

Predicting extreme wind speeds

Lee Fawcett, Amy Green, Dave Walshaw

March 2017

Motivation



Motivation



Motivation

Estimation of **return levels** is usually the key aim of any extreme value analysis.

- Examples: *British Standards Institute* – 50 year wind speed; *Dutch Delta Commission* – 10,000 year sea-surge; *Office for Nuclear Regulation* – also 10,000 year safety specifications...
- Large / very large uncertainty on estimates
- Need to get the most from our data (and other sources!):
 - ▶ Threshold-based analyses *c.f.* block maxima analyses
 - ▶ Accounting for extremal dependence *c.f.* temporal filtering
 - ▶ Re-combining seasonal analyses *c.f.* single-season approach
 - ▶ Making use of **prior information**
- How do we report our inferences back to practitioners?
- How can practitioners use cutting-edge methods?

Motivation

Estimation of **return levels** is usually the key aim of any extreme value analysis.

- Examples: *British Standards Institute* – 50 year wind speed; *Dutch Delta Commission* – 10,000 year sea-surge; *Office for Nuclear Regulation* – also 10,000 year safety specifications...
- Large / very large uncertainty on estimates
- Need to get the most from our data (and other sources!):
 - ▶ Threshold-based analyses *c.f.* block maxima analyses
 - ▶ Accounting for extremal dependence *c.f.* temporal filtering
 - ▶ Re-combining seasonal analyses *c.f.* single-season approach
 - ▶ Making use of **prior information**
- How do we report our inferences back to practitioners?
- How can practitioners use cutting-edge methods?

Motivation

Estimation of **return levels** is usually the key aim of any extreme value analysis.

- Examples: *British Standards Institute* – 50 year wind speed; *Dutch Delta Commission* – 10,000 year sea-surge; *Office for Nuclear Regulation* – also 10,000 year safety specifications...
- Large / very large uncertainty on estimates
- Need to get the most from our data (and other sources!):
 - ▶ Threshold-based analyses *c.f.* block maxima analyses
 - ▶ Accounting for extremal dependence *c.f.* temporal filtering
 - ▶ Re-combining seasonal analyses *c.f.* single-season approach
 - ▶ Making use of **prior information**
- How do we report our inferences back to practitioners?
- How can practitioners use cutting-edge methods?

Motivation

Estimation of **return levels** is usually the key aim of any extreme value analysis.

- Examples: *British Standards Institute* – 50 year wind speed; *Dutch Delta Commission* – 10,000 year sea-surge; *Office for Nuclear Regulation* – also 10,000 year safety specifications...
- Large / very large uncertainty on estimates
- Need to get the most from our data (and other sources!):
 - ▶ Threshold-based analyses *c.f.* block maxima analyses
 - ▶ Accounting for extremal dependence *c.f.* temporal filtering
 - ▶ Re-combining seasonal analyses *c.f.* single-season approach
 - ▶ Making use of **prior information**
- How do we report our inferences back to practitioners?
- How can practitioners use cutting-edge methods?

Motivation

Estimation of **return levels** is usually the key aim of any extreme value analysis.

- Examples: *British Standards Institute* – 50 year wind speed; *Dutch Delta Commission* – 10,000 year sea-surge; *Office for Nuclear Regulation* – also 10,000 year safety specifications...
- Large / very large uncertainty on estimates
- Need to get the most from our data (and other sources!):
 - ▶ Threshold-based analyses *c.f.* block maxima analyses
 - ▶ Accounting for extremal dependence *c.f.* temporal filtering
 - ▶ Re-combining seasonal analyses *c.f.* single-season approach
 - ▶ Making use of **prior information**
- How do we report our inferences back to practitioners?
- How can practitioners use cutting-edge methods?

Examples

- Wave heights at Seven stones in the English Channel: (14.2, 19.3)m for z_{50} (Davison and Smith, 1990)
- Flows at the River Nidd: (270, 1250)m³s⁻¹ for z_{100} (NERC Flood Studies Report, 1995)
- Hurricane-induced wave-heights at Grand Isle LA: (10, 22.5)m for z_{105} (Northrop and Jonathan, 2011)
- Sea-surge at Newlyn: (0.7, 6.5 $\rightarrow$ $\sim$ 30)m for z_{1000} (Fawcett and Walshaw, 2012)
- Rainfall in SW England: (114, 223)mm, changing to (128, 190)mm with informative priors, for z_{100} (Smith, 2005)

Structure of (the rest of) this talk

- ① Preliminary analysis of wind speed data
- ② Simulation study
 - ▶ Predictive return level versus other return level posterior summaries
 - ▶ Chains with different dependence structures
 - ▶ Mis-specification of dependence structure
 - ▶ Comparison to inferences using temporal filtering

Structure of (the rest of) this talk

- ① Preliminary analysis of wind speed data
- ② Simulation study
 - ▶ Predictive return level versus other return level posterior summaries
 - ▶ Chains with different dependence structures
 - ▶ Mis-specification of dependence structure
 - ▶ Comparison to inferences using temporal filtering

Modelling extremes: Starting point

Let $\{X_n\}$ denote a stationary sequence of random variables with common distribution function F , and let $M_n = \max\{X_1, \dots, X_n\}$.

It is typically the case that, as $n \rightarrow \infty$,

$$\Pr(M_n \leq x) \approx F^{n\theta}(x), \quad (1)$$

where $\theta \in (0, 1)$ is known as the **extremal index**; see, for e.g., [Leadbetter & Rootzén \(1988\)](#).

As $\theta \rightarrow 0$ there is increasing dependence in the extremes of the process; for an independent process, $\theta = 1$.

Modelling extremes: Block maxima

The **Generalised Extreme Value** (GEV) distribution is the limiting model for F^n , with distribution function (d.f.)

$$\mathcal{G}(y; \mu, \varsigma, \xi) = \begin{cases} \exp \left[- (1 + \xi(y - \mu)/\varsigma)^{-1/\xi} \right], & \xi \neq 0 \\ \exp \left[-\exp \left(-(y - \mu)/\varsigma \right) \right] & \xi = 0, \end{cases}$$

defined on $\{y : 1 + \xi(y - \mu)/\varsigma > 0\}$, where $-\infty < \mu < \infty$, $\varsigma > 0$ and $-\infty < \xi < \infty$ are parameters of location, scale and shape.

