Excesses over threshold method for wind speed

Eulalia Nualart

University of Paris 13 and Universidad Pública de Navarra

with Fermín Mallor (Universidad Pública de Navarra), Edouard Omey (H. University of Brussels) and CENER (Centro de Energías Renovables)

Basque Center for Applied Mathematics
March 13th 2009
The Problem

- **Goal**: estimate the reference wind speed \( V_{\text{ref}} \), that is, the extreme 10-min average wind speed that will occur in a given location with a recurrence period of 50 years.
The Problem

- **Goal**: estimate the reference with speed (\(V_{\text{ref}}\)), that is, the extreme 10-min average wind speed that will occur in a given location with a recurrence period of 50 years.

- This is a basic parameter for wind turbine classes and therefore strongly related to design of wind turbines.
The Problem

- **Goal**: estimate the reference with speed ($V_{ref}$), that is, the extreme 10-min average wind speed that will occur in a given location with a recurrence period of 50 years.

- This is a basic parameter for wind turbine classes and therefore strongly related to design of wind turbines.

- In general $V_{ref}$ has to be determined statistically on the basis of either on-site measurement or long-term measurements, e.g. meteorological stations or reanalysis data.
Wind turbine classes and methodology

- Basic parameters for wind turbine generators classes

**Table:** Classification of wind turbine generators (WTGS) according to Vref:

<table>
<thead>
<tr>
<th>WTGS class</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vref (m/s)</td>
<td>50</td>
<td>42.5</td>
<td>37.5</td>
<td>30</td>
</tr>
</tbody>
</table>

Method to estimate Vref:
- Fit the data to a theoretical distribution (extreme value distribution) in order to calculate quantiles.
- Approach designed specifically to deal with short data sets (usually not more than 10-20 years).
Wind turbine classes and methodology

- Basic parameters for wind turbine generators classes

Table: Classification of wind turbine generators (WTGS) according to $V_{ref}$:

<table>
<thead>
<tr>
<th>WTGS class</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{ref}$ (m/s)</td>
<td>50</td>
<td>42.5</td>
<td>37.5</td>
<td>30</td>
</tr>
</tbody>
</table>

- Method to estimate $V_{ref}$: Fit the data to a theoretical distribution (extreme value distribution) in order to calculate quantiles.
Wind turbine classes and methodology

- Basic parameters for wind turbine generators classes

  **Table:** Classification of wind turbine generators (WTGS) according to Vref:

<table>
<thead>
<tr>
<th>WTGS class</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vref (m/s)</td>
<td>50</td>
<td>42.5</td>
<td>37.5</td>
<td>30</td>
</tr>
</tbody>
</table>

- Method to estimate Vref: Fit the data to a theoretical distribution (extreme value distribution) in order to calculate quantiles.

- Approach designed specifically to deal with short data sets (usually not more that 10-20 years).
Extreme value distributions: its definition

- \((X_i, i \geq 1)\) i.i.d with d.f. \(F\).
Extreme value distributions: its definition

- \((X_i, i \geq 1)\) i.i.d with d.f. \(F\).
- Then, \(M_n = \max(X_1, \ldots, X_n), \ n \geq 1\), has d.f. \(F^n\).
Extreme value distributions: its definition

- $(X_i, i \geq 1)$ i.i.d with d.f. $F$.
- Then, $M_n = \max(X_1, \ldots, X_n)$, $n \geq 1$, has d.f. $F^n$.
- Thus, $\forall x : F(x) < 1$, $F^n(x) \to 0$, $n \to \infty$. 

$G$ is an extreme value distribution if $\exists a_n > 0, b_n, n = 1, 2, \ldots$:

$$F_n(a_n x + b_n) \to G(x), \quad n \to \infty$$

$\forall x$ point of continuity of $G$.

If $G$ exists then $F \in D(G)$. 

(\(X_i, i \geq 1\)) i.i.d with d.f. \(F\).

Then, \(M_n = \max(X_1, ..., X_n), n \geq 1\), has d.f. \(F^n\).

Thus, \(\forall x : F(x) < 1, F^n(x) \to 0, n \to \infty\).

\(G\) is an extreme value distribution if \(\exists a_n > 0, b_n, n = 1, 2, .. :\)

\[F^n(a_n x + b_n) \to G(x), n \to \infty\]

\(\forall x\) point of continuity of \(G\).
Extreme value distributions: its definition

- $(X_i, i \geq 1)$ i.i.d with d.f. $F$.
- Then, $M_n = \max(X_1, \ldots, X_n), n \geq 1,$ has d.f. $F^n$.
- Thus, $\forall x : F(x) < 1, \ F^n(x) \to 0, \ n \to \infty$.
- $G$ is an extreme value distribution if $\exists a_n > 0, \ b_n, \ n = 1, 2, \ldots$
  
  $$F^n(a_n x + b_n) \to G(x), \ n \to \infty$$

$\forall x$ point of continuity of $G$.
- If $G$ exists then $F \in D(G)$. 
Characterization of extreme value distributions

Theorem (Gnedenko’43)

If \( F \in D(G) \), then \( G \) is of the type \( G_\gamma(ax + b) \) with \( a > 0 \) and \( b \in \mathbb{R} \), where

\[
G_\gamma(x) = \begin{cases} 
\exp \left( -\left(1 + \gamma x \right)^{-1/\gamma} \right), & \text{if } \gamma \neq 0, \\
\exp(-e^{-x}), & \text{if } \gamma = 0,
\end{cases}
\]

where \( y_+ = \max(y, 0) \).

\( \gamma \) is called the extreme value index, we write \( F \in D(G_\gamma) \).
Behaviour of the tails

- Behaviour of $1 - G_\gamma(x)$ for large $x$?

  - Case $\gamma > 0$: Fréchet distribution. Long-tailed case: the tail of $1 - G_\gamma(x)$ decreases polynomially as $x^{-1/\gamma}$.
  
  - Case $\gamma = 0$: Gumbel distribution. Medium-tailed case: the tail of $1 - G_\gamma(x)$ decreases exponentially.
  
  - Case $\gamma < 0$: Weibull distribution. Short-tailed case: the tail of $1 - G_\gamma(x)$ has a finite endpoint, the minimum value of $x$ for which $G_\gamma(x) = 1$. 
Behaviour of the tails

- Behaviour of $1 - G_\gamma(x)$ for large $x$?

