Modelling extreme values of processes observed at irregular time step

Extreme Environmental Events

Nicolas Raillard

Université de Bretagne Occidentale
IFREMER
Brest, France

16th December 2010

Introduction

Objectives

- Spatio-temporal peaks-over-thresholds
- Application: Satellite data, in-situ observations, hindcast data...

Introduction

Objectives

- Spatio-temporal peaks-over-thresholds
- Application: Satellite data, in-situ observations, hindcast data...

Data with irregular sampling?

- Deal with missing observations
- Cope with satellite data
- Generalize to spatial data

Introduction

Objectives

- Spatio-temporal peaks-over-thresholds
- Application: Satellite data, in-situ observations, hindcast data...

Data with irregular sampling?

- Deal with missing observations
- Cope with satellite data
- Generalize to spatial data

Introduction

Objectives

- Spatio-temporal peaks-over-thresholds
- Application: Satellite data, in-situ observations, hindcast data...

Data with irregular sampling?

- Deal with missing observations
- Cope with satellite data
- Generalize to spatial data

Contents

Contents

Univariate probabilistic theory

- Limit theorems for $\frac{\sum_{i=1}^{n} X_i - b_n}{a_n}$: CLT, α-stable distributions...
- For the tails: $S_n \leftrightarrow M_n = \max_{i=1,\dots,n} X_i$
- If $F^n(a_n x + b_n) = \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \rightarrow G(x)$, then G is a max-stable distribution: $G^n(\alpha_n x + \beta_n) = G(x)$
- Fisher-Tippett (1928), Gnedenko (1943): $G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \in \text{GEV}(\mu, \sigma, \xi)$
- Example:
  - Type Weibull: $\xi < 0$ (including Rayleigh...)
  - Heavy tails (Fréchet): $\xi > 0$; Unit Fréchet $\xi = 1$
  - $\xi \rightarrow 0$: $G(x) = \exp(-e^{-\frac{x-\mu}{\sigma}})$ (Gumbel distribution)

Univariate probabilistic theory

- Limit theorems for $\frac{\sum_{i=1}^{n} X_i - b_n}{a_n}$: CLT, α-stable distributions...
- For the tails: $S_n \leftrightarrow M_n = \max_{i=1,\ldots,n} X_i$
- If $F^n(a_n x + b_n) = \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x)$, then G is a max-stable distribution: $G^n(\alpha_n x + \beta_n) = G(x)$
- Fisher-Tippett (1928), Gnedenko (1943): $G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \in \text{GEV}(\mu, \sigma, \xi)$
- Example:
  - Type Weibull: $\xi < 0$ (including Rayleigh...)
  - Heavy tails (Fréchet): $\xi > 0$; Unit Fréchet $\xi = 1$
  - $\xi \to 0$: $G(x) = \exp(-e^{-\frac{x-\mu}{\sigma}})$ (Gumbel distribution)

Univariate probabilistic theory

- Limit theorems for $\frac{\sum_{i=1}^{n} X_i - b_n}{a_n}$: CLT, α-stable distributions...
- For the tails: $S_n \leftrightarrow M_n = \max_{i=1,\ldots,n} X_i$
- If $F^n(a_n x + b_n) = \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x)$, then G is a max-stable distribution: $G^n(\alpha_n x + \beta_n) = G(x)$
- Fisher-Tippett (1928), Gnedenko (1943): $G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi \frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \in \mathrm{GEV}(\mu, \sigma, \xi)$
- Example:
  - Type Weibull: $\xi < 0$ (including Rayleigh...)
  - Heavy tails (Fréchet): $\xi > 0$; Unit Fréchet $\xi = 1$
  - $\xi \to 0$: $G(x) = \exp(-e^{-\frac{x-\mu}{\sigma}})$ (Gumbel distribution)

