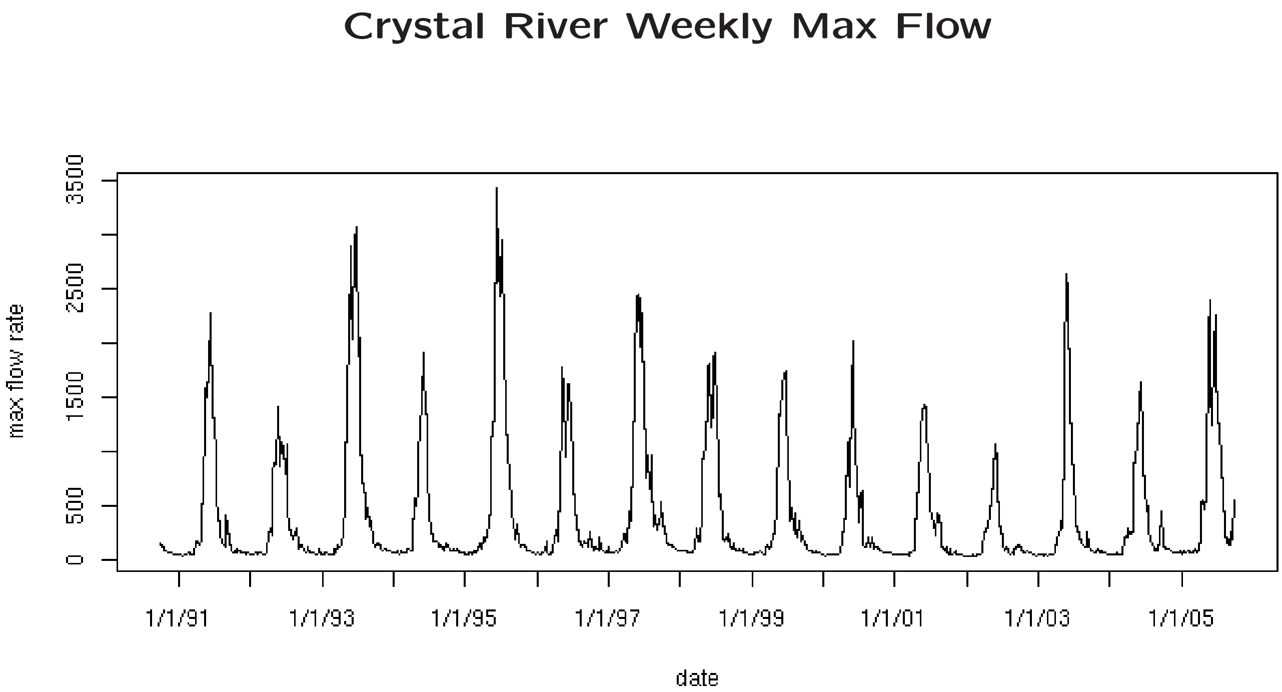
Prediction for Max-Stable Processes via an Approximated Conditional Density

Daniel Cooley
Colorado State University

Richard A. Davis
Columbia University

Philippe Naveau
LSCE, IPSL-CNRS, France

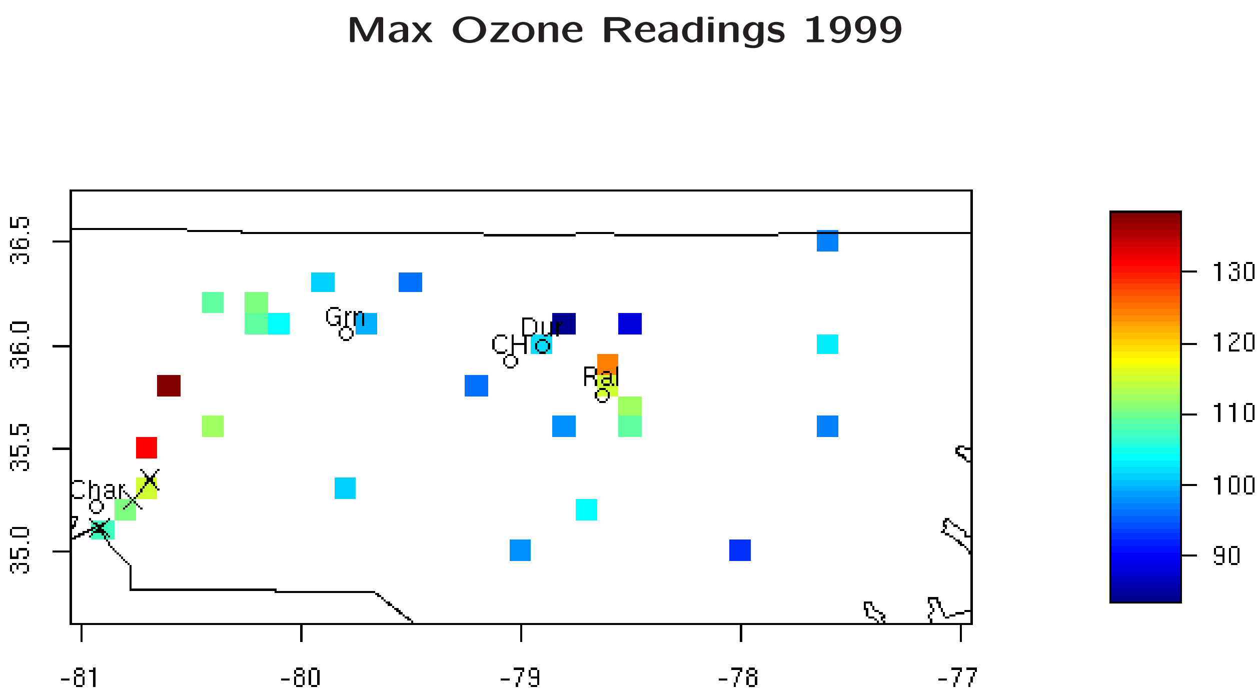
Motivating Example 1: Time Series Data



Q: Can we use the previous weekly maximum flows to predict the next weekly maximum flow?