- **Case $\gamma > 0$:** Fréchet distribution.
  Long-tailed case: the tail of $1 - G_\gamma(x)$ decreases polynomially as $x^{-1/\gamma}$.
Behaviour of the tails

- Behaviour of $1 - G_\gamma(x)$ for large $x$?

- **Case $\gamma > 0$:** Fréchet distribution.
  Long-tailed case: the tail of $1 - G_\gamma(x)$ decreases polynomially as $x^{-1/\gamma}$.

- **Case $\gamma = 0$:** Gumbel distribution.
  Medium-tailed case: the tail of $1 - G_\gamma(x)$ decreases exponentially.

- **Case $\gamma < 0$:** Weibull distribution.
  Short-tailed case: the tail of $1 - G_\gamma(x)$ has a finite endpoint, the minimum value of $x$ for which $G_\gamma(x) = 1$. 
Behaviour of the tails

- Behaviour of $1 - G_\gamma(x)$ for large $x$?

- Case $\gamma > 0$: Fréchet distribution.
  Long-tailed case: the tail of $1 - G_\gamma(x)$ decreases polynomially as $x^{-1/\gamma}$.

- Case $\gamma = 0$: Gumbel distribution.
  Medium-tailed case: the tail of $1 - G_\gamma(x)$ decreases exponentially.

- Case $\gamma < 0$: Weibull distribution.
  Short-tailed case: the tail of $1 - G_\gamma(x)$ has a finite endpoint, the minimum value of $x$ for which $G_\gamma(x) = 1$. 
Consider the distribution of $Y = X - u$ conditional on exceeding a threshold $u$:

$$F_\mathcal{U}(x) = P\{Y \leq x | Y > 0\} = \frac{F(u + x) - F(u)}{1 - F(u)}.$$
Generalized Pareto distribution

Consider the distribution of $Y = X - u$ conditional on exceeding a threshold $u$:

$$F_u(x) = P\{Y \leq x | Y > 0\} = \frac{F(u + x) - F(u)}{1 - F(u)}.$$ 

Theorem (Pickands’75): $F \in D(G_\gamma) \iff \exists \sigma > 0$:

$$F_u(x) \rightarrow H_{\gamma,\sigma}(x) := \begin{cases} 
1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-1/\gamma}, & \text{if } \gamma \neq 0, \\
1 - \exp(-x/\sigma), & \text{if } \gamma = 0.
\end{cases}$$

as $u \rightarrow \sup\{x : F(x) < 1\}$. 

$H_{\gamma,\sigma}$ is the generalized Pareto distribution function.

$\sigma$ is the scale parameter.

When $\gamma \geq 0$ the support is $[0, +\infty[$. When $\gamma < 0$ is $[0, \sigma|\gamma|[$.

When $\gamma = 0$ the g.P.d. is the Exponential $(1/\sigma)$. 

Generalized Pareto distribution

- Consider the distribution of $Y = X - u$ conditional on exceeding a threshold $u$:

  $$F_u(x) = \mathbb{P}\{Y \leq x \mid Y > 0\} = \frac{F(u + x) - F(u)}{1 - F(u)}.$$ 

- **Theorem (Pickands’75):** $F \in D(G_\gamma) \iff \exists \sigma > 0$:

  $$F_u(x) \to H_{\gamma,\sigma}(x) := \begin{cases} 
  1 - \left(1 + \frac{\gamma}{\sigma} x\right)^{-1/\gamma}, & \text{if } \gamma \neq 0, \\
  1 - \exp(-x/\sigma), & x > 0, \quad \text{if } \gamma = 0.
  \end{cases}$$

  as $u \to \sup\{x : F(x) < 1\}$.

- $H_{\gamma,\sigma}$ is the generalized Pareto distribution function.
Generalized Pareto distribution

Consider the distribution of \( Y = X - u \) conditional on exceeding a threshold \( u \):

\[
F_u(x) = P\{ Y \leq x | Y > 0 \} = \frac{F(u + x) - F(u)}{1 - F(u)}.
\]

Theorem (Pickands’75): \( F \in D(G_\gamma) \iff \exists \sigma > 0 : \)

\[
F_u(x) \to H_{\gamma, \sigma}(x) := \begin{cases} 
1 - \left(1 + \frac{\gamma}{\sigma} x\right)^{-1/\gamma}, & \text{if } \gamma \neq 0, \\
1 - \exp(-x/\sigma), & x > 0, \text{ if } \gamma = 0.
\end{cases}
\]

as \( u \to \sup\{ x : F(x) < 1 \} \).

\( H_{\gamma, \sigma} \) is the generalized Pareto distribution function.

\( \sigma \) is the scale parameter.
Generalized Pareto distribution

Consider the distribution of $Y = X - u$ conditional on exceeding a threshold $u$:

$$F_u(x) = \Pr\{Y \leq x \mid Y > 0\} = \frac{F(u + x) - F(u)}{1 - F(u)}.$$

Theorem (Pickands’75): $F \in D(G_\gamma) \iff \exists \sigma > 0:\$

$$F_u(x) \to H_{\gamma, \sigma}(x) := \begin{cases} 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-1/\gamma}, & \text{if } \gamma \neq 0, \\ 1 - \exp(-x/\sigma), & x > 0, \text{ if } \gamma = 0. \end{cases}$$

as $u \to \sup\{x : F(x) < 1\}$.

$H_{\gamma, \sigma}$ is the generalized Pareto distribution function.

$\sigma$ is the scale parameter.

When $\gamma \geq 0$ the support is $[0, +\infty[. \text{ When } \gamma < 0 \text{ is }]0, \frac{\sigma}{|\gamma|}].$
Consider the distribution of $Y = X - u$ conditional on exceeding a threshold $u$:

$$F_u(x) = P\{ Y \leq x | Y > 0 \} = \frac{F(u + x) - F(u)}{1 - F(u)}.$$

**Theorem (Pickands’75):** $F \in D(G_\gamma) \iff \exists \sigma > 0$:

$$F_u(x) \to H_{\gamma,\sigma}(x) := \begin{cases} 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-1/\gamma}, & \text{if } \gamma \neq 0, \\ 1 - \exp(-x/\sigma), & \text{if } \gamma = 0. \end{cases}$$

as $u \to \sup\{x : F(x) < 1\}$.