Univariate probabilistic theory

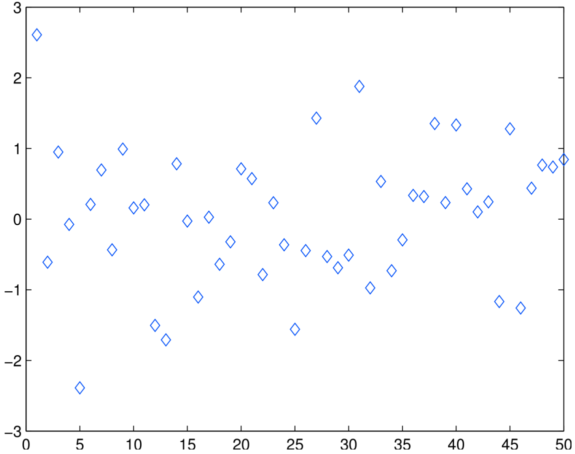
- Limit theorems for $\frac{\sum_{i=1}^{n} X_i - b_n}{a_n}$: CLT, α-stable distributions...
- For the tails: $S_n \leftrightarrow M_n = \max_{i=1,\dots,n} X_i$
- If $F^n(a_n x + b_n) = \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x)$, then G is a max-stable distribution: $G^n(\alpha_n x + \beta_n) = G(x)$
- Fisher-Tippett (1928), Gnedenko (1943):
  $G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi \frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \in \text{GEV}(\mu, \sigma, \xi)$
- Example:
  - Type Weibull: $\xi < 0$ (including Rayleigh...)
  - Heavy tails (Fréchet): $\xi > 0$; Unit Fréchet $\xi = 1$
  - $\xi \to 0$: $G(x) = \exp(-e^{-\frac{x-\mu}{\sigma}})$ (Gumbel distribution)

Univariate probabilistic theory

- Limit theorems for $\frac{\sum_{i=1}^{n} X_i - b_n}{a_n}$: CLT, α-stable distributions...
- For the tails: $S_n \leftrightarrow M_n = \max_{i=1,\dots,n} X_i$
- If $F^n(a_n x + b_n) = \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x)$, then G is a max-stable distribution: $G^n(\alpha_n x + \beta_n) = G(x)$
- Fisher-Tippett (1928), Gnedenko (1943):
  $G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi \frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right] \in \mathrm{GEV}(\mu, \sigma, \xi)$
- Example:
  - Type Weibull: $\xi < 0$ (including Rayleigh...)
  - Heavy tails (Fréchet): $\xi > 0$; Unit Fréchet $\xi = 1$
  - $\xi \to 0$: $G(x) = \exp(-e^{-\frac{x-\mu}{\sigma}})$ (Gumbel distribution)

Statistical modelling of univariate data

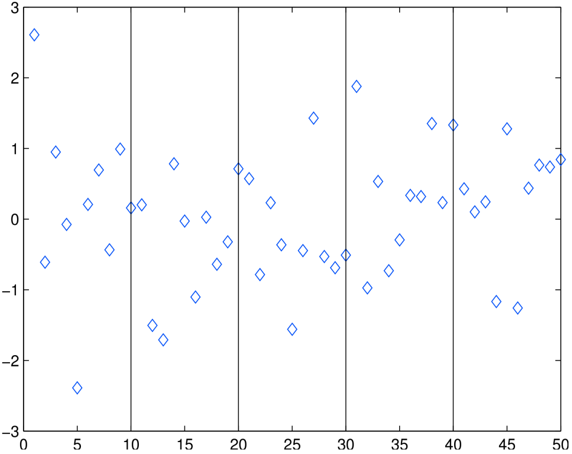
- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$



- **Improvement**: keep all excesses over a high threshold u

Statistical modelling of univariate data

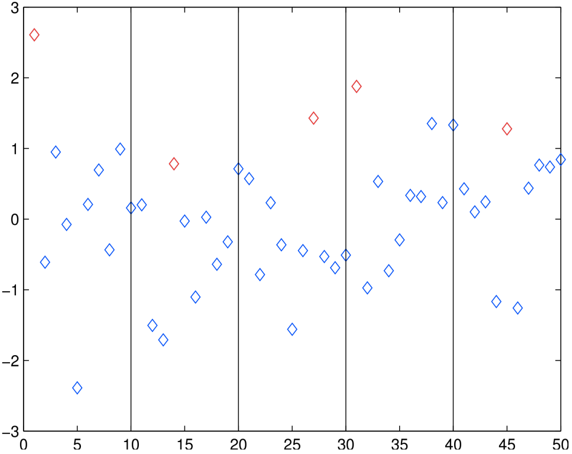
- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$



