# Statistical Analysis of Extremes in Geophysical Science reading group

Fawcett L and D Walshaw. (*in press*). Markov chain models for extreme wind speeds. *Environmetrics*, **DOI: 10.1002/env.794**.

Presented by Eric Gilleland on Thursday, July 6, 2006.

http://www.ral.ucar.edu/staff/ericg/readinggroup.html

# Statistical Analysis of Extremes in Geophysical Science

## A couple of relevant papers (not explicitly cited)


Coles SG, JA Tawn, and RL Smith. (1994). A seasonal Markov model for extremely low temperatures. *Environmetrics*, **5**:221–239.

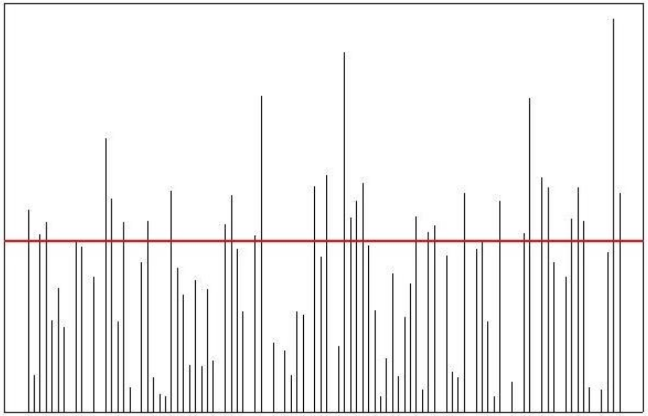
Smith RL, JA Tawn, and SG Coles. (1997). Markov chain models for threshold exceedances. *Biometrika*, **84**(2):249–268.

# Markov chain models for extreme wind speeds

## Outline

- Background and Motivation
  - Dependence in threshold exceedances
  - Wind Speed Data
  - Generalized Pareto Distribution (GPD)
- Markov chain model
  - Joint density
  - Bivariate GPD
- Bayesian Inference
- Results for High Bradfield
- Further model considerations

# Background and Motivation: Dependence...



Exceedances are not usually independent. Can *decluster* exceedances to remove dependencies, but ...

- Also removes data.
- Parameter estimates are sensitive to filtering scheme.
- Information about the dependence may be important.
- May introduce biases in parameter estimates.

# Background and Motivation: Dependence...

**Solutions?**

- Make adjustments to standard errors and confidence intervals.
- Explicitly model the dependence (e.g., Coles *et al*. (1994), Smith *et al*. (1997))

# Background and Motivation: Data



- Hourly *gust* maximum wind speeds (knots).
- Location: High Bradfield (high-altitude site in Pennines).
- 10-year record from 1-Jan-1975 to 31-Dec-1984 ($\approx$ 86 000 observations).

# Background and Motivation: GPD

For $X_1, X_2, \ldots$ a seq. of **independent** random variables with common cdf $F$, the limiting distribution, if it exists, as $u \longrightarrow \infty$ of $(X - u | X > u)$ is GPD:

$$G(y) = 1 - \left(1 + \frac{\xi^* y}{\sigma^*}\right)_+^{-1/\xi^*} \tag{1}$$

with $z_+ = \max\{z, 0\}$, $\sigma^* > 0$ and $\xi^*$ are scale and shape parameters, respectively.

# Background and Motivation: GPD

For $X_1, X_2, \ldots$ a seq. of **independent** random variables with common cdf $F$, the limiting distribution, if it exists, as $u \longrightarrow \infty$ of $(X - u|X > u)$ is GPD:

$$G(y) = 1 - \left(1 + \frac{\xi^* y}{\sigma^*}\right)_+^{-1/\xi^*} \tag{1}$$

with $z_+ = \max\{z, 0\}$, $\sigma^* > 0$ and $\xi^*$ are scale and shape parameters, respectively.

Relating to the GEV (Coles *et al.*, 1994),

$$\sigma^* = \sigma + \xi(\mu - u)$$

$$\xi^* = \xi$$

$$\lambda = 1 - \exp\left\{-\frac{1}{N}\left[1 + \xi\left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

with $N$ the number of observations per year, and $\lambda$ the exceedance rate.

# Background and Motivation: GPD

Nonstationarity: Seasonality is often present with meteorological data.

- Stratify by month, and fit models to each month separately.
- Incorporate seasonality into the model (e.g., Coles *et al*. (1994)).

# Background and Motivation: GPD

Nonstationarity: Seasonality is often present with meteorological data.

- Stratify by month, and fit models to each month separately.
- Incorporate seasonality into the model (e.g., Coles *et al*. (1994)).

The former approach is taken in this paper, where the interest is in modeling the temporal dependence through a Markov chain model.