Q: What is a measure of risk (e.g., an estimated conditional 0.95 quantile) given that observed values are high ?

Motivating Example 2: Spatial Data



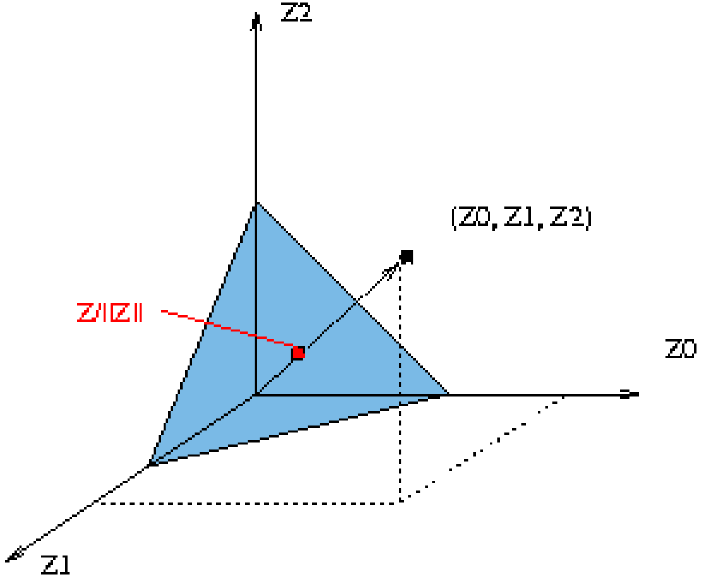
Q: Can we use the observed annual maxima to predict (interpolate) the annual max at an unobserved location?

Q: What is probability that it exceeds some standard?

Our Approach for Project

1. Assume observations arise from a max-stable process from which we have a finite number of observations.
2. Find and fit a useful model for the observations plus the unmonitored location/time.
3. Approximate the *conditional density* of the unmonitored location given the "nearby" observations.

Multivariate Max-Stable Distributions



Basic idea: Given a nice common marginal, MMSD's can be described in terms of a point process whose intensity measure is a product measure of "radial" and "angular" components.

Multivariate Max-Stable Distributions

If $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_p))^T$ has a multivariate max-stable distribution with unit Fréchet margins ($\mathbb{P}(Z(\boldsymbol{x}_i) \leq z) = \exp(-z^{-1})$) then:

$$G(\boldsymbol{z}) = \mathbb{P}(\boldsymbol{Z} \leq \boldsymbol{z}) = \exp[-V(\boldsymbol{z})], \text{ where}$$

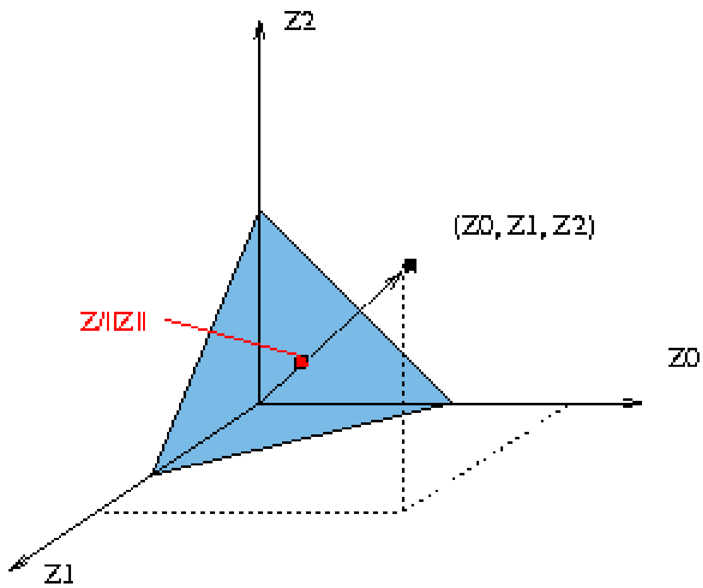
$$V(\boldsymbol{z}) = \int_{S_p} \max_i \left(\frac{w_i}{z_i}\right) dH(\boldsymbol{w}),$$

H is a positive measure on S_p, s.t.

$$\int_{S_p} w_i dH(\boldsymbol{w}) = 1,$$

and $S_p = \{\boldsymbol{w} \in \mathbb{R}_+^p | w_1 + \ldots + w_p = 1\}$.