- **Improvement**: keep all excesses over a high threshold u

Statistical modelling of univariate data

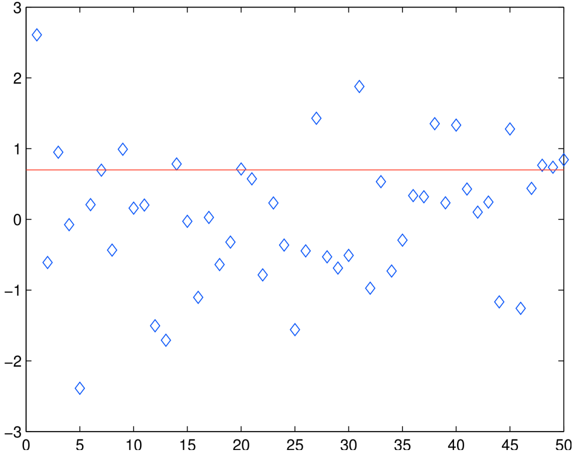
- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$



- **Improvement**: keep all excesses over a high threshold u

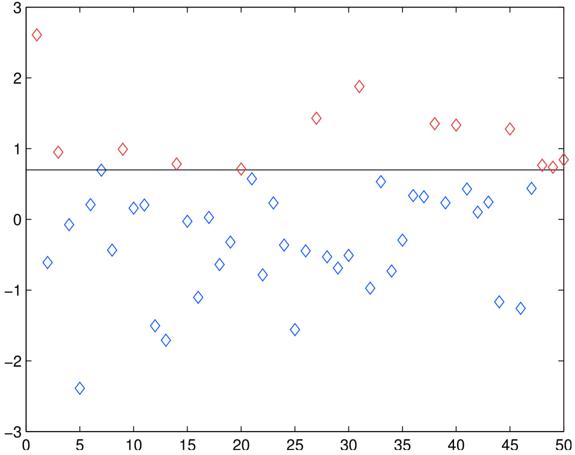
Statistical modelling of univariate data

- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \mathrm{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u



Statistical modelling of univariate data

- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u



Statistical modelling of univariate data

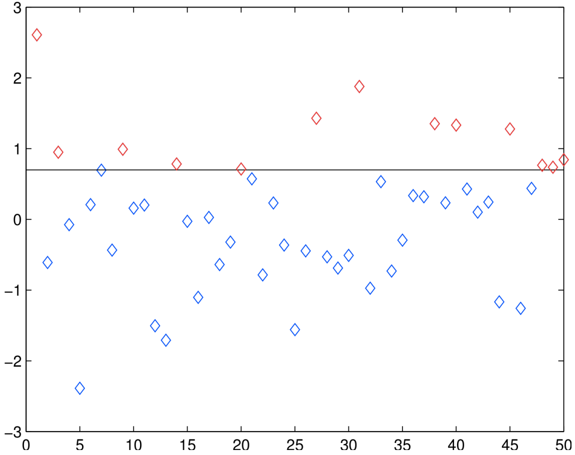
- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u
  - POT: for $x > u$, $F(x) \approx 1 - \lambda(1 + \xi\frac{x-u}{\sigma})^{-1/\xi} \in \text{GPD}(\lambda, \sigma, \xi)$
    $L(x_1, \ldots, x_n; \lambda, \sigma, \xi) = (1-\lambda)^{N_u}\lambda^{n-N_u}\prod_{i/x_i>u} p(x_i; \sigma, \xi)$
    $N_u = |\{i/x_i \leq u\}|$
  - Censoring: If $X_i > u$, $X_i = \widetilde{X}_i$ where $\tilde{X}_i \sim \text{GEV}(\mu, \sigma, \xi)$
    $L(x_1, \ldots, x_n; \mu, \sigma, \xi) = G(u; \mu, \sigma, \xi)^{N_u}\prod_{i/x_i>u}\tilde{p}(x_i; \mu, \sigma, \xi)$.
- **Problem**: dependence between observations $\Rightarrow$ dynamical model on the excesses

Statistical modelling of univariate data

- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u
  - POT: for $x > u$, $F(x) \approx 1 - \lambda(1 + \xi\frac{x-u}{\sigma})^{-1/\xi} \in \text{GPD}(\lambda, \sigma, \xi)$
    $L(x_1, \ldots, x_n; \lambda, \sigma, \xi) = (1-\lambda)^{N_u}\lambda^{n-N_u}\prod_{i/x_i>u} p(x_i; \sigma, \xi)$
    $N_u = |\{i/x_i \leq u\}|$
  - Censoring: If $X_i > u$, $X_i = \widetilde{X}_i$ where $\tilde{X}_i \sim \text{GEV}(\mu, \sigma, \xi)$
    $L(x_1, \ldots, x_n; \mu, \sigma, \xi) = G(u; \mu, \sigma, \xi)^{N_u}\prod_{i/x_i>u} \tilde{p}(x_i; \mu, \sigma, \xi)$.
- **Problem**: dependence between observations $\Rightarrow$ dynamical model on the excesses