- It can be shown that $F^{n\theta}$ is also GEV with d.f. $\mathcal{G}(y; \mu^*, \varsigma^*, \xi)$, provided long-range dependence is negligible¹
- In practice, the GEV is used to model maxima over some convenient calendar unit – usually years

¹Technically, this requires **Leadbetter's $D(u_n)$ condition** to hold; see **Leadbetter et al. (1983)**.

Modelling extremes: Threshold excesses

Pickands (1975) showed that for $\theta = 1$ and u large, $(X - u | X > u)$ follows a **Generalised Pareto** (GP) distribution with d.f.

$$\mathcal{H}(y) = \begin{cases} 1 - (1 + \xi y / \sigma)^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp[-y/\sigma] & \xi = 0, \end{cases}$$

defined on $\{y : y > 0 \text{ and } (1 + \xi y / \sigma) > 0\}$, where $\sigma = \varsigma + \xi(u - \mu) > 0$ and ξ are the scale and shape.

- Simple graphical diagnostics (not discussed here) available for selection of threshold u
- Extremal index θ can't be absorbed into the parameters of the GPD; careful consideration of extremal dependence required

Modelling extremes: Return level estimation

The **r -year return level z_r** can then be obtained by inversion of $\mathcal{G}(z_r)$ or $\mathcal{H}^\theta(z_r)$.

For threshold excesses, on equating to $1 - r^{-1}$ this gives

$$z_r = \begin{cases} u + \sigma \xi^{-1} \left[(\lambda_u^{-1} w_r)^{-\xi} - 1 \right] & \xi \neq 0 \\ u - \sigma \log (\lambda_u^{-1} w_r) & \xi = 0, \end{cases}$$

where $w_r = 1 - [1 - (rn_y)^{-1}]^{1/\theta}$, λ_u is the rate of threshold excess and n_y is the (average) number of observations per year.

For seasonally varying extremes, solve

$$\prod_{m=1}^M \mathcal{H}_m^{\theta_m}(z_r) = 1 - r^{-1}, \quad (2)$$

where H_m is the GP distribution function in season m with parameter set $(\lambda_{u_m}, \sigma_m, \xi_m)$, and θ_m is the extremal index in season m .

Modelling extremes: Return level estimation

The **r -year return level z_r** can then be obtained by inversion of $\mathcal{G}(z_r)$ or $\mathcal{H}^\theta(z_r)$.

For threshold excesses, on equating to $1 - r^{-1}$ this gives

$$z_r = \begin{cases} u + \sigma \xi^{-1} \left[(\lambda_u^{-1} w_r)^{-\xi} - 1 \right] & \xi \neq 0 \\ u - \sigma \log(\lambda_u^{-1} w_r) & \xi = 0, \end{cases}$$

where $w_r = 1 - [1 - (rn_y)^{-1}]^{1/\theta}$, λ_u is the rate of threshold excess and n_y is the (average) number of observations per year.

For seasonally varying extremes, solve

$$\prod_{m=1}^M \mathcal{H}_m^{\theta_m}(z_r) = 1 - r^{-1}, \quad (2)$$

where H_m is the GP distribution function in season m with parameter set $(\lambda_{u_m}, \sigma_m, \xi_m)$, and θ_m is the extremal index in season m .

Modelling extremes: Return level estimation

The **r -year return level z_r** can then be obtained by inversion of $\mathcal{G}(z_r)$ or $\mathcal{H}^\theta(z_r)$.

For threshold excesses, on equating to $1 - r^{-1}$ this gives

$$z_r = \begin{cases} u + \sigma \xi^{-1} \left[(\lambda_u^{-1} w_r)^{-\xi} - 1 \right] & \xi \neq 0 \\ u - \sigma \log(\lambda_u^{-1} w_r) & \xi = 0, \end{cases}$$

where $w_r = 1 - [1 - (rn_y)^{-1}]^{1/\theta}$, λ_u is the rate of threshold excess and n_y is the (average) number of observations per year.

For seasonally varying extremes, solve

$$\prod_{m=1}^M \mathcal{H}_m^{\theta_m}(z_r) = 1 - r^{-1}, \quad (2)$$

where H_m is the GP distribution function in season m with parameter set $(\lambda_{u_m}, \sigma_m, \xi_m)$, and θ_m is the extremal index in season m .

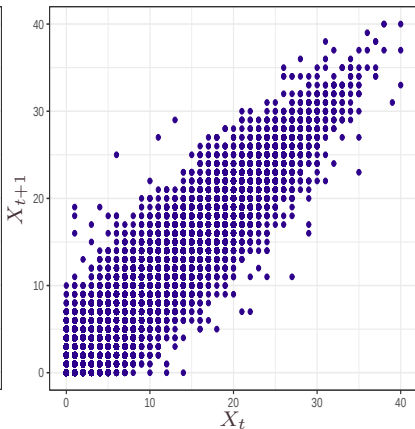
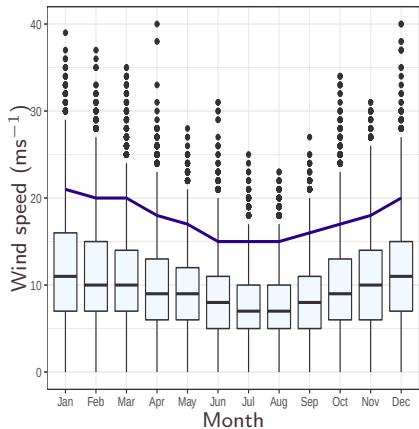
- Hourly gust wind speed maxima observed at 8 locations in southwest England (obtained from NOAA:
<https://data.noaa.gov/>)
- Mix of series lengths: start date – January 1st 1967, end date – December 31st 2015
- All sites close to the planned nuclear development at Hinkley Point
- Clear seasonal (monthly) variations and first-order serial correlation
- For now, we focus on Yeovilton only

- Hourly gust wind speed maxima observed at 8 locations in southwest England (obtained from NOAA:
<https://data.noaa.gov/>)
- Mix of series lengths: start date – January 1st 1967, end date – December 31st 2015
- All sites close to the planned nuclear development at Hinkley Point
- Clear seasonal (monthly) variations and first-order serial correlation
- For now, we focus on Yeovilton only