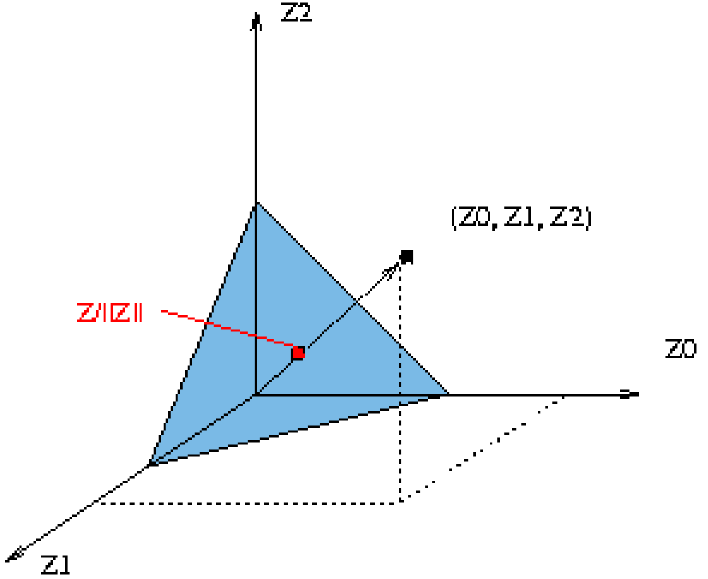
Multivariate Max-Stable Distributions



$\boldsymbol{z} \in \mathbb{R}^p$; $\boldsymbol{w} \in S_p$

$V(\boldsymbol{z})$: "exponent measure function" – relates to point process intensity, $G(\boldsymbol{z}) = \exp[-V(\boldsymbol{z})]$
$H(\boldsymbol{w})$: "spectral measure" – lives on unit simplex, meets center of mass condition, $V(\boldsymbol{z}) = \int_{S_p} \max_i \left(\frac{w_i}{z_i}\right) dH(\boldsymbol{w})$
$h(\boldsymbol{w})$: "spectral density" – exists if $H(\boldsymbol{w})$ is differentiable, not Fourier!

Multivariate Max-Stable Distributions



- no parametric form for entire family
- a few useful parametric sub-families suggested

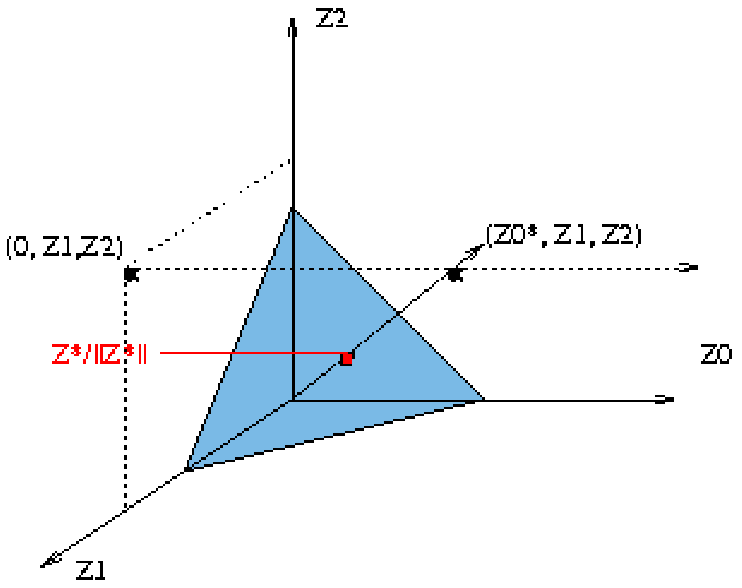
Models for Multivariate MSD's

Exponent measure function $V(\boldsymbol{z})$	Spectral density $h(\boldsymbol{w})$
• Logistic • Asymmetric Logistic (Tawn, 88) • Negative Logistic (Joe, 90)	• Dirichlet (Coles & Tawn, 91)
+ Can obtain $G(\boldsymbol{z})$ – Overparametrized? – Less flexible?	+ More flexibility? – Cannot directly get $G(\boldsymbol{z})$

Models for Multivariate MSD's

Exponent measure function $V(\boldsymbol{z})$	Spectral density $h(\boldsymbol{w})$
• Logistic • Asymmetric Logistic (Tawn, 88) • Negative Logistic (Joe, 90)	• Dirichlet (Coles & Tawn, 91) • *Pairwise Beta* * meets COM condition * parameters interpretable * pairwise specification
+ Can obtain $G(\boldsymbol{z})$ – Overparametrized? – Less flexible?	+ More flexibility? – Cannot directly get $G(\boldsymbol{z})$

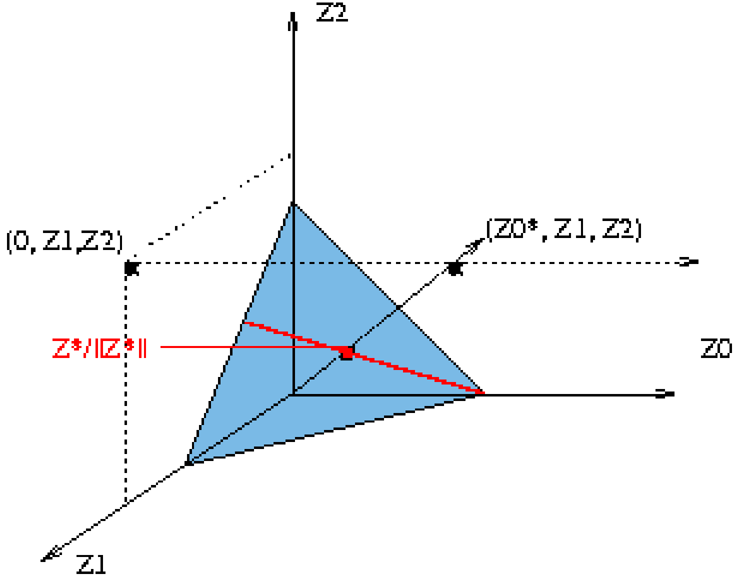
Prediction: Approximating the conditional density?



If $V(\boldsymbol{z})$ is known and differentiable, then joint density can be obtained exactly. However, we are modeling $h(\boldsymbol{w})$.

Assume Z_1, Z_2 are observed and Z_0 is unobserved. Any predictor Z_0^* will yield a point $\boldsymbol{Z}^* = (Z_0^*, Z_1, Z_2)$ which can be mapped back to S_p as $\frac{\boldsymbol{Z}^*}{\|\boldsymbol{Z}^*\|_1}$.

Approximating the conditional density?