Statistical modelling of univariate data

- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u



Statistical modelling of univariate data

- **Bloc maxima** (e.g.: annual maxima) $M_n \sim \text{GEV}(\mu, \sigma, \xi)$: data is wasted
- **Improvement**: keep all excesses over a high threshold u
  - POT: for $x > u$, $F(x) \approx 1 - \lambda(1 + \xi \frac{x-u}{\sigma})^{-1/\xi} \in \text{GPD}(\lambda, \sigma, \xi)$
    $L(x_1, \dots, x_n; \lambda, \sigma, \xi) = (1-\lambda)^{N_u} \lambda^{n-N_u} \prod_{i/x_i>u} p(x_i; \sigma, \xi)$
    $N_u = |\{i/x_i \leq u\}|$
  - Censoring: If $X_i > u$, $X_i = \widetilde{X}_i$ where $\tilde{X}_i \sim \text{GEV}(\mu, \sigma, \xi)$
    $L(x_1, \dots, x_n; \mu, \sigma, \xi) = G(u; \mu, \sigma, \xi)^{N_u} \prod_{i/x_i>u} \tilde{p}(x_i; \mu, \sigma, \xi)$.
- **Problem**: dependence between observations $\Rightarrow$ dynamical model on the excesses

Solutions to time-dependence in POT

2 approaches:

1. **Declustering** (Coles, 2001)
   - Identify Clusters
   - Fit the marginal distribution to maximum excess within each cluster
   - Estimate the extremal index to **correct** the return period
2. **Modelling all exceedances** of the threshold (Smith,1997; Ribatet 2009)
   - Advantages: no need to define a cluster; all excess are kept
   - Drawbacks: Need for a dependence modelling; less usual
   - Usually: Markov Chain model for consecutive observations: need for constant time-step; strong assumption

Solutions to time-dependence in POT

2 approaches:

1. **Declustering** (Coles, 2001)
   - Identify Clusters
   - Fit the marginal distribution to maximum excess within each cluster
   - Estimate the extremal index to **correct** the return period
2. **Modelling all exceedances** of the threshold (Smith,1997; Ribatet 2009)
   - Advantages: no need to define a cluster; all excess are kept
   - Drawbacks: Need for a dependence modelling; less usual
   - Usually: Markov Chain model for consecutive observations: need for constant time-step; strong assumption

Solutions to time-dependence in POT

2 approaches:

1. **Declustering** (Coles, 2001)
   - Identify Clusters
   - Fit the marginal distribution to maximum excess within each cluster
   - Estimate the extremal index to **correct** the return period
2. **Modelling all exceedances** of the threshold (Smith,1997; Ribatet 2009)
   - Advantages: no need to define a cluster; all excess are kept
   - Drawbacks: Need for a dependence modelling; less usual
   - Usually: Markov Chain model for consecutive observations: need for constant time-step; strong assumption

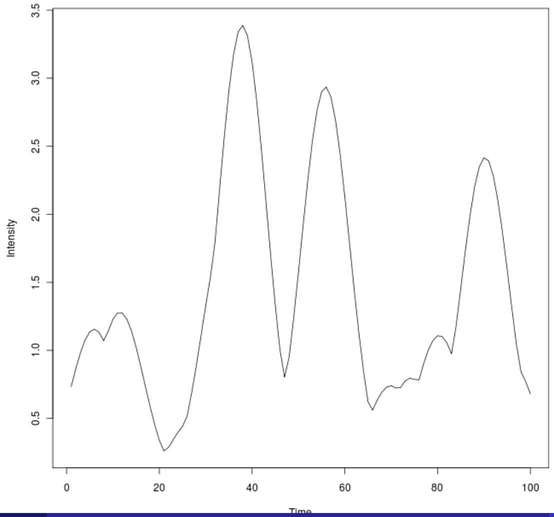
Modelling dependence
Smith's max-stable process

Extension of **max-stable concept for processes** (de Haan, 1986: Smith, 1990), with unit Fréchet marginal distributions.

$$\forall t \in \mathbb{R},\ Z(t) = \max_{i=1,\dots} \{\lambda_i f(t, Y_i)\}$$

with: $f(s,t) = \varphi(s-t)$, φ p.d.f of $\mathcal{N}(0,1)$, (λ_i, Y_i) points of a Poisson process with intensity $\lambda^{-2} d\lambda / ds$