- Hourly gust wind speed maxima observed at 8 locations in southwest England (obtained from NOAA:
<https://data.noaa.gov/>)
- Mix of series lengths: start date – January 1st 1967, end date – December 31st 2015
- All sites close to the planned nuclear development at Hinkley Point
- Clear seasonal (monthly) variations and first-order serial correlation
- For now, we focus on Yeovilton only

- Hourly gust wind speed maxima observed at 8 locations in southwest England (obtained from NOAA:
<https://data.noaa.gov/>)
- Mix of series lengths: start date – January 1st 1967, end date – December 31st 2015
- All sites close to the planned nuclear development at Hinkley Point
- Clear seasonal (monthly) variations and first-order serial correlation
- For now, we focus on Yeovilton only

Data



Modelling template

- Use the GP distribution to model excesses over high thresholds (using standard graphical diagnostics to select these thresholds)
- Use extremes from all seasons in a **separate seasons** approach to modelling
- Model the dependence between consecutive extremes and use the **extremal index** to aid return level estimation
- Incorporate external information through **informative prior distributions**
- Report the **posterior predictive return level** as perhaps the most useful single number summary of our inference for z_r

Dependence structure: Possible approaches

We want an estimate of the extremal index θ to plug in to our expression for z_r (Eq. 2). Two main approaches:

- 1 Direct estimation using, for example: (i) runs, (ii) blocks, (iii) Gomes, (iv) Süveges, (v) intervals,... estimators
- 2 Model the temporal evolution of the process using, for example, an extreme value Markov process or an AR(1) process; then estimate θ from the estimates of the dependence parameter(s) in the fitted process

Dependence structure: Possible approaches

We want an estimate of the extremal index θ to plug in to our expression for z_r (Eq. 2). Two main approaches:

- 1 Direct estimation using, for example: (i) runs, (ii) blocks, (iii) Gomes, (iv) Süveges, (v) intervals,... estimators
- 2 Model the temporal evolution of the process using, for example, an extreme value Markov process or an AR(1) process; then estimate θ from the estimates of the dependence parameter(s) in the fitted process

Dependence structure: Possible approaches

We want an estimate of the extremal index θ to plug in to our expression for z_r (Eq. 2). Two main approaches:

- 1 Direct estimation using, for example: (i) runs, (ii) blocks, (iii) Gomes, (iv) Süveges, (v) intervals,... estimators
- 2 Model the temporal evolution of the process using, for example, an extreme value Markov process or an AR(1) process; then estimate θ from the estimates of the dependence parameter(s) in the fitted process

Dependence structure: Possible approaches

We want an estimate of the extremal index θ to plug in to our expression for z_r (Eq. 2). Two main approaches:

- 1 Direct estimation using, for example: (i) runs, (ii) blocks, (iii) Gomes, (iv) **Süveges**, (v) intervals,... estimators
- 2 Model the temporal evolution of the process using, for example, an **extreme value Markov process** or an AR(1) process; then estimate θ from the estimates of the dependence parameter(s) in the fitted process

Dependence structure: Extreme value Markov process

Assuming a first-order Markov structure for our monthly wind speed processes enables us to write down the likelihood as:

$$\prod_{i=1}^{n-1} f(x_i, x_{i+1}; \psi) \bigg/ \prod_{i=2}^{n-1} f(x_i; \psi), \quad m = 1, \dots, 12. \quad (3)$$

- Transformation from GP to unit Fréchet margins gives a range of bivariate models to use for the numerator in Eq. 3
- Most commonly-used is the **logistic family**:

$$G(x_i, x_{i+1}) = \exp \left\{ - \left(x_i^{-1/\alpha} + x_{i+1}^{-1/\alpha} \right)^\alpha \right\}, \quad (4)$$

$$\alpha \in [0, 1)$$

- Differentiation of RHS of Eq. 4, with careful censoring, gives pairwise contributions to the numerator in Eq. 3
- Univariate contributions to the denominator are given through the GP distribution

Dependence structure: Extreme value Markov process

Assuming a first-order Markov structure for our monthly wind speed processes enables us to write down the likelihood as:

$$\prod_{i=1}^{n-1} f(x_i, x_{i+1}; \psi) \bigg/ \prod_{i=2}^{n-1} f(x_i; \psi), \quad m = 1, \dots, 12. \quad (3)$$

- Transformation from GP to unit Fréchet margins gives a range of bivariate models to use for the numerator in [Eq. 3](#)
- Most commonly-used is the **logistic family**:

$$G(x_i, x_{i+1}) = \exp \left\{ - \left(x_i^{-1/\alpha} + x_{i+1}^{-1/\alpha} \right)^\alpha \right\}, \quad (4)$$

$$\alpha \in [0, 1)$$

- Differentiation of RHS of [Eq. 4](#), with careful censoring, gives pairwise contributions to the numerator in [Eq. 3](#)
- Univariate contributions to the denominator are given through the GP distribution

Dependence structure: Extreme value Markov process

Assuming a first-order Markov structure for our monthly wind speed processes enables us to write down the likelihood as:

$$\prod_{i=1}^{n-1} f(x_i, x_{i+1}; \psi) \bigg/ \prod_{i=2}^{n-1} f(x_i; \psi), \quad m = 1, \dots, 12. \quad (3)$$

- Transformation from GP to unit Fréchet margins gives a range of bivariate models to use for the numerator in [Eq. 3](#)
- Most commonly-used is the **logistic family**:

$$G(x_i, x_{i+1}) = \exp \left\{ - \left(x_i^{-1/\alpha} + x_{i+1}^{-1/\alpha} \right)^\alpha \right\}, \quad (4)$$

$$\alpha \in [0, 1)$$

- Differentiation of RHS of [Eq. 4](#), with careful censoring, gives pairwise contributions to the numerator in [Eq. 3](#)
- Univariate contributions to the denominator are given through the GP distribution

Dependence structure: Extreme value Markov process

Assuming a first-order Markov structure for our monthly wind speed processes enables us to write down the likelihood as:

$$\prod_{i=1}^{n-1} f(x_i, x_{i+1}; \psi) \bigg/ \prod_{i=2}^{n-1} f(x_i; \psi), \quad m = 1, \dots, 12. \quad (3)$$