Given the radius is large, by knowing the values of the spectral density at $\frac{Z^*}{\|Z^*\|_1}$ and the value of the "radius" $\|Z^*\|_1$, we can approximate the values of the joint density and in turn the *conditional density*.

Approximating the conditional density?

If $V(\boldsymbol{z}) = \mu\{(0, z]^c\}$ is small (i.e. the radius is large), then

$$G(\boldsymbol{z}) = \exp(-V(\boldsymbol{z})) \approx 1 - V(\boldsymbol{z}).$$

Using Coles and Tawn (91) result to estimate the density at $\boldsymbol{z}$:

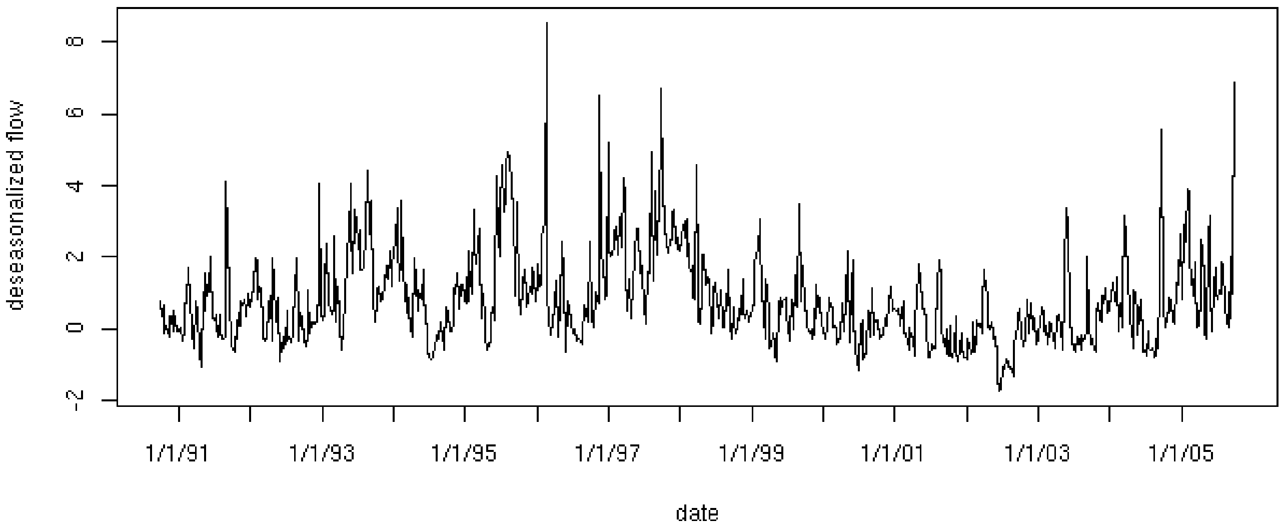
$$g(\boldsymbol{z}) \approx \frac{\partial}{\partial z_1, \ldots, \partial z_p}[1 - V(\boldsymbol{z})] = \frac{1}{||\boldsymbol{z}||^{-(p+1)}} h\left(\frac{\boldsymbol{z}}{||\boldsymbol{z}||}\right)$$

So conditional density can be approximated by

$$g_{Z_p|Z_1,\ldots Z_{p-1}}(z_p|z_1, \ldots, z_{p-1}) \approx \frac{\frac{1}{||\boldsymbol{z}||^{-(p+1)}} h\left(\frac{\boldsymbol{z}}{||\boldsymbol{z}||}\right)}{\int_0^\infty \frac{1}{||\boldsymbol{z}^*||^{-(p+1)}} h\left(\frac{\boldsymbol{z}^*}{||\boldsymbol{z}^*||}\right) d\zeta}$$

where $\boldsymbol{z}^* = (z_1, \ldots, z_{p-1}, \zeta)$.

Time Series Example



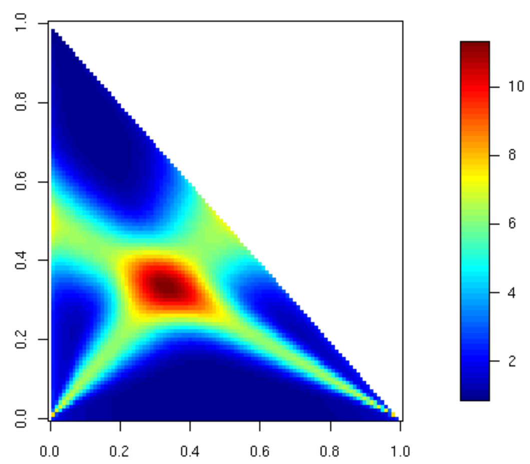
1. Data are deseasonalized.
2. ACF/PACF plots $\longrightarrow$ base prediction on previous two observations.
3. GEV is fit to marginal distribution, then transformed.

Fitting the spectral density model

Time series broken into non-overlapping triples. Dependence measured by the extremal coefficient (Schlather & Tawn 03).