Modelling dependence
Smith's max-stable process



Modelling dependence
Smith's max-stable process

Extension of **max-stable concept for processes** (de Haan, 1986: Smith, 1990), with unit Fréchet marginal distributions.

$$\forall t \in \mathbb{R},\ Z(t) = \max_{i=1,\dots} \{\lambda_i f(t, Y_i)\}$$

with: $f(s,t) = \varphi(s-t)$, φ p.d.f of $\mathcal{N}(0,1)$, (λ_i, Y_i) points of a Poisson process with intensity $\lambda^{-2} d\lambda/ds$

We have **only access to the bivariate c.d.f**:

$$\mathbb{P}\{Z(s) \le z_i,\ Z(t) \le z_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_j}{z_i}\right) - \frac{1}{z_j}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where Φ is $\mathcal{N}(0,1)$ c.d.f and $a^2 = \frac{s-t}{\nu}$

Modelling dependence

Smith's max-stable process

Extension of **max-stable concept for processes** (de Haan, 1986: Smith, 1990), with unit Fréchet marginal distributions.

$$\forall t \in \mathbb{R},\ Z(t) = \max_{i=1,\ldots} \{\lambda_i f(t, Y_i)\}$$

with: $f(s,t) = \varphi(s-t)$, φ p.d.f of $\mathcal{N}(0,1)$, (λ_i, Y_i) points of a Poisson process with intensity $\lambda^{-2} d\lambda / ds$

We have **only access to the bivariate c.d.f**:

$$\mathbb{P}\{Z(s) \leq z_i,\ Z(t) \leq z_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_j}{z_i}\right) - \frac{1}{z_j}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where Φ is $\mathcal{N}(0,1)$ c.d.f and $a^2 = \frac{s-t}{\nu}$

$\Rightarrow$ use of a **composite likelihood method** (Varin, 2008):

$$CL(z_{t_1}, \ldots, z_{t_n}; \theta) = \prod_{i=1}^{n-K} \prod_{j=i+1}^{i+K} p(z_{t_i}, z_{t_j}; \theta)$$

Application to peaks over a high threshold
Model

Model for excess

- Observations $\{X_{t_i}\}_{i=1,\dots,n}$
- If $X_{t_i} > u$, $X_{t_i} = \widetilde{X}_{t_i}$ where $\widetilde{X}_t$ is a max-stable process:
  - $\widetilde{X}_t \sim G(.;\mu,\sigma,\xi) \in GEV(\mu,\sigma,\xi)$
    $\Rightarrow Z_t = -1/\log G(X_t;\mu,\sigma,\xi) \sim$ Fréchet(1)
  - $\{Z_t\}$ is a Smith process
- Thus, for $x_i > u$ and $x_j > u$:

$$\mathbb{P}\{X_{t_i} \leq x_i,\ X_{t_j} \leq x_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_j}{z_i}\right) - \frac{1}{z_j}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where $a = \frac{t_i - t_j}{\nu}$ and $z_i = -\frac{1}{\log G(x_i;\mu,\sigma,\xi)}$

Application to peaks over a high threshold
Model

Model for excess

- Observations $\{X_{t_i}\}_{i=1,\dots,n}$
- If $X_{t_i} > u$, $X_{t_i} = \widetilde{X}_{t_i}$ where $\widetilde{X}_t$ is a max-stable process:
  - $\widetilde{X}_t \sim G(.;\mu,\sigma,\xi) \in GEV(\mu,\sigma,\xi)$
    $\Rightarrow Z_t = -1/\log G(X_t;\mu,\sigma,\xi) \sim$ Fréchet(1)
  - $\{Z_t\}$ is a Smith process
- Thus, for $x_i > u$ and $x_j > u$:

$$\mathbb{P}\{X_{t_i} \le x_i,\ X_{t_j} \le x_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_j}{z_i}\right) -\frac{1}{z_j}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where $a = \frac{t_i - t_j}{\nu}$ and $z_i = -\frac{1}{\log G(x_i;\mu,\sigma,\xi)}$

