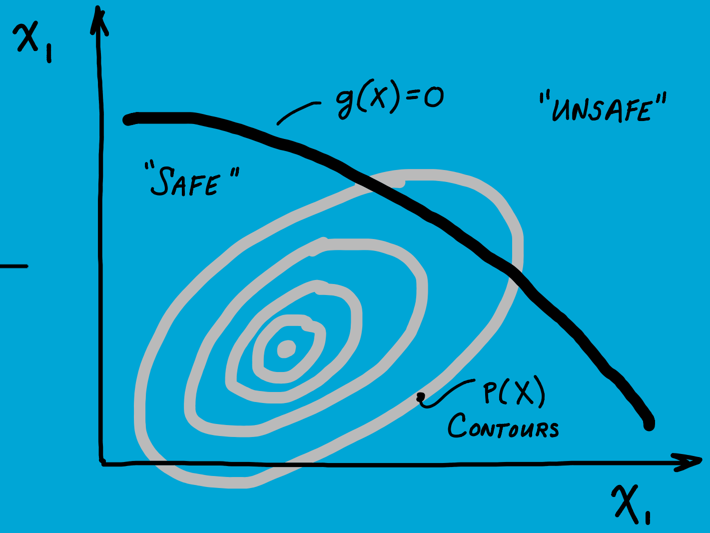
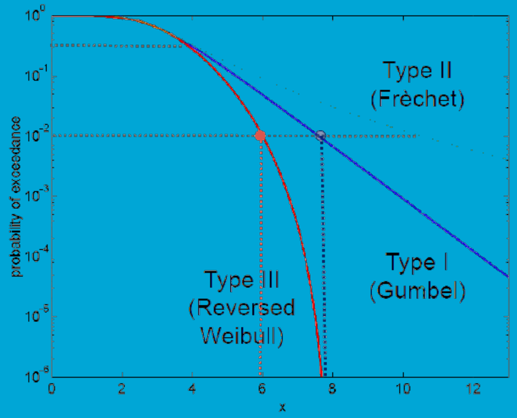
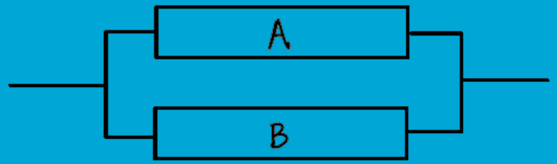
CIEM42X0 Probabilistic Design

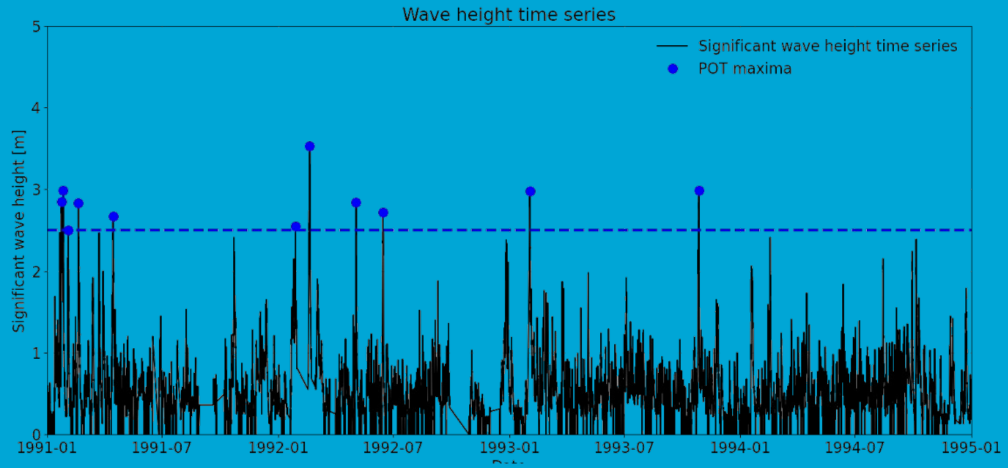
Hydraulic and Offshore Structures (HOS) Track

Civil Engineering MSc Program









EVA: Introduction

Patricia Mares Nasarre

TU Delft

Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context

2. Interpret and apply the concept of **return period and design life**

3. Apply **extreme value analysis** to datasets:

   a. Block maxima - GEV

   b. Peak over threshold (POT) – GPD

4. Apply techniques to **support the threshold selection** in POT

Extremes and Extreme Value Analysis

An **extreme observation** is an observation that **deviates from the average observations.**

Infrastructures and systems are designed to **withstand extreme conditions (ULS)**.

- Breakwater → wave storm
- Flood defences → floods, droughts

To properly design and assess infrastructures and system **we need to characterize the uncertainty of the loads**.