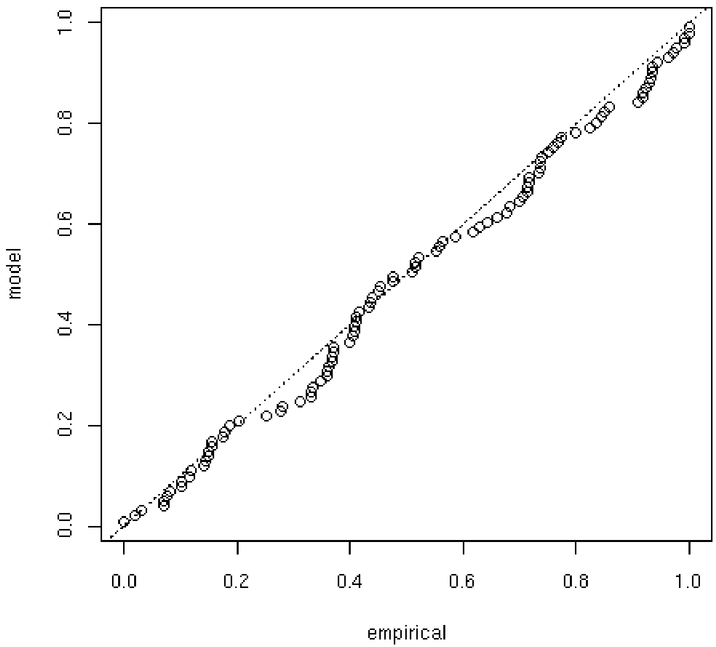
$$\phi_{1,2} = V(1, 1, \infty, \ldots, \infty); \ \phi_{i,j} \in [1, 2]$$

Ext. coefficients estimated at lags 1 and 2, $(\hat{\phi}_{-1}, \hat{\phi}_{-2}) = (1.36, 1.49)$, and pairwise beta model parameters found to match the extremal coefficient estimates. $(\hat{\alpha}; \hat{\beta}) = (1, 16, 0.7, 16)$.



Time series prediction

75 largest triples selected for prediction. Conditional density of 3rd component given 1st and 2nd components is approximated.



Time series prediction

75 largest triples selected for prediction. Conditional density of 3rd component given 1st and 2nd components is approximated.

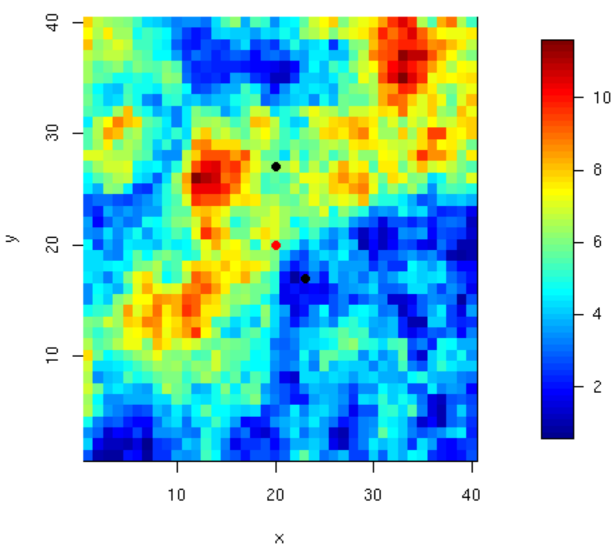
Assessing Risk

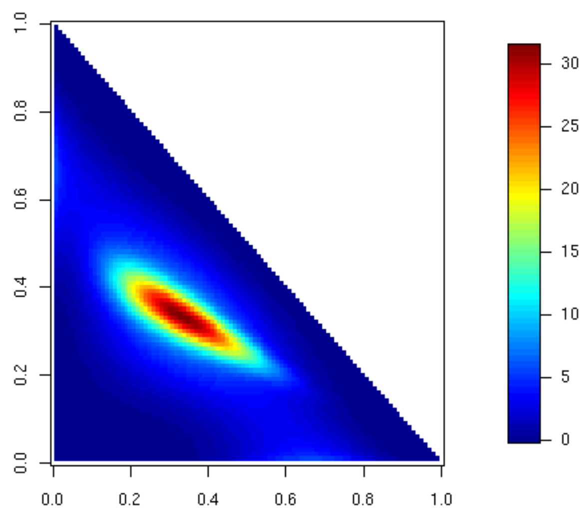
How well is the 95% quantile predicted?

- Our method: 6 observations (8%) exceed the estimated quantile
- AR(2): 9 observations (12%) exceed the estimated quantile