Application to peaks over a high threshold
Model

Model for excess

- Observations $\{X_{t_i}\}_{i=1,\dots,n}$
- If $X_{t_i} > u$, $X_{t_i} = \widetilde{X}_{t_i}$ where $\widetilde{X}_t$ is a max-stable process:
  - $\widetilde{X}_t \sim G(.;\mu,\sigma,\xi) \in GEV(\mu,\sigma,\xi)$
    $\Rightarrow Z_t = -1/\log G(X_t;\mu,\sigma,\xi) \sim \text{Fréchet}(1)$
  - $\{Z_t\}$ is a Smith process
- Thus, for $x_i > u$ and $x_j > u$:

$$\mathbb{P}\{X_{t_i} \leq x_i,\ X_{t_j} \leq x_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_j}{z_i}\right) - \frac{1}{z_j}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where $a = \frac{t_i - t_j}{\nu}$ and $z_i = -\frac{1}{\log G(x_i;\mu,\sigma,\xi)}$

Application to peaks over a high threshold
Model

Model for excess

- Observations $\{X_{t_i}\}_{i=1,\dots,n}$
- If $X_{t_i} > u$, $X_{t_i} = \widetilde{X}_{t_i}$ where $\widetilde{X}_t$ is a max-stable process:
  - $\widetilde{X}_t \sim G(.;\mu,\sigma,\xi) \in GEV(\mu,\sigma,\xi)$
    $\Rightarrow Z_t = -1/\log G(X_t;\mu,\sigma,\xi) \sim$ Fréchet(1)
  - $\{Z_t\}$ is a Smith process
- Thus, for $x_i > u$ and $x_j > u$:

$$\mathbb{P}\{X_{t_i} \le x_i,\ X_{t_j} \le x_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_j}{z_i}\right) -\frac{1}{z_j}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where $a = \frac{t_i - t_j}{\nu}$ and $z_i = -\frac{1}{\log G(x_i;\mu,\sigma,\xi)}$

Application to peaks over a high threshold

Model

Model for excess

- Observations $\{X_{t_i}\}_{i=1,\dots,n}$
- If $X_{t_i} > u$, $X_{t_i} = \widetilde{X}_{t_i}$ where $\widetilde{X}_t$ is a max-stable process:
  - $\widetilde{X}_t \sim G(.;\mu,\sigma,\xi) \in GEV(\mu,\sigma,\xi)$
    $\Rightarrow Z_t = -1/\log G(X_t;\mu,\sigma,\xi) \sim \text{Fréchet}(1)$
  - $\{Z_t\}$ is a Smith process
- Thus, for $x_i > u$ and $x_j > u$:

$$\mathbb{P}\{X_{t_i} \leq x_i,\ X_{t_j} \leq x_j\} = \exp\left[-\frac{1}{z_i}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_j}{z_i}\right) -\frac{1}{z_j}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_i}{z_j}\right)\right]$$

where $a = \frac{t_i - t_j}{\nu}$ and $z_i = -\frac{1}{\log G(x_i;\mu,\sigma,\xi)}$

Contents

Properties of the CL estimator
Properties

- The Smith process is ergodic (Stoev, 2008)
- Under study: consistency of the uncensored one-lagged CL estimator (i.e. with $K = 1$)
- Perspectives:
  - Obtaining asymptotic normality
  - Choice of K?
  - Results with censoring
- Can the behaviour be guessed by simulation?
  - Convergence
  - Influence of parameters values
  - Influence of K
  - Robustness to missing values and censoring

Properties of the CL estimator
Properties

- The Smith process is ergodic (Stoev, 2008)
- Under study: consistency of the uncensored one-lagged CL estimator (i.e. with $K = 1$)
- Perspectives:
  - Obtaining asymptotic normality
  - Choice of K?
  - Results with censoring
- Can the behaviour be guessed by simulation?
  - Convergence
  - Influence of parameters values
  - Influence of K
  - Robustness to missing values and censoring

Properties of the CL estimator
Properties

- The Smith process is ergodic (Stoev, 2008)
- Under study: consistency of the uncensored one-lagged CL estimator (i.e. with $K = 1$)
- Perspectives:
  - Obtaining asymptotic normality
  - Choice of K?
  - Results with censoring
- Can the behaviour be guessed by simulation?
  - Convergence
  - Influence of parameters values
  - Influence of K
  - Robustness to missing values and censoring

Properties of the CL estimator

Properties

- The Smith process is ergodic (Stoev, 2008)
- Under study: consistency of the uncensored one-lagged CL estimator (i.e. with $K = 1$)
- Perspectives:
  - Obtaining asymptotic normality
  - Choice of K?
  - Results with censoring
- Can the behaviour be guessed by simulation?
  - Convergence
  - Influence of parameters values
  - Influence of K
  - Robustness to missing values and censoring