- Transformation from GP to unit Fréchet margins gives a range of bivariate models to use for the numerator in [Eq. 3](#)
- Most commonly-used is the **logistic family**:

$$G(x_i, x_{i+1}) = \exp \left\{ - \left(x_i^{-1/\alpha} + x_{i+1}^{-1/\alpha} \right)^\alpha \right\}, \quad (4)$$

$$\alpha \in [0, 1)$$

- Differentiation of RHS of [Eq. 4](#), with careful censoring, gives pairwise contributions to the numerator in [Eq. 3](#)
- Univariate contributions to the denominator are given through the GP distribution

Dependence structure: Extreme value Markov process

Assuming a first-order Markov structure for our monthly wind speed processes enables us to write down the likelihood as:

$$\prod_{i=1}^{n-1} f(x_i, x_{i+1}; \psi) \bigg/ \prod_{i=2}^{n-1} f(x_i; \psi), \quad m = 1, \dots, 12. \quad (3)$$

- Transformation from GP to unit Fréchet margins gives a range of bivariate models to use for the numerator in [Eq. 3](#)
- Most commonly-used is the **logistic family**:

$$G(x_i, x_{i+1}) = \exp \left\{ - \left(x_i^{-1/\alpha} + x_{i+1}^{-1/\alpha} \right)^\alpha \right\}, \quad (4)$$

$$\alpha \in [0, 1)$$

- Differentiation of RHS of [Eq. 4](#), with careful censoring, gives pairwise contributions to the numerator in [Eq. 3](#)
- Univariate contributions to the denominator are given through the GP distribution

Dependence structure: Extreme value Markov process

Define (arbitrarily) x_n such that $F^n(x_n) = 1/2$ in Eq. 1.

Then, using Eq. 1, we can define

$$\theta_n = -\frac{\log \Pr(\max\{X_1, \dots, X_n\} \leq x_n)}{\log 2}, \quad (5)$$

and so $\theta_n \rightarrow \theta$ as $n \rightarrow \infty$.

We can use Eq. 5 to investigate the relationship between θ and α in the logistic model:

- Simulate N first-order Markov chains, each of length n , with logistic dependence α
- The probability in the numerator of Eq. 5 is estimated as the proportion of simulated chains whose maximum does not exceed x_n
- We find values for θ_n for $\alpha = \{0.05, 0.10, \dots, 0.95\}$, with $N = n = 10^6$

Dependence structure: Extreme value Markov process

Define (arbitrarily) x_n such that $F^n(x_n) = 1/2$ in Eq. 1.

Then, using Eq. 1, we can define

$$\theta_n = -\frac{\log \Pr(\max\{X_1, \dots, X_n\} \leq x_n)}{\log 2}, \quad (5)$$

and so $\theta_n \rightarrow \theta$ as $n \rightarrow \infty$.

We can use Eq. 5 to investigate the relationship between θ and α in the logistic model:

- Simulate N first-order Markov chains, each of length n , with logistic dependence α
- The probability in the numerator of Eq. 5 is estimated as the proportion of simulated chains whose maximum does not exceed x_n
- We find values for θ_n for $\alpha = \{0.05, 0.10, \dots, 0.95\}$, with $N = n = 10^6$

Dependence structure: Extreme value Markov process

Define (arbitrarily) x_n such that $F^n(x_n) = 1/2$ in Eq. 1.

Then, using Eq. 1, we can define

$$\theta_n = -\frac{\log \Pr(\max\{X_1, \dots, X_n\} \leq x_n)}{\log 2}, \quad (5)$$

and so $\theta_n \rightarrow \theta$ as $n \rightarrow \infty$.

We can use Eq. 5 to investigate the relationship between θ and α in the logistic model:

- Simulate N first-order Markov chains, each of length n , with logistic dependence α
- The probability in the numerator of Eq. 5 is estimated as the proportion of simulated chains whose maximum does not exceed x_n
- We find values for θ_n for $\alpha = \{0.05, 0.10, \dots, 0.95\}$, with $N = n = 10^6$

Dependence structure: Extreme value Markov process

Define (arbitrarily) x_n such that $F^n(x_n) = 1/2$ in Eq. 1.

Then, using Eq. 1, we can define

$$\theta_n = -\frac{\log \Pr(\max\{X_1, \dots, X_n\} \leq x_n)}{\log 2}, \quad (5)$$

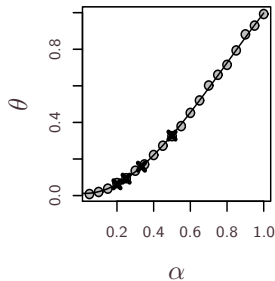
and so $\theta_n \rightarrow \theta$ as $n \rightarrow \infty$.

We can use Eq. 5 to investigate the relationship between θ and α in the logistic model:

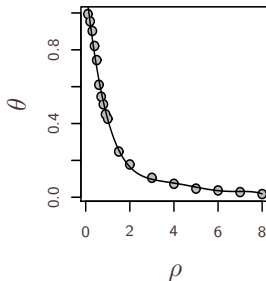
- Simulate N first-order Markov chains, each of length n , with logistic dependence α
- The probability in the numerator of Eq. 5 is estimated as the proportion of simulated chains whose maximum does not exceed x_n
- We find values for θ_n for $\alpha = \{0.05, 0.10, \dots, 0.95\}$, with $N = n = 10^6$

Dependence structure: Extreme value Markov process

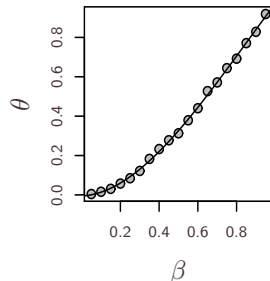
Logistic



Negative logistic



Bilogistic

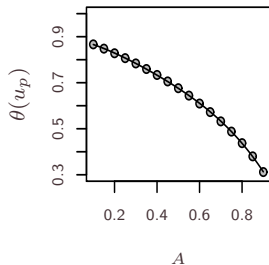


Giving, for example,

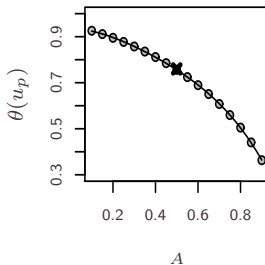
$$\theta \approx 0.013 - 0.092\alpha + 1.833\alpha^2 - 0.756\alpha^3$$