$H_{\gamma,\sigma}$ is the generalized Pareto distribution function.

$\sigma$ is the scale parameter.

When $\gamma \geq 0$ the support is $]0, +\infty[$. When $\gamma < 0$ is $]0, \frac{\sigma}{|\gamma|}[$.

When $\gamma = 0$ the g.P.d. is the Exponential($\frac{1}{\sigma}$).
Properties of the g.P.d.

- If $X$ has a g.P.d. and $u > 0$, then the conditional distribution of $X - u$ given $X > u$ has also a g.P.d.
Properties of the g.P.d.

- If $X$ has a g.P.d. and $u > 0$, then the conditional distribution of $X - u$ given $X > u$ has also a g.P.d.
- Let $N$ denote the number of excesses over a threshold $u$ during $T$ years. Assume that $N$ has a Poisson($\lambda$) and $X_1, ..., X_N$ are i.i.d. with g.P.d. ($\lambda$=crossing rate per year). Then $\max(X_1, ..., X_N)$ has the extreme value d.f.

$$Z_{\lambda, \gamma, \sigma}(x) = \exp \left( - \left( 1 + \gamma \frac{x - u}{\sigma} \right)^{-1/\gamma} \right).$$

In particular, this reduces to $G_\gamma$ with

$$\sigma = a + \lambda(u - b), \quad \lambda = \left( 1 + \gamma \frac{u - b}{a} \right)^{-1/\gamma}.$$
Properties of the g.P.d.

- If $X$ has a g.P.d. and $u > 0$, then the conditional distribution of $X - u$ given $X > u$ has also a g.P.d.
- Let $N$ denote the number of excesses over a threshold $u$ during $T$ years. Assume that $N$ has a Poisson($\lambda$) and $X_1, ..., X_N$ are i.i.d. with g.P.d. ($\lambda$=crossing rate per year). Then $\max(X_1, ..., X_N)$ has the extreme value d.f. $Z_{\lambda, \gamma, \sigma}(x) = \exp \left( - \left( 1 + \gamma \frac{x - u}{\sigma} \right)^{1/\gamma} \right)$.

In particular, this reduces to $G_{\gamma}$ with

$$\sigma = a + \lambda(u - b), \quad \lambda = \left( 1 + \gamma \frac{u - b}{a} \right)^{-1/\gamma}.$$ 

- These properties characterize the g.P.d.
Independence

Use separation of 48 hours for European wind climates (Davison and Smith’90) to choose the peaks excesses ($y_1 = x_1 - u$, ..., $y_k = x_k - u$).

Wind speed data portion of year 2004
$u=18.3$

$> 48$ hours $> 48$ hours
Choice of the threshold

- **Mean excess plot method**: (Davison and Smith’90)

\[
\begin{align*}
Y = X - u & \text{ has a g.P.d. with } \gamma < 1, \text{ then,} \\
\forall x > 0: \sigma + \gamma x > 0, \\
E[Y - x | Y > x] &= \sigma + \gamma x - 1 - \gamma.
\end{align*}
\]

- Draw the conditional mean excess plot: \( x = \text{the threshold}, y = \text{sample mean of all peak excesses over that threshold.} \)
- Select the lowest threshold where a departure of linearity starts.
- **Difficulties**: The mean excess plot may present high variability, particularly at high thresholds.
- A validation is needed looking at the effect of the choice on the estimates of the Vref.
- In practice it is recommended the threshold be chosen to fix a value between one and five peaks per year.
Choice of the threshold

- **Mean excess plot method**: (Davison and Smith’90)
- **Fact**: if $Y = X - u$ has a g.P.d. with $\gamma < 1$, then, $\forall x > 0$:
  \[ \sigma + \gamma x > 0, \]
  \[ E[Y - x|Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}. \]
Choice of the threshold

- Mean excess plot method: (Davison and Smith’90)

- Fact: if \( Y = X - u \) has a g.P.d. with \( \gamma < 1 \), then, \( \forall x > 0 : \sigma + \gamma x > 0 \),

  \[ E[Y - x | Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}. \]

- Draw the conditional mean excess plot: \( x \) = the threshold, \( y \) = sample mean of all peak excesses over that threshold.
Choice of the threshold

- **Mean excess plot method:** (Davison and Smith’90)
- **Fact:** if \( Y = X - u \) has a g.P.d. with \( \gamma < 1 \), then, \( \forall x > 0 \):
  \[ \sigma + \gamma x > 0, \]
  \[ E[Y - x | Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}. \]
- Draw the **conditional mean excess plot**: \( x \) = the threshold, \( y \) = sample mean of all peak excesses over that threshold.
- Select the **lowest** threshold where a departure of linearity starts.
Choice of the threshold

- **Mean excess plot method**: (Davison and Smith’90)
- **Fact**: if \( Y = X - u \) has a g.P.d. with \( \gamma < 1 \), then, \( \forall x > 0 : \sigma + \gamma x > 0 \),
  \[
  E[Y - x|Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}.
  \]
- **Draw the conditional mean excess plot**: \( x = \)the threshold, \( y = \)sample mean of all peak excesses over that threshold.
- **Select the lowest** threshold where a departure of linearity starts.
- **Difficulties**: The mean excess plot may present high variability, particularly at high thresholds.
Choice of the threshold

- Mean excess plot method: (Davison and Smith’90)
- Fact: if $Y = X - u$ has a g.P.d. with $\gamma < 1$, then, $\forall x > 0$:
  $\sigma + \gamma x > 0,$
  $$E[Y - x|Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}.$$ 
- Draw the conditional mean excess plot: $x =$the threshold, $y =$sample mean of all peak excesses over that threshold.
- Select the lowest threshold where a departure of linearity starts.
- Difficulties: The mean excess plot may present high variability, particularly at high thresholds.
- A validation is needed looking at the effect of the choice on the estimates of the Vref.
Choice of the threshold