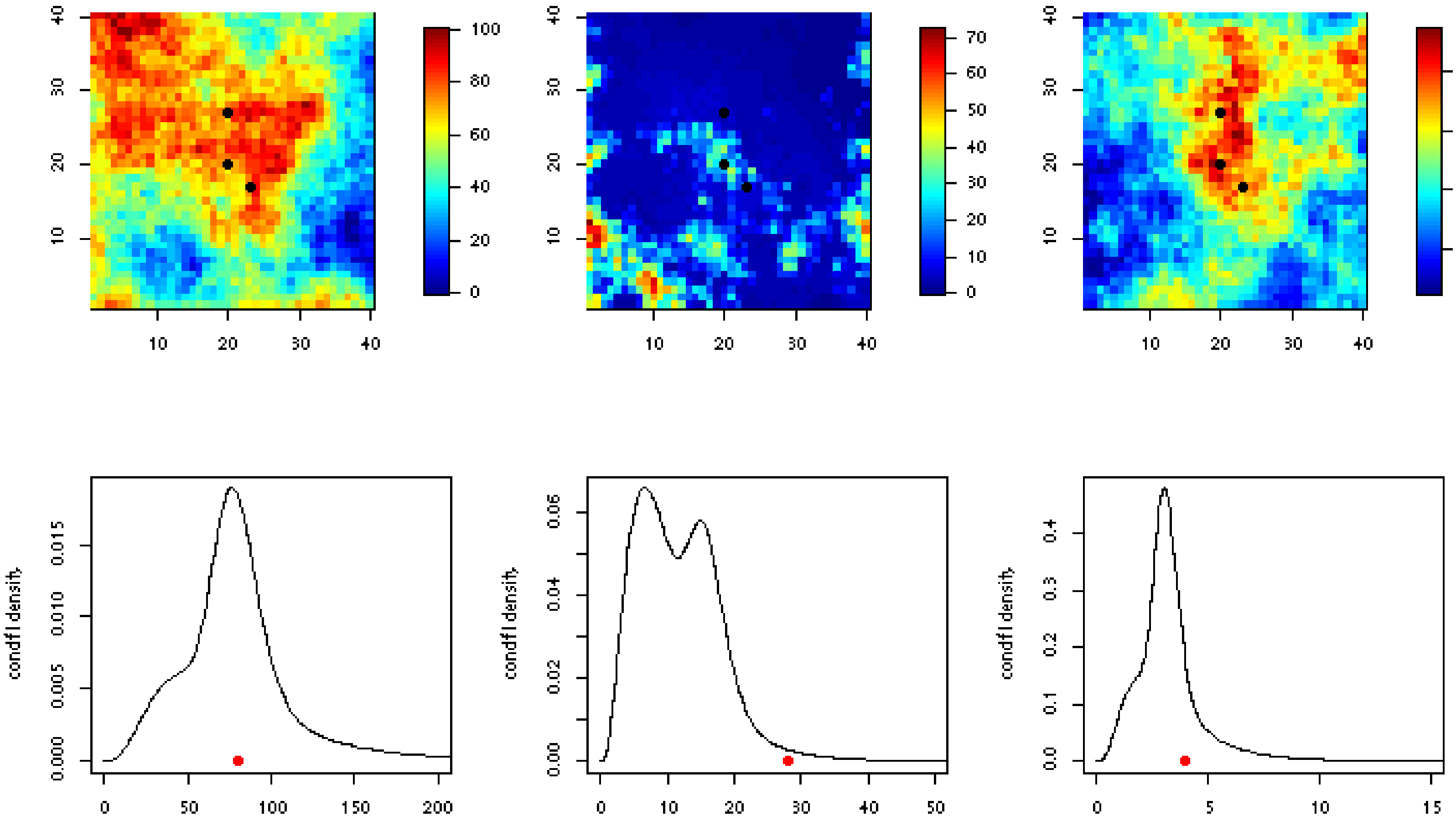
Spatial interpolation

- MS fields w/ Fréchet marginals simulated (Schlather 02).
- Known bivariate dependence structure: $\phi = (1.34, 1.28, 1.22)$.
- Pairwise beta model fit as before: $(\alpha; \beta) = (4.3; 0.87, 4.4, 74)$.
- Conditional density approximated for largest 300 of 1000 simulated fields.





Spatial interpolation examples

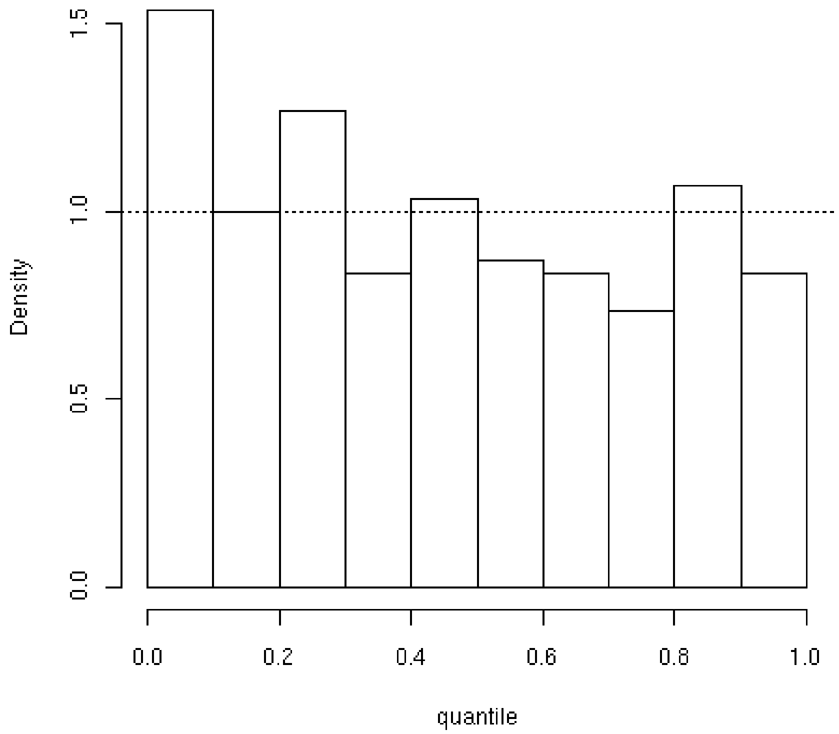


(80.69, 79.95, 80.45)

(4.50, 16.85, 28.14)

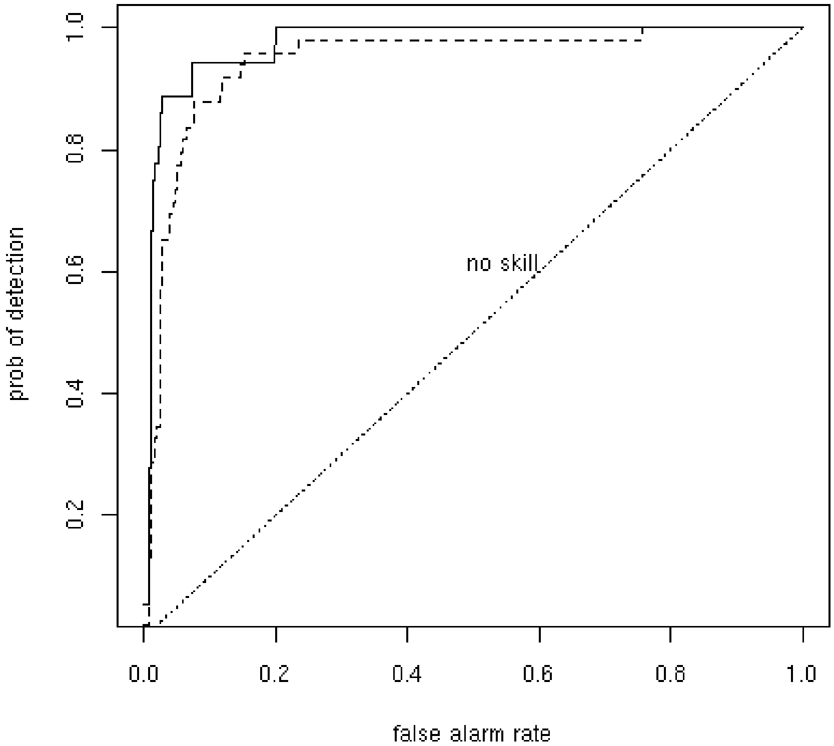
(3.47, 3.15, 3.93)