Dependence structure: AR(1) process

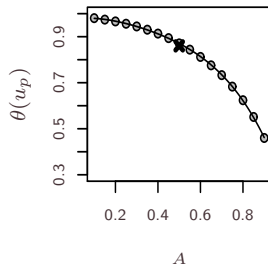
AR(1) with $p = 0.9$



AR(1) with $p = 0.95$



AR(1) with $p = 0.99$



Dependence structure: Extreme value Markov process

- Inference proceeds for $\psi = (\eta_m, \xi_m, \alpha_m)^T$ using the EVMP likelihood, leading to estimates of $\psi' = (\eta_m, \xi_m, \theta_m)^T$, $m = 1, \dots, 12$.
- Prior specification:
 - ▶ We use a re-parameterised GP scale: $\eta_m = \log(\sigma_m - \xi_m u_m)$
 - ▶ We assume independence between marginal and dependence parameters, specifying

$$(\eta_m, \xi_m)^T \sim N_{24}(\mu, \Sigma) \quad \text{and} \quad \alpha_m \sim \text{Beta}(10, 19).$$

- ▶ Components of μ chosen based on a previous study of wind speeds at a nearby location (Keele)
- ▶ We specify values for $\text{cov}(\eta_m, \xi_m)$ according to our observations at Keele (crudely scaled to acknowledge our between-site uncertainty)
- ▶ Priors for α_m chosen for similar reasons

Dependence structure: Extreme value Markov process

- Inference proceeds for $\psi = (\eta_m, \xi_m, \alpha_m)^T$ using the EVMP likelihood, leading to estimates of $\psi' = (\eta_m, \xi_m, \theta_m)^T$, $m = 1, \dots, 12$.
- Prior specification:
 - ▶ We use a re-parameterised GP scale: $\eta_m = \log(\sigma_m - \xi_m u_m)$
 - ▶ We assume independence between marginal and dependence parameters, specifying

$$(\eta_m, \xi_m)^T \sim N_{24}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{and} \quad \alpha_m \sim \text{Beta}(10, 19).$$

- ▶ Components of $\boldsymbol{\mu}$ chosen based on a previous study of wind speeds at a nearby location (Keele)
- ▶ We specify values for $\text{cov}(\eta_m, \xi_m)$ according to our observations at Keele (crudely scaled to acknowledge our between-site uncertainty)
- ▶ Priors for α_m chosen for similar reasons

Bayesian sampling and return level inference

- Simple Metropolis-Hastings sampling used to obtain successive draws from

$$\psi'^{[j]} = (\eta_m^{[j]}, \xi_m^{[j]}, \theta_m^{[j]})^T, \quad m = 1, \dots, 12; j = 1, \dots, 10^6$$

- Within each Metropolis step we use random walk updates to generate candidate values for the parameters, tuning the innovation variances to optimise efficiency
- Convergence assessed by starting each chain at multiple initial values and observing trace plots; thinning applied to reduce autocorrelation
- $$\begin{array}{ccc} (\eta_m^{[1]}, \xi_m^{[1]}, \theta_m^{[1]}) & \xrightarrow{\text{Solve Eq. 2}} & z_r^{[1]} \\ (\eta_m^{[2]}, \xi_m^{[2]}, \theta_m^{[2]}) & \xrightarrow{\text{Solve Eq. 2}} & z_r^{[2]} \\ \vdots & & \vdots \end{array}$$
- This leads to standard reporting of, e.g., **posterior mean** plus **posterior st. dev. / credible interval**

Bayesian sampling and return level inference

		z_{10}	z_{50}	z_{1000}	z_{10000}
Yeovilton (knots)	Mean	61.7	66.1	82.2	95.8
	Median	60.9	65.8	74.9	83.4
	95% CI	(56.0, 69.7)	(58.8, 77.3)	(67.8, 107.0)	(74.1, 169.8)
Bradfield (knots)	Mean	96.9	105.5	132.1	166.7
	Median	95.0	103.5	128.4	155.2
	95% CI	(91.7, 102.8)	(96.6, 115.9)	(104.6, 150.4)	(112.3, 193.1)
Rainfall SWE (mm)	Mean	92.3	—	150.3	231.8
	Median	—	—	—	—
	95% CI	(83.3, 103.2)	—	(127.8, 178.2)	(178.1, 340.2)

Predictive inference

- Bayesian framework offers natural extension to prediction through the **posterior predictive distribution**
- Let Y denote a future extreme of our wind speed process. Then we can write

$$\Pr(Y \leq y|x) = \int_{\Psi} \Pr(Y \leq y|\psi) \pi(\psi|x) d\psi \quad (6)$$

- Equating to $1 - r^{-1}$ and solving for $y = z_{r,\text{pred}}$ gives a return level estimate which incorporates uncertainty due to parameter estimation and randomness in future observations
- Eq. 6 analytically intractable; approximate using

$$\frac{1}{S} \sum_{j=1}^S \Pr \left\{ Y \leq z_{r,\text{pred}} | \psi^{[j]} \right\} = \frac{1}{S} \sum_{j=1}^S \prod_{m=1}^{12} \mathcal{H}_m^{\theta_m}(z_{r,\text{pred}} | \eta_m^{[j]}, \xi_m^{[j]}, \alpha_m^{[j]}),$$

where S is the number of MCMC iterations after burn-in

Predictive inference

- Bayesian framework offers natural extension to prediction through the **posterior predictive distribution**
- Let Y denote a future extreme of our wind speed process. Then we can write

$$\Pr(Y \leq y|x) = \int_{\Psi} \Pr(Y \leq y|\psi) \pi(\psi|x) d\psi \quad (6)$$

- Equating to $1 - r^{-1}$ and solving for $y = z_{r,\text{pred}}$ gives a return level estimate which incorporates uncertainty due to parameter estimation and randomness in future observations
- Eq. 6 analytically intractable; approximate using

$$\frac{1}{S} \sum_{j=1}^S \Pr\{Y \leq z_{r,\text{pred}}|\psi^{[j]}\} = \frac{1}{S} \sum_{j=1}^S \prod_{m=1}^{12} \mathcal{H}_m^{\theta_m}(z_{r,\text{pred}}|\eta_m^{[j]}, \xi_m^{[j]}, \alpha_m^{[j]}),$$

where S is the number of MCMC iterations after burn-in

Predictive inference

- Bayesian framework offers natural extension to prediction through the **posterior predictive distribution**
- Let Y denote a future extreme of our wind speed process. Then we can write

$$\Pr(Y \leq y | x) = \int_{\Psi} \Pr(Y \leq y | \psi) \pi(\psi | x) d\psi \quad (6)$$