▶ Mean excess plot method: (Davison and Smith’90)
▶ Fact: if $Y = X - u$ has a g.P.d. with $\gamma < 1$, then, $\forall x > 0:$
$\sigma + \gamma x > 0,$

$$E[Y - x|Y > x] = \frac{\sigma + \gamma x}{1 - \gamma}.$$ 

▶ Draw the conditional mean excess plot: $x =$the threshold, $y =$sample mean of all peak excesses over that threshold.
▶ Select the lowest threshold where a departure of linearity starts.
▶ Difficulties: The mean excess plot may present high variability, partiuculary at high thresholds.
▶ A validation is needed looking at the effect of the choice on the estimates of the Vref.
▶ In practice it is recommended the threshold be chosen to fix a value between one and five peaks per year.
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a graphical estimate of $\sigma$ and $\gamma$.
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a **graphical estimate** of $\sigma$ and $\gamma$.
- **De Haan moment estimation method (MEM):** gives $\hat{\sigma}$ and $\hat{\gamma}$, where $\hat{\gamma}$ is an asymptotically normal estimator.
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a graphical estimate of $\sigma$ and $\gamma$.
- De Haan moment estimation method (MEM): gives $\hat{\sigma}$ and $\hat{\gamma}$, where $\hat{\gamma}$ is an asymptotically normal estimator.
- The maximum likelihood estimation (MLE): gives $\hat{\sigma}$ and $\hat{\gamma}$. Asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold for $\hat{\gamma}$ if $\gamma > -\frac{1}{2}$.
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a **graphical estimate** of $\sigma$ and $\gamma$.

- **De Haan moment estimation method** (MEM): gives $\hat{\sigma}$ and $\hat{\gamma}$, where $\hat{\gamma}$ is an asymptotically normal estimator.

- **The maximum likelihood estimation** (MLE): gives $\hat{\sigma}$ and $\hat{\gamma}$. Asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold for $\hat{\gamma}$ if $\gamma > -\frac{1}{2}$.

- **Hill’s estimator** for $\gamma$ ($\gamma > 0$) (Hill’75).
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a **graphical estimate** of $\sigma$ and $\gamma$.
- **De Haan moment estimation method** (MEM): gives $\hat{\sigma}$ and $\hat{\gamma}$, where $\hat{\gamma}$ is an asymptotically normal estimator.
- **The maximum likelihood estimation** (MLE): gives $\hat{\sigma}$ and $\hat{\gamma}$. Asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold for $\hat{\gamma}$ if $\gamma > -\frac{1}{2}$.
- **Hill’s estimator** for $\gamma$ ($\gamma > 0$) (Hill’75).
- **Negative weighted moments** method for $\gamma$ ($\gamma < 1$) (Hosking and Wallis’87).
Methods of estimation for the parameters of the g.P.d.

- The conditional mean excess plot suggests a **graphical estimate** of \( \sigma \) and \( \gamma \).
- **De Haan moment estimation method** (MEM): gives \( \hat{\sigma} \) and \( \hat{\gamma} \), where \( \hat{\gamma} \) is an asymptotically normal estimator.
- **The maximum likelihood estimation** (MLE): gives \( \hat{\sigma} \) and \( \hat{\gamma} \). Asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold for \( \hat{\gamma} \) if \( \gamma > -\frac{1}{2} \).
- **Hill’s estimator** for \( \gamma (\gamma > 0) \) (Hill’75).
- **Negative weighted moments** method for \( \gamma (\gamma < 1) \) (Hosking and Wallis’87).
- The simplest and oldest estimator for \( \gamma \) is the **Pickands’ estimator** (Pickands’75).
De Haan moment estimation method

- By means of moments of the excesses obtained from the log-transformed data:

\[ \hat{\gamma} = M_k^{(1)} + 1 - \frac{1}{2} \left( 1 - \frac{(M_k^{(1)})^2}{M_k^{(2)}} \right)^{-1}, \]

where

\[ M_k^{(r)} = \frac{1}{k} \sum_{i=1}^{k} (\log(x_k) - u)^r, \quad r = 1, 2. \]
De Haan moment estimation method

- By means of moments of the excesses obtained from the log-transformed data:

\[ \hat{\gamma} = M_k^{(1)} + 1 - \frac{1}{2} \left( 1 - \frac{(M_k^{(1)})^2}{M_k^{(2)}} \right)^{-1}, \]

where

\[ M_k^{(r)} = \frac{1}{k} \sum_{i=1}^{k} (\log(x_k) - u)^r, \quad r = 1, 2. \]

- \[ \hat{\sigma} = u \frac{M_k^{(1)}}{\rho}, \]

where \[ \rho = 1 \text{ if } \hat{\gamma} \geq 0 \text{ and } \rho = \frac{1}{1 - \hat{\gamma}} \text{ if } \hat{\gamma} < 0. \]
Estimation of the Vref

- $V_{ref}$ = extreme 10-min average wind speed with a recurrence period of 50 years.

Unbiased estimator:

$\hat{\lambda} = \frac{K}{T}$, $k =$ number of peak excesses collected over $T$ years.

The $V_{ref}$ (50-year return level) can be estimated as the level which is exceeded on average once in 50 years.

If $x > u$ the mean crossing rate per year of level $x$ is

$\hat{\lambda} P(Y > x - u) = \frac{1}{50}$, $Y \sim g.P.d.$

$V_{ref}$ is estimated as the $(1 - \frac{1}{50})$-quantile of the $g.P.d.$:

$V_{ref} = H^{-\frac{1}{\hat{\gamma}}}, \hat{\gamma} = \sqrt{\frac{1}{\hat{\sigma}} T \frac{50}{k}} + u$
Estimation of the Vref

- $V_{ref}$ = extreme 10-min average wind speed with a recurrence period of 50 years.
- We assume that the excesses process has a Poisson distribution with rate $\lambda$.

$\hat{\lambda} = \frac{K}{T}, \ k =$ number of peak excesses collected over $T$ years.

The $V_{ref}$ (50-year return level) can be estimated as the level which is exceeded on average once in 50 years.