Repeated Simulation Results

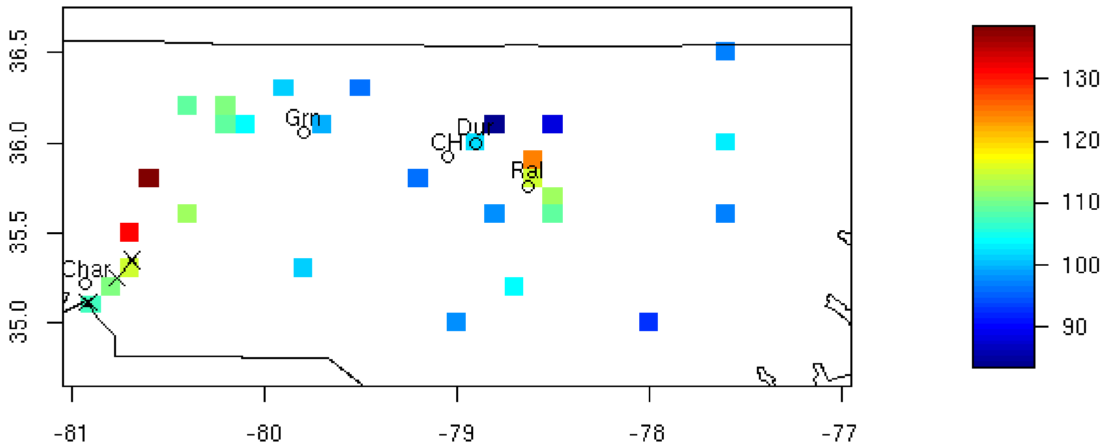


Spatial interpolation results

How well does the method assess exceeding some standard?



Ground level ozone

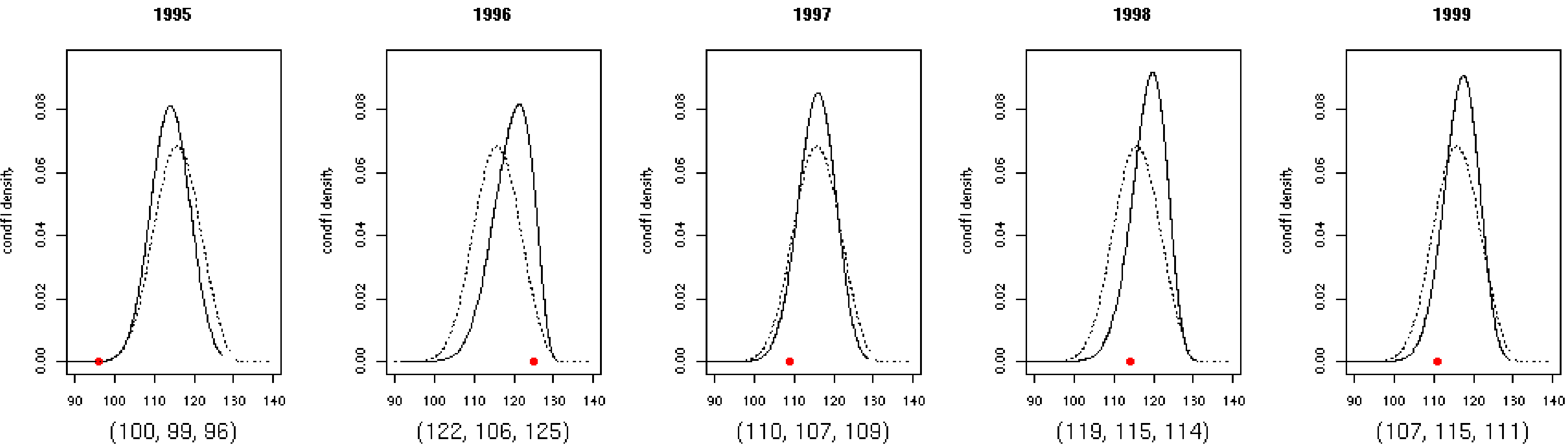


- Only 5 years of data.
- Marginal distributions from (Gilleland, et al 06).
- Dependence estimated as function of distance using madogram (Naveau et al 06).
- Weak dependence estimated $\phi = (1.95, 1.84, 1.67)$.

Ground level ozone

Unable to restrict attention to the "large" years. An adjustment is made to earlier approximations.

$$g(\boldsymbol{z}) \approx \exp(-1/z_3) \frac{1}{||\boldsymbol{z}||^{(p+1)}} h\left(\frac{\boldsymbol{z}}{||\boldsymbol{z}||}\right)$$



Summary

- Method for approximating the conditional density of an unobserved component of a max-stable vector given the other components.
- Method designed specifically for dealing with extremes; alternative to standard time series and spatial prediction methods which are better suited for central tendancies.
- Uses only the spectral density, which (possibly) allows for more flexiblity in modeling.
- Applied to both time series and spatial contexts.
- Introduced pairwise beta model.
- Future work: extend from max-stable case to exceedances over a threshold case.