- Equating to $1 - r^{-1}$ and solving for $y = z_{r,\text{pred}}$ gives a return level estimate which incorporates uncertainty due to parameter estimation and randomness in future observations
- Eq. 6 analytically intractable; approximate using

$$\frac{1}{S} \sum_{j=1}^S \Pr \left\{ Y \leq z_{r,\text{pred}} | \psi^{[j]} \right\} = \frac{1}{S} \sum_{j=1}^S \prod_{m=1}^{12} \mathcal{H}_m^{\theta_m} (z_{r,\text{pred}} | \eta_m^{[j]}, \xi_m^{[j]}, \alpha_m^{[j]}),$$

where S is the number of MCMC iterations after burn-in

Predictive inference

- Bayesian framework offers natural extension to prediction through the **posterior predictive distribution**
- Let Y denote a future extreme of our wind speed process. Then we can write

$$\Pr(Y \leq y | x) = \int_{\Psi} \Pr(Y \leq y | \psi) \pi(\psi | x) d\psi \quad (6)$$

- Equating to $1 - r^{-1}$ and solving for $y = z_{r,\text{pred}}$ gives a return level estimate which incorporates uncertainty due to parameter estimation and randomness in future observations
- Eq. 6 analytically intractable; approximate using

$$\frac{1}{S} \sum_{j=1}^S \Pr \left\{ Y \leq z_{r,\text{pred}} | \psi^{[j]} \right\} = \frac{1}{S} \sum_{j=1}^S \prod_{m=1}^{12} \mathcal{H}_m^{\theta_m}(z_{r,\text{pred}} | \eta_m^{[j]}, \xi_m^{[j]}, \alpha_m^{[j]}),$$

where S is the number of MCMC iterations after burn-in

Predictive inference

		z_{10}	z_{50}	z_{1000}	z_{10000}
Yeovilton (knots)	Mean	61.7	66.1	82.2	95.8
	Median	60.9	65.8	74.9	83.4
	95% CI	(56.0, 69.7)	(58.8, 77.3)	(67.8, 107.0)	(74.1, 169.8)
	Pred.	62.2	67.1	87.3	109.4
Bradfield (knots)	Mean	96.9	105.5	132.1	166.7
	Median	95.0	103.5	128.4	155.2
	95% CI	(91.7, 102.8)	(96.6, 115.9)	(104.6, 150.4)	(112.3, 193.1)
	Pred.	101.4	113.1	128.1	195.0
Rainfall SWE (mm)	Mean	92.3	—	150.3	231.8
	Median	—	—	—	—
	95% CI	(83.3, 103.2)	—	(127.8, 178.2)	(178.1, 340.2)
	Pred.	94.1	—	163.4	302.4

Simulation study: Aims and design

- Investigate performance of $z_{r,\text{pred}}$ relative to other point summaries from the return level posterior distribution: specifically, the
 - ▶ posterior mean $\bar{z}_r$
 - ▶ posterior mode $\dot{z}_r$
 - ▶ 95% credible upper bound $z_{r,\text{upp}}$
- **Idea:** Investigate the **return probabilities** associated with $z_{r,\text{pred}}$ relative to the other posterior summaries, and the intended value r^{-1}

Simulation study: Aims and design

1. Simulate a (stationary) reference series $\mathbf{y}^{(\text{Ref})}$ of length N years, with N_y observations per year (N very large)
 - ▶ First-order Markov chain
 - ▶ Asymptotically dependent: **EVMP** – logistic(α) and bilogistic(α, β)
 - ▶ Asymptotically independent: **AR(1)** process with lag 1 autocorrelation A
 - ▶ Marginally, $\mathbf{y}^{(\text{Ref})} \sim \text{GP}(\sigma, \xi)$, giving $\text{GP}(\sigma^* = \sigma + \xi u, \xi^* = \xi)$ for excesses over some threshold u

Simulation study: Aims and design

1. Simulate a (stationary) reference series $\mathbf{y}^{(\text{Ref})}$ of length N years, with N_y observations per year (N very large)
 - ▶ First-order Markov chain
 - ▶ Asymptotically dependent: **EVMP** – logistic(α) and bilogistic (α, β)
 - ▶ Asymptotically independent: **AR(1)** process with lag 1 autocorrelation A
 - ▶ Marginally, $\mathbf{y}^{(\text{Ref})} \sim \text{GP}(\sigma, \xi)$, giving $\text{GP}(\sigma^* = \sigma + \xi u, \xi^* = \xi)$ for excesses over some threshold u

Simulation study: Aims and design

2. At each replication ℓ , $\ell = 1, \dots, L$, simulate a stationary series $\mathbf{y}^{(\ell)}$ of length n years, with n_y observations per year
 - ▶ First-order Markov chain: EVMP or AR(1) process
 - ▶ Perform a full MCMC procedure on $\mathbf{y}^{(\ell)}$ using objective (and, where available, conjugate) priors
 - ▶ Obtain (approximate) samples from the posteriors for σ^* , ξ^* and the associated dependence parameters (e.g. α or $A \rightarrow \theta$)
 - ▶ Hence obtain $\bar{z}_r^{(\ell)}$, $\dot{z}_r^{(\ell)}$, $z_{r,\text{upp}}^{(\ell)}$ and $z_{r,\text{pred}}^{(\ell)}$

Simulation study: Aims and design

3. Define p : Proportion of annual maxima in $\mathbf{y}^{(\text{Ref})}$ exceeding each of $\bar{z}_r^{(\ell)}$, $\dot{z}_r^{(\ell)}$, $z_{r,\text{upp}}^{(\ell)}$ and $z_{r,\text{pred}}^{(\ell)}$, $\ell = 1, \dots, L$