If $x > u$ the mean crossing rate per year of level $x$ is $\hat{\lambda} P(Y > x - u) = \frac{1}{50}$, $Y \sim \text{P.d.}$

$V_{ref}$ is estimated as the $(1 - \frac{1}{50})$-quantile of the g.P.d.:

$V_{ref} = H - \frac{1}{\hat{\gamma}}\hat{\sigma}(1 - \frac{T}{50} k) + u$
Estimation of the $V_{ref}$

- $V_{ref}$ = extreme 10-min average wind speed with a recurrence period of 50 years.
- We assume that the excesses process has a Poisson distribution with rate $\lambda$.
- **Unbiased estimator**: $\hat{\lambda} = \frac{K}{T}$, $k =$ number of peak excesses collected over $T$ years.
Estimation of the Vref

- **Vref** = extreme 10-min average wind speed with a recurrence period of 50 years.
- We assume that the **excesses process** has a Poisson distribution with rate $\lambda$.
- **Unbiased estimator**: $\hat{\lambda} = \frac{K}{T}$, $k =$ number of peak excesses collected over $T$ years.
- The Vref (**50-year return level**) can be estimated as the level which is exceeded on average once in 50 years.
Estimation of the Vref

- **Vref** = extreme 10-min average wind speed with a recurrence period of 50 years.
- We assume that the excesses process has a Poisson distribution with rate $\lambda$.
- **Unbiased estimator**: $\hat{\lambda} = \frac{K}{T}$, $k =$ number of peak excesses collected over $T$ years.
- The Vref (50-year return level) can be estimated as the level which is exceeded on average once in 50 years.
- If $x > u$, the mean crossing rate per year of level $x$ is

  \[ \hat{\lambda} P(Y > x - u) = \frac{1}{50}, \quad Y \text{ g.P.d.} \]
Estimation of the Vref

- \( \text{Vref} \) = extreme 10-min average wind speed with a recurrence period of 50 years.
- We assume that the excesses process has a Poisson distribution with rate \( \lambda \).
- Unbiased estimator: \( \hat{\lambda} = \frac{K}{T} \), \( K \) = number of peak excesses collected over \( T \) years.
- The \( \text{Vref} \) (50-year return level) can be estimated as the level which is exceeded on average once in 50 years. 
- If \( x > u \) the mean crossing rate per year of level \( x \) is
  \[
  \hat{\lambda} P(Y > x - u) = \frac{1}{50}, \quad Y \text{ g.P.d.}
  \]
- \( \text{Vref} \) is estimated as the \( (1 - \frac{1}{50\hat{\lambda}}) \)-quantile of the g.P.d.:
  \[
  \text{Vref} = H_{\hat{\gamma}, \hat{\sigma}}^{-1} \left( 1 - \frac{T}{50k} \right) + u
  \]
Example 1

Smits’01 applies this method for wind speed from different stations at the Netherlands and chooses 10 as a minimal value of the threshold.

![Graph](image)

**Figure 2**: Example of CME graph.

Because both distributions (CWD and GPD) are conditional distributions, it is necessary to estimate the crossing rate per year $\lambda$ of the threshold in order to calculate exceedance frequencies per year. If the exceedance process above threshold $\omega$ can be assumed to be Poisson distributed (which is the case for sufficient high thresholds), the crossing rate can be estimated by the total number of exceedances of $\omega$ divided by the number of real years (exclusive periods of gaps). With this definition, it seemed that an appropriate value of the lowest threshold value corresponds with a crossing rate of about 7 per year (for each season). The highest threshold value has set to the level that corresponds with a crossing rate of 2 per year, because this frequency is the highest frequency for a aim in the KNMI HYDRA project.

With the help of the distribution parameters and the crossing rate of the threshold value, return levels (or exceedance frequencies per year) can be calculated for each threshold value as follows:

$$\omega \lambda = \omega \lambda$$

where $j$ represents the season (Smits, 2001a), $G$ the exceedance frequency per year, $j \omega$ the exceedance frequency per season based on threshold $\omega$ and $F_j \omega$ the cumulative distribution function per season based on threshold $\omega$. Both the CWD and the GPD can be used to calculate (2.4).

Also return periods per threshold and season can now be calculated:

$$T_j \omega = \omega \lambda$$

where $T$ represents the return period in years.

Final return levels per season are calculated by averaging the return levels obtained with the use of the several threshold values. The preference has been given to the calculation of the mean of the return levels above the calculation of the mean of the distribution parameters, because of the strong correlation between the scale and shape parameter of the CWD. Calculating the mean of both of them independently can result in parameters that are not correctly related.

An example of the averaging is given in Figure 2.3.
Example 1

They apply the g.P.d. with MLE and compare the return levels with a non-asymptotic conditional Weibul distribution, which turns out to be more appropriate for their data.
Example 2

Davison and Smith’90 consider 154 excesses of level 65 $m^3/s$ by the River Nidd at Husingore Wier from 1934 to 1969 years. Around threshold 110 the levels off but the variability observed in the higher threshold shows the existence of a mixture of two populations.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.

- They observe a high variability in all their estimates.

- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.

- Their results are not significant as they obtain very large and variable confidence intervals.

- They conclude that this asymptotic theory may not be applicable to their data and perform a non-asymptotic method based on a Bayesian analysis and compare their results to the g.P.d. method.

- This example shows the difficulties which sometimes arise using the excesses over threshold method to estimate extreme quantiles.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.
- They observe a high variability in all their estimates.
- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.
- Their results are not significant as they obtain very large and variable confidence intervals.
- They conclude that this asymptotic theory may not be applicable to their data and perform a non-asymptotic method based on a Bayesian analysis and compare their results to the g.P.d. method.
- This example shows the difficulties which sometimes arise using the excesses over threshold method to estimate extreme quantiles.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.
- They observe a high variability in all their estimates.
- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.
- Their results are not significant as they obtain very large and variable confidence intervals.
- They conclude that this asymptotic theory may not be applicable to their data and perform a non-asymptotic method based on a Bayesian analysis and compare their results to the g.P.d. method.

This example shows the difficulties which sometimes arise using the excesses over threshold method to estimate extreme quantiles.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.
- They observe a high variability in all their estimates.
- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.
- Their results are not significant as they obtain very large and variable confidence intervals.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.
- They observe a high variability in all their estimates.
- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.
- Their results are not significant as they obtain very large and variable confidence intervals.
- They conclude that this asymptotic theory may not be applicable to their data and perform a non-asymptotic method based on a Bayesian analysis and compare their results to the g.P.d. method.