References

Coles, S. and Tawn, J. (1991). Modeling multivariate extreme events. *Journal of the Royal Statistical Society, Series B*, 53:377–92.

Cooley, D., Naveau, P., and Poncet, P. (2006). Variograms for spatial max-stable random fields. In Bertail, P., Doukhan, P., and Soulier, P., editors, *Dependence in Probability and Statistics*, Springer Lecture Notes in Statistics. Springer, New York.

Gilleland, E., Nychka, D., and Schneider, U. (2006). Spatial models for the distribution of extremes. In Clark, J. and Gelfand, A., editors, *Hierarchical Modelling for the Environmental Sciences: Statistical Methods and Applications*, pages 170–183. Oxford University Press, New York.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistics and Probability Letters*, 9:75–81.

Naveau, P., Guillou, A., Cooley, D., and Diebolt, J. (2006). Modeling pairwise dependence of maxima in space. *Submitted*.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1):33–44.

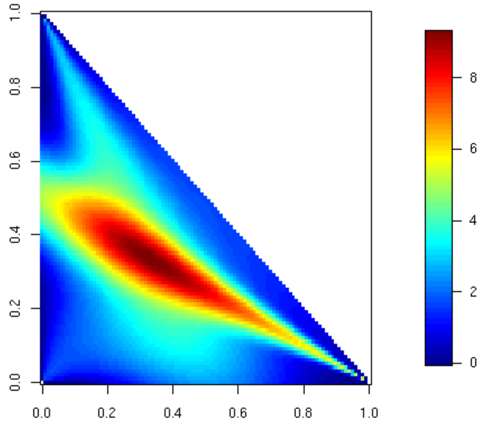
Tawn, J. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, 75:397–415.

Pairwise Beta Model

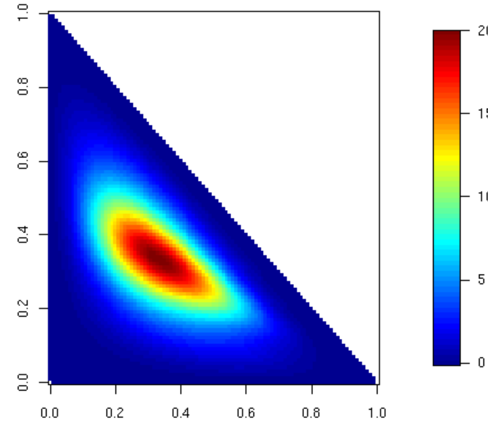
$$
\begin{aligned}
h_p(\boldsymbol{w};\alpha,\boldsymbol{\beta}) &= K_p(\alpha)\sum_{i\neq j} h_{i,j}(w_i,w_j;\alpha,\beta_{i,j}), \text{ where} \\
h_{i,j}(w_i,w_j;\alpha,\beta_{i,j}) &= (w_i+w_j)^{(p-1)(\alpha-1)}(1-(w_i+w_j))^{\alpha-1}\times \\
&\quad \frac{\Gamma(2\beta_{i,j})}{(\Gamma(\beta_{i,j}))^2}\left(\frac{w_i}{w_i+w_j}\right)^{\beta_{i,j}-1}\left(\frac{w_j}{w_i+w_j}\right)^{\beta_{i,j}-1} \\
K_p(\alpha) &= \frac{2(p-3)!}{(p-1)\sqrt{p}}\frac{\Gamma(\alpha p)}{\Gamma(\alpha p-\alpha-p+3)\Gamma(\alpha+p-3)}
\end{aligned}
$$

Advantages:

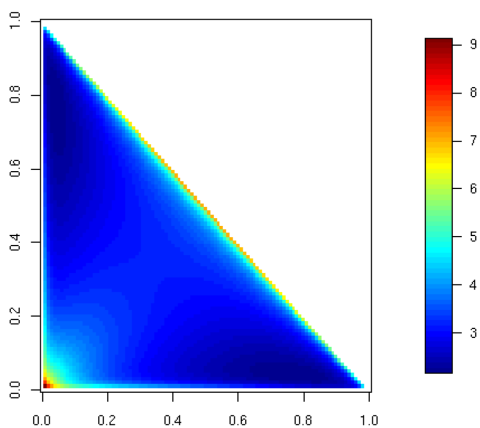
- no adjustment necessary to get center of mass condition
- parameters have some interpretation: α controls overall dependence, $\beta_{i,j}$'s control pairwise dependence
- largely specified by pairwise parameters



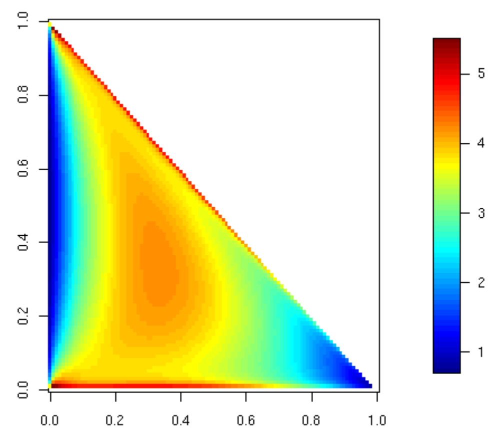
$\alpha = 1, \beta = (2, 4, 15)$



$\alpha = 4, \beta = (2, 4, 15)$



$\alpha = 1, \beta = (2, .5, .5)$



$\alpha = 1, \beta = (2, 2, .5)$