- ▶ Gives sampling distributions for $p_{\bar{z}_r}$, $p_{\dot{z}_r}$, $p_{z_{r,\text{upp}}}$ and $p_{z_{r,\text{pred}}}$
- ▶ Can compare p_* to each other, and to r^{-1}

4. We use

- ▶ $N = 10^6$ and $N_y = 365.25 \times 24$
- ▶ *Asymptotic dependence*: $\alpha = \{0.3, \dots, 0.9\}$ and $\alpha = 0.5$; $\beta = \{0.3, \dots, 0.9\}$ for the logistic and bilogistic models respectively
- ▶ *Asymptotic independence*: $A = \{0.2, \dots, 0.8\}$
- ▶ *Marginals*: $\sigma = 1$ and $\xi = \{-0.4, -0.1, 0, 0.1, 0.3\}$, yielding $\sigma^* = \xi u + 1$ and $\xi^* = \xi$, where we use $u = \{u_{0.9}, u_{0.95}, u_{0.99}\}$
- ▶ $L = 10^4$; $n = 50$ and $n_y = N_y$
- ▶ *MCMC*: $S = 10^4$, pilot runs for suitable MCMC tuning at each replication ℓ – acceptance probabilities in the range 20%–40%

Simulation study: Aims and design

3. Define p : Proportion of annual maxima in $\mathbf{y}^{(\text{Ref})}$ exceeding each of $\bar{z}_r^{(\ell)}$, $\dot{z}_r^{(\ell)}$, $z_{r,\text{upp}}^{(\ell)}$ and $z_{r,\text{pred}}^{(\ell)}$, $\ell = 1, \dots, L$

- ▶ Gives sampling distributions for $p_{\bar{z}_r}$, $p_{\dot{z}_r}$, $p_{z_{r,\text{upp}}}$ and $p_{z_{r,\text{pred}}}$
- ▶ Can compare p_* to each other, and to r^{-1}

4. We use

- ▶ $N = 10^6$ and $N_y = 365.25 \times 24$
- ▶ *Asymptotic dependence*: $\alpha = \{0.3, \dots, 0.9\}$ and $\alpha = 0.5$; $\beta = \{0.3, \dots, 0.9\}$ for the logistic and bilogistic models respectively
- ▶ *Asymptotic independence*: $A = \{0.2, \dots, 0.8\}$
- ▶ *Marginals*: $\sigma = 1$ and $\xi = \{-0.4, -0.1, 0, 0.1, 0.3\}$, yielding $\sigma^* = \xi u + 1$ and $\xi^* = \xi$, where we use $u = \{u_{0.9}, u_{0.95}, u_{0.99}\}$
- ▶ $L = 10^4$; $n = 50$ and $n_y = N_y$
- ▶ *MCMC*: $S = 10^4$, pilot runs for suitable MCMC tuning at each replication ℓ – acceptance probabilities in the range 20%–40%

Simulation study: Aims and design

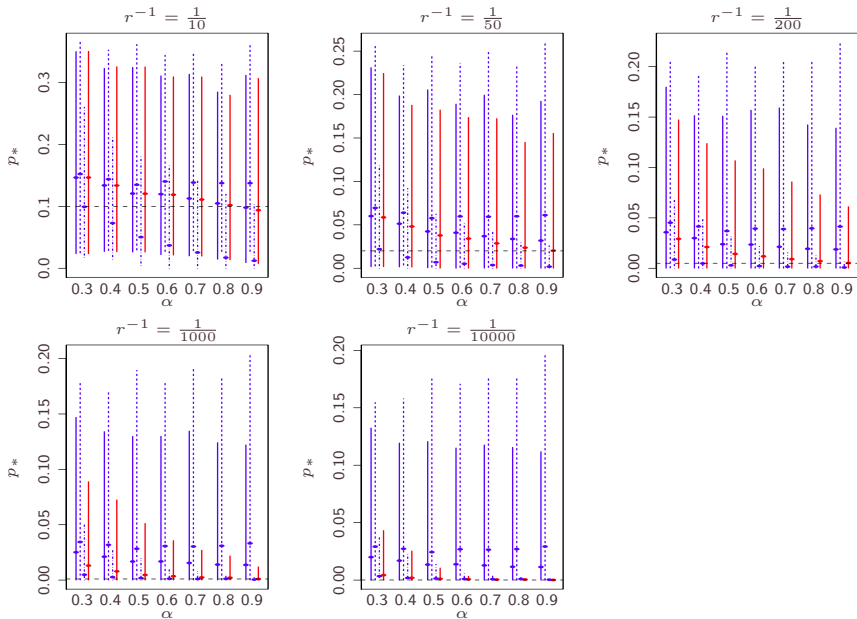
3. Define p : Proportion of annual maxima in $\mathbf{y}^{(\text{Ref})}$ exceeding each of $\bar{z}_r^{(\ell)}$, $\dot{z}_r^{(\ell)}$, $z_{r,\text{upp}}^{(\ell)}$ and $z_{r,\text{pred}}^{(\ell)}$, $\ell = 1, \dots, L$

- Gives sampling distributions for $p_{\bar{z}_r}$, $p_{\dot{z}_r}$, $p_{z_{r,\text{upp}}}$ and $p_{z_{r,\text{pred}}}$
- Can compare p_* to each other, and to r^{-1}

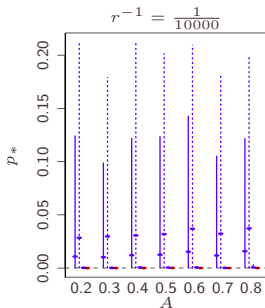
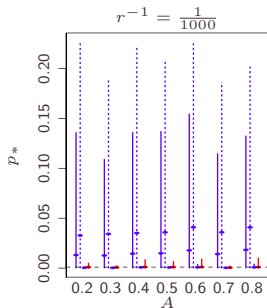
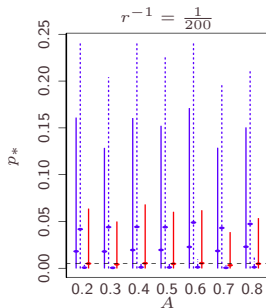
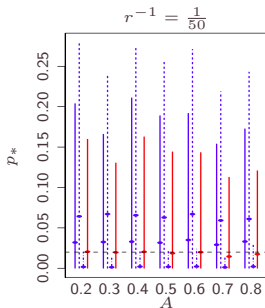
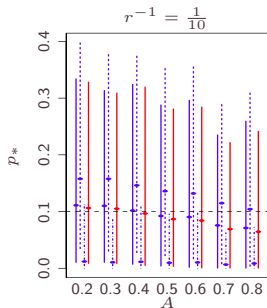
4. We use