This example shows the difficulties which sometimes arise using the excesses over threshold method to estimate extreme quantiles.
Example 2

- Davison and Smith’90 apply the MLE and calculate the return levels around 110.
- They observe a high variability in all their estimates.
- They test a goodness-of-fit to the exponential distribution of the excesses and compute confidence intervals for the return levels.
- Their results are not significant as they obtain very large and variable confidence intervals.
- They conclude that this asymptotic theory may not be applicable to their data and perform a non-asymptotic method based on a Bayesian analysis and compare their results to the g.P.d. method.
- This example shows the difficulties which sometimes arise using the excesses over threshold method to estimate extreme quantiles.
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
- Technology center in Pamplona (Navarra, Spain) specialized in applied research and development as well as the promotion of renewable energies.
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
- Technology center in Pamplona (Navarra, Spain) specialized in applied research and development as well as the promotion of renewable energies.
- The file contains data from a wind mast with three levels 37, 25 and 10 m. for 2000-first half 2006 (6.5 years).
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
- Technology center in Pamplona (Navarra, Spain) specialized in applied research and development as well as the promotion of renewable energies.
- The file contains data from a wind mast with three levels 37, 25 and 10 m. for 2000-first half 2006 (6.5 years).
- The data acquisition system records data each 1 second, average each 10 minutes, and records 10 minutes mean wind speed.
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
- Technology center in Pamplona (Navarra, Spain) specialized in applied research and development as well as the promotion of renewable energies.
- The file contains data from a wind mast with three levels 37, 25 and 10 m. for 2000-first half 2006 (6.5 years).
- The data acquisition system records data each 1 second, average each 10 minutes, and records 10 minutes mean wind speed.
- With this information we must obtain the $V_{ref}$ (extreme 10-min average wind speed with recurrence period of 50 years) in order to know what kind of turbine generator we can install in the area of study.
Application to CENER data

- We apply the excesses over threshold method to data provided by the Wind Energy Department of the National Renewable Energy Center (CENER).
- Technology center in Pamplona (Navarra, Spain) specialized in applied research and development as well as the promotion of renewable energies.
- The file contains data from a wind mast with three levels 37, 25 and 10 m. for 2000-first half 2006 (6,5 years).
- The data acquisition system records data each 1 second, average each 10 minutes, and records 10 minutes mean wind speed.
- With this information we must obtain the $V_{ref}$ (extreme 10-min average wind speed with recurrence period of 50 years) in order to know what kind of turbine generator we can install in the area of study.
- We have done the analysis with the 37 meters level.
Figure: Wind speed CENER data for year 2000
Figure: Wind speed CENER data for year 2001
Figure: Wind speed CENER data for year 2002
Figure: Wind speed CENER data for year 2003
Figure: Wind speed CENER data for year 2004
Figure: Wind speed CENER data for year 2005
Figure: Wind speed CENER data for first half of year 2006
Figure: Wind speed CENER data for years 2000-2006
When extreme storms mostly occur in a given season it is better to do the analysis separately for each season and combine to estimate annual return levels (as in Examples 1 and 2).
Application to CENER data

- When extreme storms mostly occur in a given *season* it is better to do the analysis separately for each season and combine to estimate annual return levels (as in Examples 1 and 2).
- We do *not* perceive *seasonality* in our data.
Application to CENER data

- When extreme storms mostly occur in a given season it is better to do the analysis separately for each season and combine to estimate annual return levels (as in Examples 1 and 2).
- We do not perceive seasonality in our data.
- We observe quite regular wind speed data with a maximum wind speed 29 m/s in year 2005.
Application to CENER data

- When extreme storms mostly occur in a given season it is better to do the analysis separately for each season and combine to estimate annual return levels (as in Examples 1 and 2).
- We do not perceive seasonality in our data.
- We observe quite regular wind speed data with a maximum wind speed 29 m/s in year 2005.
- Intuitively this suggest to choose a WTGS of class III.
Application to CENER data

- When extreme storms mostly occur in a given season it is better to do the analysis separately for each season and combine to estimate annual return levels (as in Examples 1 and 2).
- We do not perceive seasonality in our data.
- We observe quite regular wind speed data with a maximum wind speed 29 m/s in year 2005.
- Intuitively this suggest to choose a WTGS of class III.
- We now apply the excesses over threshold method for wind speed.
Conditional mean excess plot

We choose the peak excesses \((y_1 = x_1 - u, \ldots, y_k = x_k - u)\) separated by 48 hours, i.e. 288 10-minutes intervals.
## Number of peaks per year

<table>
<thead>
<tr>
<th>threshold</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>16</td>
<td>16</td>
<td>21</td>
<td>12</td>
<td>17</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>17,5</td>
<td>13</td>
<td>12</td>
<td>15</td>
<td>12</td>
<td>15</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>18,5</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>11</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>19,5</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>20,5</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>21,5</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>22,5</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Conditional mean excess plot

Graphical estimates: $\hat{\gamma} = -0, 21$ and $\hat{\sigma} = 2, 16$.
Optimal choice of threshold: $u = 22, 5$. 
MEM of the g.P.d. parameters and the Vref