Properties of the CL estimator
Properties

- The Smith process is ergodic (Stoev, 2008)
- Under study: consistency of the uncensored one-lagged CL estimator (i.e. with $K = 1$)
- Perspectives:
  - Obtaining asymptotic normality
  - Choice of K?
  - Results with censoring
- Can the behaviour be guessed by simulation?
  - Convergence
  - Influence of parameters values
  - Influence of K
  - Robustness to missing values and censoring

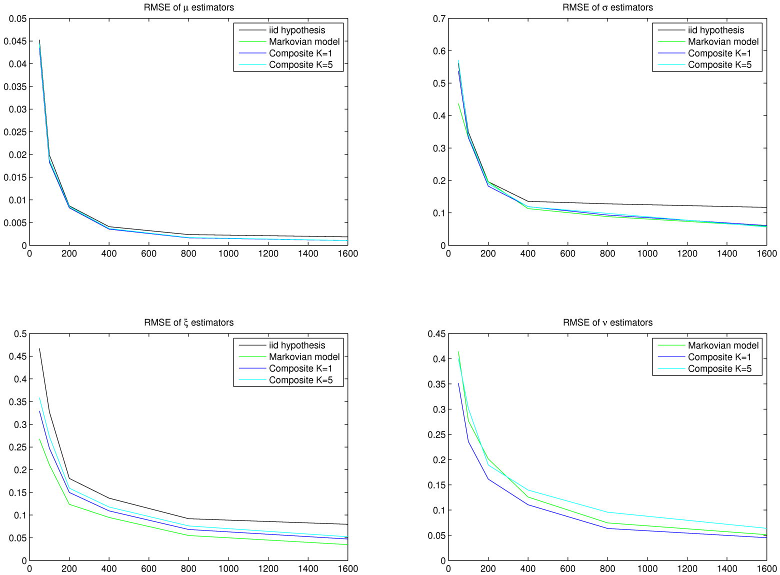


Figure: RMSE of the different estimators

Robustness

Question: Is this model able to catch extremal properties of usual time series models?

- 10 years return level: level which is exceeded on average once every 10 years
- Frequency and length of extremal events
- Sum of the values above the threshold

Robustness

Question: Is this model able to catch extremal properties of usual time series models?

- 10 years return level: level which is exceeded on average once every 10 years
- Frequency and length of extremal events
- Sum of the values above the threshold

Method:

- Simulation of $X_t = \varphi X_{t-1} + \mathcal{N}\left(0, \sqrt{1-\varphi^2}\right)$ of length 100 years (with 1 obs./day)
- Fitting the model
- 100 times:
  - Simulation of the fitted model over 10 years, and computation of the statistics
  - Simulation of the $AR(1)$ over 10 years, and computation of the statistics

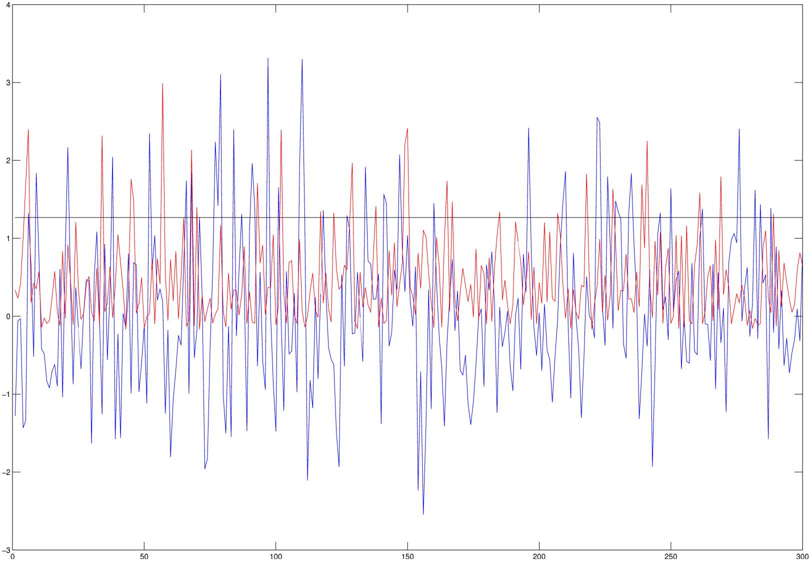


Figure: Example of simulated data (Blue) vs. fitted model (Red) and selected threshold (Black)