- $N = 10^6$ and $N_y = 365.25 \times 24$
- *Asymptotic dependence*: $\alpha = \{0.3, \dots, 0.9\}$ and $\alpha = 0.5$; $\beta = \{0.3, \dots, 0.9\}$ for the logistic and bilogistic models respectively
- *Asymptotic independence*: $A = \{0.2, \dots, 0.8\}$
- *Marginals*: $\sigma = 1$ and $\xi = \{-0.4, -0.1, 0, 0.1, 0.3\}$, yielding $\sigma^* = \xi u + 1$ and $\xi^* = \xi$, where we use $u = \{u_{0.9}, u_{0.95}, u_{0.99}\}$
- $L = 10^4$; $n = 50$ and $n_y = N_y$
- *MCMC*: $S = 10^4$, pilot runs for suitable MCMC tuning at each replication ℓ – acceptance probabilities in the range 20%–40%

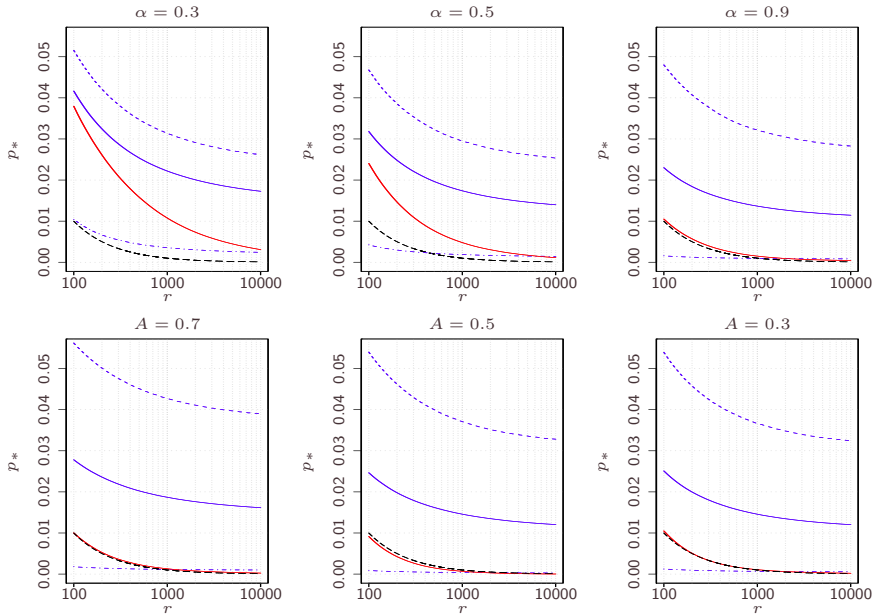
Simulation study: Some results (EMVP, $\xi = -0.1$)



Simulation study: Some results (AR(1), $\xi = -0.1$)



Simulation study: Some results

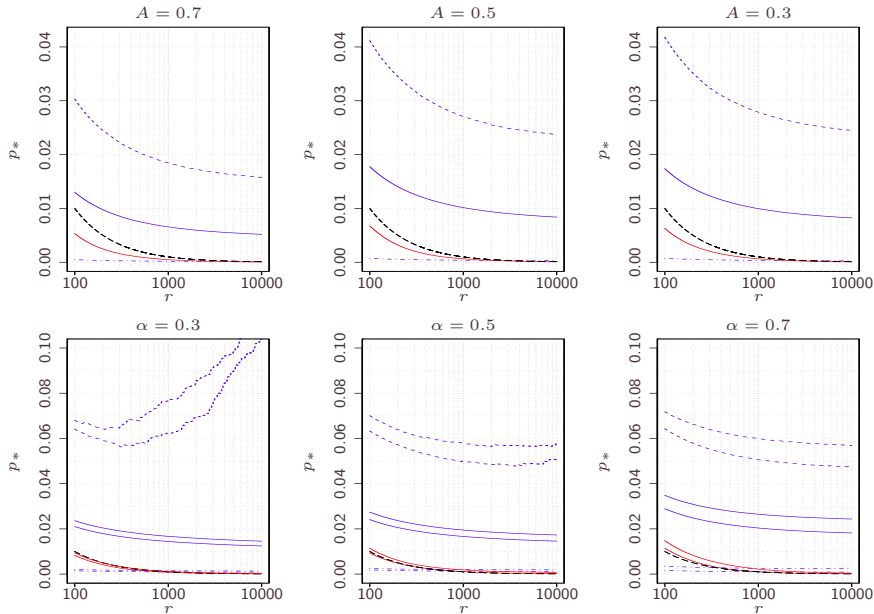


Simulation study: Some results

Regardless of the shape parameter, dependence structure and threshold used, results are largely consistent with what we have seen here

- Predictive return levels give estimates of return probabilities more in line with what we'd expect to see (r^{-1}), and with greater precision
- Most noticeable for series exhibiting asymptotic independence, and for longer return periods
- Larger (positive) biases in estimates of return probabilities given by $\bar{z}_r$ and $\dot{z}_r$ than those given by $z_{r,\text{pred}}$ $\longrightarrow$
underprotection?

Simulation study: Some results



Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!

Recommendations

- Use the predictive return level!
 - ▶ Gives a **convenient**, single number summary for our return level, accounting for uncertainty in *both* parameter estimation *and* future observations
 - ▶ Could be **superior** to point summaries from the return level posterior distribution
- At the same time, try to get the most from your data / other sources of information
 - ▶ Use **threshold excesses**
 - ▶ Use **all** threshold excesses
 - ▶ Use extremes from **all seasons**, if that seems sensible
 - ▶ Use **informative priors**
- Use our **Shiny app**!