<table>
<thead>
<tr>
<th>threshold</th>
<th>$\hat{\lambda} = \frac{k}{6.5}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\sigma}$</th>
<th>Vref</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>8,1538</td>
<td>-0,2985</td>
<td>3,1207</td>
<td>27,7165</td>
</tr>
<tr>
<td>19,5</td>
<td>7,0769</td>
<td>-0,2651</td>
<td>2,8952</td>
<td>28,1174</td>
</tr>
<tr>
<td>20</td>
<td>5,6923</td>
<td>-0,4507</td>
<td>3,4026</td>
<td>26,9587</td>
</tr>
<tr>
<td>21</td>
<td>4,7692</td>
<td>-0,1473</td>
<td>2,1232</td>
<td>28,9777</td>
</tr>
<tr>
<td>22</td>
<td>3,2308</td>
<td>-0,1518</td>
<td>1,9369</td>
<td>28,8621</td>
</tr>
<tr>
<td>22,1</td>
<td>3,2308</td>
<td>-0,0725</td>
<td>1,7042</td>
<td>29,3474</td>
</tr>
<tr>
<td>22,2</td>
<td>3,2308</td>
<td>0,0014</td>
<td>1,4959</td>
<td>29,8329</td>
</tr>
<tr>
<td>22,3</td>
<td>3,0769</td>
<td>0,0177</td>
<td>1,4751</td>
<td>30,0692</td>
</tr>
<tr>
<td>22,4</td>
<td>3,0769</td>
<td>0,0849</td>
<td>1,3815</td>
<td>31,0809</td>
</tr>
<tr>
<td>22,5</td>
<td>2,9231</td>
<td>0,1061</td>
<td>1,3573</td>
<td>31,4174</td>
</tr>
<tr>
<td>22,6</td>
<td>2,6154</td>
<td>0,0760</td>
<td>1,4169</td>
<td>30,9565</td>
</tr>
<tr>
<td>22,7</td>
<td>2,3077</td>
<td>0,0207</td>
<td>1,5050</td>
<td>30,2099</td>
</tr>
<tr>
<td>22,8</td>
<td>2,1538</td>
<td>0,0157</td>
<td>1,5149</td>
<td>30,1557</td>
</tr>
<tr>
<td>22,9</td>
<td>2,0000</td>
<td>0,0039</td>
<td>1,5333</td>
<td>30,0248</td>
</tr>
<tr>
<td>23</td>
<td>1,8462</td>
<td>-0,0175</td>
<td>1,5899</td>
<td>29,9173</td>
</tr>
<tr>
<td>23,1</td>
<td>1,6923</td>
<td>-0,0529</td>
<td>1,6907</td>
<td>29,7873</td>
</tr>
<tr>
<td>23,2</td>
<td>1,5385</td>
<td>-0,1106</td>
<td>1,8511</td>
<td>29,5843</td>
</tr>
<tr>
<td>23,8</td>
<td>1,0769</td>
<td>-0,2248</td>
<td>2,0907</td>
<td>29,3038</td>
</tr>
</tbody>
</table>
Goodness-of-Fit Anderson-Darling test of the exponential distribution

Very good fit, specially for thresholds 22.7, 22.8 and 22.9. This test does not fit for thresholds below 22.2 and above 22.9.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Mean</th>
<th>N</th>
<th>AD</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.2</td>
<td>1.597</td>
<td>21</td>
<td>0.227</td>
<td>0.916</td>
</tr>
<tr>
<td>22.3</td>
<td>1.574</td>
<td>20</td>
<td>0.270</td>
<td>0.858</td>
</tr>
<tr>
<td>22.4</td>
<td>1.474</td>
<td>20</td>
<td>0.255</td>
<td>0.879</td>
</tr>
<tr>
<td>22.5</td>
<td>1.448</td>
<td>19</td>
<td>0.368</td>
<td>0.683</td>
</tr>
<tr>
<td>22.6</td>
<td>1.512</td>
<td>17</td>
<td>0.237</td>
<td>0.852</td>
</tr>
<tr>
<td>22.7</td>
<td>1.606</td>
<td>15</td>
<td>0.148</td>
<td>0.988</td>
</tr>
<tr>
<td>22.8</td>
<td>1.616</td>
<td>14</td>
<td>0.158</td>
<td>0.983</td>
</tr>
<tr>
<td>22.9</td>
<td>1.636</td>
<td>13</td>
<td>0.164</td>
<td>0.980</td>
</tr>
</tbody>
</table>

Probability Plot of 22.2; 22.3; 22.4; 22.5; 22.6; 22.7; 22.8; 22.9

Exponential - 95% CI
## Estimate of the Vref with exponential distribution

<table>
<thead>
<tr>
<th>threshold</th>
<th>$\hat{\sigma}_{MEM}$</th>
<th>Vref</th>
<th>$\hat{\sigma}_{MLE}$</th>
<th>Vref</th>
</tr>
</thead>
<tbody>
<tr>
<td>22,1</td>
<td>1,5890</td>
<td>30,1796</td>
<td>1,6967</td>
<td>30,7273</td>
</tr>
<tr>
<td>22,2</td>
<td>1,4959</td>
<td>29,8065</td>
<td>1,5967</td>
<td>30,3188</td>
</tr>
<tr>
<td>22,3</td>
<td>1,4751</td>
<td>29,7285</td>
<td>1,5740</td>
<td>30,2268</td>
</tr>
<tr>
<td>22,4</td>
<td>1,3815</td>
<td>29,3571</td>
<td>1,4740</td>
<td>29,8232</td>
</tr>
<tr>
<td>22,5</td>
<td>1,3573</td>
<td>29,2656</td>
<td>1,4485</td>
<td>29,7200</td>
</tr>
<tr>
<td>22,6</td>
<td>1,4169</td>
<td>29,5050</td>
<td>1,5122</td>
<td>29,9698</td>
</tr>
<tr>
<td>22,7</td>
<td>1,5050</td>
<td>29,8463</td>
<td>1,6063</td>
<td>30,3271</td>
</tr>
<tr>
<td>22,8</td>
<td>1,5149</td>
<td>29,8884</td>
<td>1,6165</td>
<td>30,3639</td>
</tr>
<tr>
<td>22,9</td>
<td>1,5333</td>
<td>29,9612</td>
<td>1,6358</td>
<td>30,4332</td>
</tr>
<tr>
<td>23</td>
<td>1,5626</td>
<td>30,0710</td>
<td>1,6666</td>
<td>30,5417</td>
</tr>
</tbody>
</table>
Threshold/return level plot using g.P.d. and De Haan MEM

Effect of the choice of the range of thresholds on the estimates of the Vref.
Threshold/return level plot using g.P.d. and De Haan MEM

Effect of the choice of the range of thresholds on the estimates of the Vref.
MLE standard errors and 95% confidence intervals for $\hat{\sigma}_{\text{MLE}}$ with exponential distribution