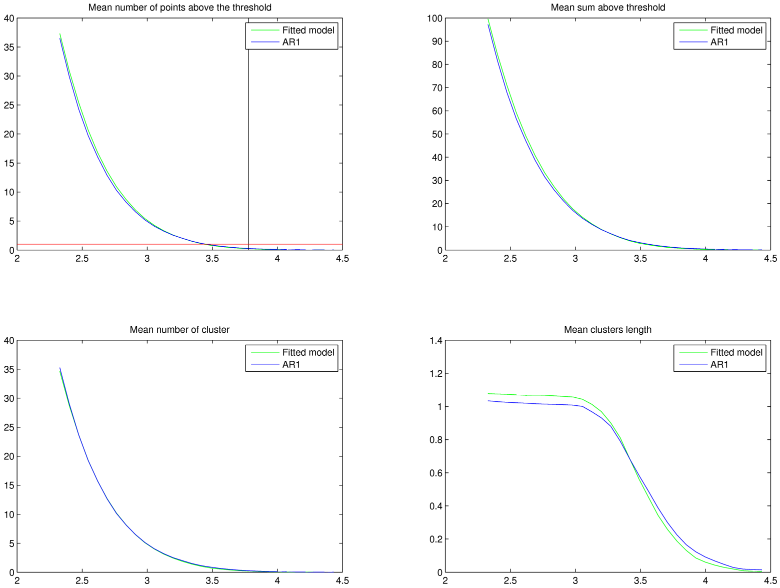


Figure: Statistics computed for an AR(1) process with $\varphi = 0.2$

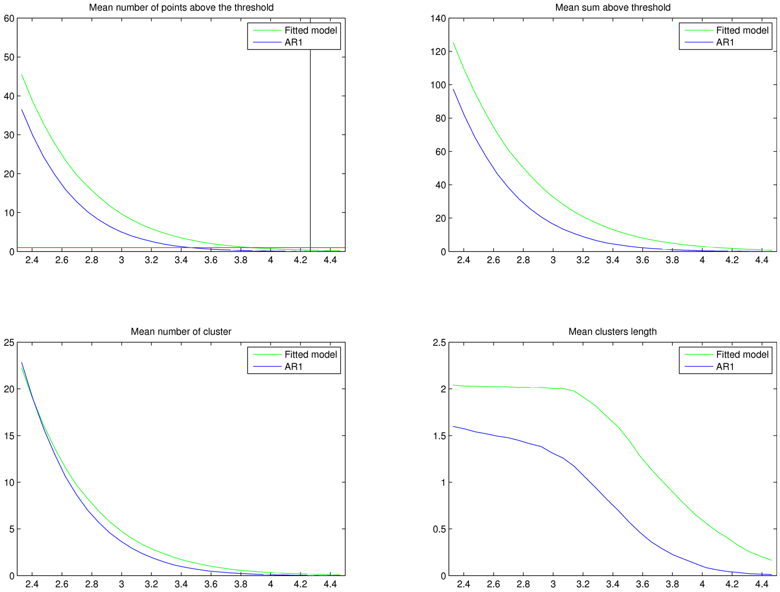


Figure: Statistics computed for an AR(1) process with $\varphi = 0.8$

Contents

Conclusion

… and prospects

- We obtained a flexible model that can account for irregular data spacing and missing values
- Ongoing: application to satellite and in-situ data
- Perspectives:
    - Theoretical results :asymptotic normality, choice of K, influence of censoring…
    - Extensions: space-time modelling…

D. R. Cox.
A note on pseudolikelihood constructed from marginal densities.
Biometrika, 92:729–737, 2004.

L. de Haan.
A spectral representation for max-stable processes.
Ann. Probab., 12(4):1194–1204, 1984.

S. A. Padoan, M. Ribatet, and S.A. Sisson.
Likelihood-based inference for max-stable processes.
In review, 2009.

M. Ribatet, T. B. M. J. Ouarda, E. Sauquet, and J. M. Gresillon.
Modeling all exceedances above a threshold using an extremal dependence structure: Inferences on several flood characteristics.
Water Resources Research, 45(3), March 2009.

Richard L. Smith, Jonathan A. Tawn, and Stuart G. Coles.
Markov chain models for threshold exceedances.
Biometrika, 84(2):249–268, 1997.

R.L. Smith.
Max-stable processes and spatial extremes.
1990.

Cristiano Varin.
On composite marginal likelihoods.
AStA Adv. Stat. Anal., 92(1):1–28, 2008.