# Markov chain model: Joint density

For independent data, the joint density, $f(y_1, \ldots, y_n)$, can be factored into the product of the marginal distributions. That is,

$$f(y_1, \ldots, y_n) = \prod_{i=1}^{n} f(y_i)$$

# Markov chain model: Joint density

For independent data, the joint density, $f(y_1, \ldots, y_n)$, can be factored into the product of the marginal distributions. That is,

$$f(y_1, \ldots, y_n) = \prod_{i=1}^{n} f(y_i)$$

A Markov chain implies a lag-one temporal dependence so that the joint density is slightly more complicated. Specifically,

$$f(y_1, \ldots, y_n) = \frac{\prod_{i=2}^{n} f(y_{i-1}, y_i)}{\prod_{i=2}^{n-1} f(y_i)} \tag{2}$$

# Markov chain model: Joint density

For example: Take $n = 4$ so that we have the series: $y_1, \ldots, y_4$. Making no assumptions about independence, we have that

$$f(y_1, y_2, y_3, y_4) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_1, y_2) \cdot f(y_4|y_1, y_2, y_3)$$

# Markov chain model: Joint density

For example: Take $n = 4$ so that we have the series: $y_1, \ldots, y_4$. Making no assumptions about independence, we have that

$$f(y_1, y_2, y_3, y_4) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_1, y_2) \cdot f(y_4|y_1, y_2, y_3)$$

For a Markov chain, the dependence is only with the previous time step so that the above reduces to:

$$f(y_1, y_2, y_3, y_4) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_2) \cdot f(y_4|y_3)$$

# Markov chain model: Joint density

For example: Take $n = 4$ so that we have the series: $y_1, \ldots, y_4$. Making no assumptions about independence, we have that

$$f(y_1, y_2, y_3, y_4) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_1, y_2) \cdot f(y_4|y_1, y_2, y_3)$$

For a Markov chain, the dependence is only with the previous time step so that the above reduces to:

$$f(y_1, y_2, y_3, y_4) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_2) \cdot f(y_4|y_3)$$

Using the Bayes formula, this becomes:

$$\begin{aligned} f(y_1, y_2, y_3, y_4) &= f(y_1) \cdot \frac{f(y_1, y_2)}{f(y_1)} \cdot \frac{f(y_2, y_3)}{f(y_2)} \cdot \frac{f(y_3, y_4)}{f(y_3)} \\ &= \frac{f(y_1, y_2) f(y_2, y_3) f(y_3, y_4)}{f(y_2) f(y_3)} \end{aligned}$$

# Markov chain model: Joint density

The numerator of (2) can be modeled using the bivariate extreme-value distribution characterised in sections 3.1 and 3.2 of the paper.

The denominator, which involves only univariate densities, is modeled with the univariate GPD (1).

# Markov chain model: Bivariate GPD

It can be shown that the bivariate distribution function, $G(x_1, x_2)$, for $x_1$ and $x_2$ with standard Fréchet margins is:

$$G(x_1, x_2) = \exp\{-V(x_1, x_2)\}$$

for certain $V(\cdot, \cdot)$ (see paper).

# Markov chain model: Bivariate GPD

It can be shown that the bivariate distribution function, $G(x_1, x_2)$, for $x_1$ and $x_2$ with standard Fréchet margins is:

$$G(x_1, x_2) = \exp\{-V(x_1, x_2)\}$$

for certain $V(\cdot, \cdot)$ (see paper).

One widely used class of families for $G$ is the *logistic* family, where

$$V(x_1, x_2) = \left(x_1^{-1/\alpha} + x_2^{-1/\alpha}\right)^{\alpha}$$

where $x_1, x_2 > 0$ and $\alpha \in (0, 1]$. For $\alpha = 1$, we have **independence**, and as $\alpha \longrightarrow 0$, we have **complete dependence**.

# Bayesian Inference

MLE: Which parameters maximize the (assumed) likelihood of realising the data that are observed?

Bayesian Estimation:

- Assume data are distributed according to [data$|\theta$].
- Assume a prior distribution, [$\theta$], for the unknown paramter(s).
- Use Bayes formula to find the posterior distribution, [$\theta|$data].[1]

$$[\theta|\text{data}] = \frac{[\text{data}|\theta][\theta]}{[\text{data}]}$$

[1] Generally the posterior is intractable, but it is often possible to simulate from it based on the posterior and data likelihood.

# Bayesian Inference

Again, parameters are fit separately for each month, $m$.

Thresholds, $u_m$, are chosen by standard exploratory methods.

No prior information, so non-informative priors are used. To ensure positivity, the scale parameters are modeled as $\sigma_m = \exp(\eta_m)$. Specifically,

$$\pi(\eta_m) \sim N(0, 10\ 000);$$

$$\pi(\xi_m) \sim N(0, 100);$$

$$\pi(\alpha_m) \sim U(0, 1)$$

# Bayesian Inference

Because the posterior distributions are generally intractable, samples from their distributions are simulated using the prior distributions of the previous slide and the likelihoods (data given the parameters).

Predictions are subsequently obtained through functions of the estimated posterior distribution. See Eq. (19) of the paper.

# Results for High Bradfield

- Rapid convergence to apparent stationary parameter distributions was achieved.
- Good mixing properties.
- Dependence parameter[2], $\alpha_m$, in range of 0.3 to 0.45.
- $\xi < 0$ for all months except June.
- Eq. (20) is set equal to $1/r$, and solved for $z_r$ to obtain the $r$-year return level.
- Note that there are two estimates for the return level. The *predictive* estimate accounts for uncertainty in parameter estimation and future observations.

[2]From *logistic* dependence model

# Further model considerations

- Other dependence models besides the logistic to account for skewness. Nothing Gained
- Higher-order Markov models.

  $2^{\text{nd}}$-order model good for

  (i) storm length

  (ii) duration between storms

  No change for return levels