<table>
<thead>
<tr>
<th>threshold</th>
<th>$\hat{\sigma}_{\text{MLE}}$</th>
<th>standard error</th>
<th>lower C.I.</th>
<th>upper C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>22,2</td>
<td>1.59670</td>
<td>0.348427</td>
<td>1.04106</td>
<td>2.44889</td>
</tr>
<tr>
<td>22,3</td>
<td>1.57404</td>
<td>0.351965</td>
<td>1.01550</td>
<td>2.43977</td>
</tr>
<tr>
<td>22,4</td>
<td>1.47404</td>
<td>0.329604</td>
<td>0.950985</td>
<td>2.28477</td>
</tr>
<tr>
<td>22,5</td>
<td>1.44845</td>
<td>0.332298</td>
<td>0.923901</td>
<td>2.27082</td>
</tr>
<tr>
<td>22,6</td>
<td>1.51225</td>
<td>0.366774</td>
<td>0.940105</td>
<td>2.43260</td>
</tr>
<tr>
<td>22,7</td>
<td>1.60628</td>
<td>0.414741</td>
<td>0.968374</td>
<td>2.66441</td>
</tr>
<tr>
<td>22,8</td>
<td>1.61647</td>
<td>0.432020</td>
<td>0.957358</td>
<td>2.72936</td>
</tr>
<tr>
<td>22,9</td>
<td>1.63582</td>
<td>0.453696</td>
<td>0.949851</td>
<td>2.81720</td>
</tr>
</tbody>
</table>
Standard errors and 95% confidence intervals for the Vref with exponential distribution

<table>
<thead>
<tr>
<th>$u$</th>
<th>$(1 - \frac{1}{50\lambda})$</th>
<th>quantile</th>
<th>s. error</th>
<th>Vref</th>
<th>lower C.I.</th>
<th>upper C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>22,2</td>
<td>0,994</td>
<td>8,1686</td>
<td>1,7825</td>
<td>30,3686</td>
<td>27,526</td>
<td>34,7285</td>
</tr>
<tr>
<td>22,3</td>
<td>0,993</td>
<td>7,8101</td>
<td>1,7464</td>
<td>30,1108</td>
<td>27,3387</td>
<td>34,4058</td>
</tr>
<tr>
<td>22,4</td>
<td>0,993</td>
<td>7,3139</td>
<td>1,6354</td>
<td>29,7139</td>
<td>27,1186</td>
<td>33,7367</td>
</tr>
<tr>
<td>22,5</td>
<td>0,993</td>
<td>7,187</td>
<td>1,6488</td>
<td>29,687</td>
<td>27,0842</td>
<td>33,7675</td>
</tr>
<tr>
<td>22,6</td>
<td>0,992</td>
<td>7,3016</td>
<td>1,7709</td>
<td>29,9016</td>
<td>27,1391</td>
<td>34,3453</td>
</tr>
<tr>
<td>22,7</td>
<td>0,991</td>
<td>7,5664</td>
<td>1,9536</td>
<td>30,2264</td>
<td>27,2615</td>
<td>35,2508</td>
</tr>
<tr>
<td>22,8</td>
<td>0,991</td>
<td>7,6144</td>
<td>2,0350</td>
<td>30,4144</td>
<td>27,309</td>
<td>35,6567</td>
</tr>
<tr>
<td>22,9</td>
<td>0,99</td>
<td>7,5332</td>
<td>2,0893</td>
<td>30,4332</td>
<td>27,2742</td>
<td>35,8737</td>
</tr>
</tbody>
</table>
Conclusions

▶ The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.
Conclusions

- The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.
- From the conditional mean excess we obtain 22.5 as an optimal choice, with number of peaks between 1 and 5.
Conclusions

- The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.

- From the conditional mean excess we obtain 22.5 as an optimal choice, with number of peaks between 1 and 5.

- The MEM of $\gamma$ says that around this threshold the peaks behave as an exponential distribution.
Conclusions

- The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.
- From the conditional mean excess we obtain 22.5 as an optimal choice, with number of peaks between 1 and 5.
- The MEM of $\gamma$ says that around this threshold the peaks behave as an exponential $d$.
- Test of fit for the range of thresholds is $(22, 2, 22, 9)$ to the exponential $d$. 

Conclusions

- The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.
- From the conditional mean excess we obtain 22.5 as an optimal choice, with number of peaks between 1 and 5.
- The MEM of $\gamma$ says that around this threshold the peaks behave as an exponential distribution.
- Test of fit for the range of thresholds is $(22, 2, 22, 9)$ to the exponential distribution.
- 22.7 as has the largest $p$-value, and is the lowest threshold in the range where the exponential distribution fits perfectly.
Conclusions

- The application of this method needs a previous study of the data: seasonality, conditional mean excess plot, choice of the threshold, fit the peak excesses to the best distribution to estimate the return levels.
- From the conditional mean excess we obtain 22.5 as an optimal choice, with number of peaks between 1 and 5.
- The MEM of $\gamma$ says that around this threshold the peaks behave as an exponential distribution.
- Test of fit for the range of thresholds is $(22, 2, 22, 9)$ to the exponential distribution.
- 22.7 has the largest $p$-value, and is the lowest threshold in the range where the exponential distribution fits perfectly.
- We conclude an optimal estimate $V_{ref} = 30, 2$, with a 95%-confidence interval of $(27, 3, 35, 2)$, which suggest a WTGS of class III.
Further work

- Estimate $\sigma$ and $\gamma$ using MLE to compare the results obtained with MEM.
Further work

- Estimate $\sigma$ and $\gamma$ using MLE to compare the results obtained with MEM.
- Fit the peaks excesses over thresholds between 18 and 21.5 to the g.P.d with the graphical estimates $\sigma$ and $\gamma$ from the conditional mean excess plot.
Further work

- Estimate $\sigma$ and $\gamma$ using MLE to compare the results obtained with MEM.
- Fit the peaks excesses over thresholds between 18 and 21.5 to the g.P.d with the graphical estimates $\sigma$ and $\gamma$ from the conditional mean excess plot.
- Fit the Poisson distribution for the number of peak excesses and obtain confidence intervals for $\hat{\lambda}$ to see the effect to the confidence intervals of the return levels.
Further work

- Estimate $\sigma$ and $\gamma$ using MLE to compare the results obtained with MEM.
- Fit the peaks excesses over thresholds between 18 and 21.5 to the g.P.d with the graphical estimates $\sigma$ and $\gamma$ from the conditional mean excess plot.
- Fit the Poisson distribution for the number of peak excesses and obtain confidence intervals for $\hat{\lambda}$ to see the effect to the confidence intervals of the return levels.
- Check the extreme value condition $F \in D(G_\gamma)$. This domain of attraction is characterized for several authors (see De Haan and Ferreira’06).
References

References


References


References


References


References

▶ Smith, R.L. (2003), Statistics of extremes, with applications in environment, insurance and finances.