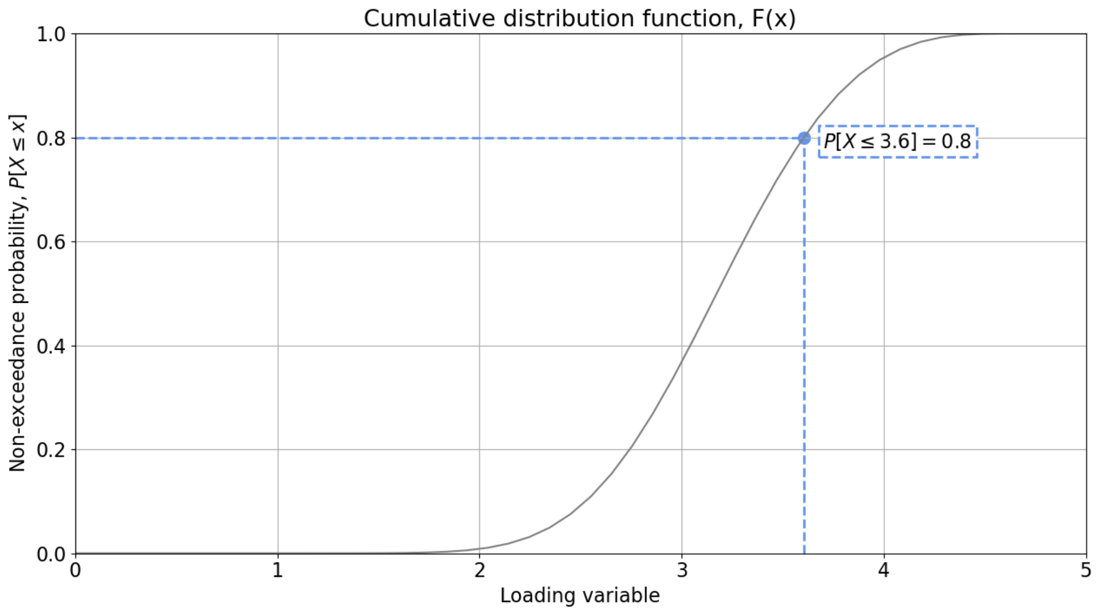
Extreme Value Analysis

Based on historical observed extremes (limited)…

- Allows us to **model** the stochastic behaviour of extreme events
- Allows us to **infer** extremes we have not observed yet (extrapolation)


Time series of **observations** of the loading variable

EVA



Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context

2. Interpret and apply the concept of **return period and design life**

3. Apply **extreme value analysis** to datasets:

   a. Block maxima - GEV

   b. Peak over threshold (POT) – GPD

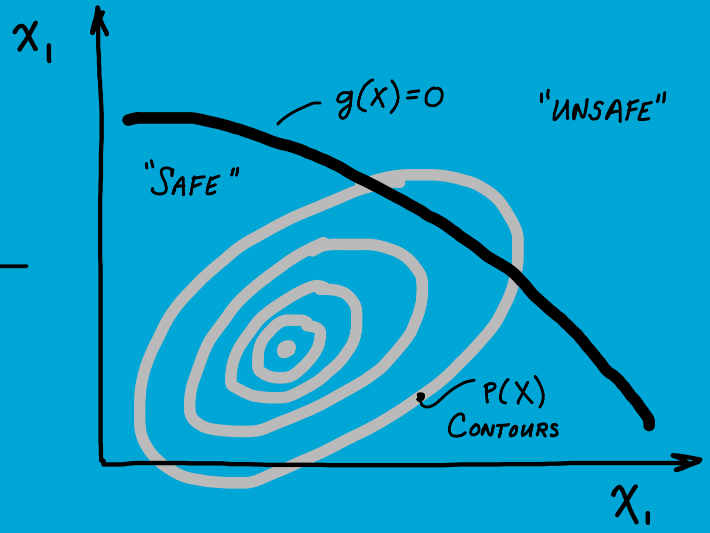
4. Apply techniques to **support the threshold selection** in POT

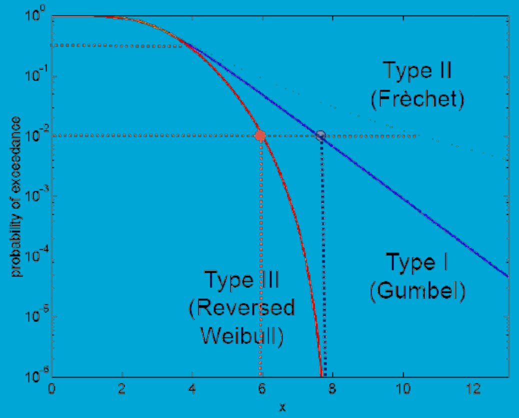
TU Delft

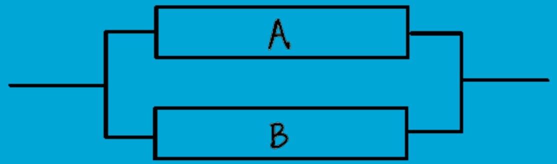
CIEM42X0 Probabilistic Design

Hydraulic and Offshore Structures (HOS) Track

Civil Engineering MSc Program



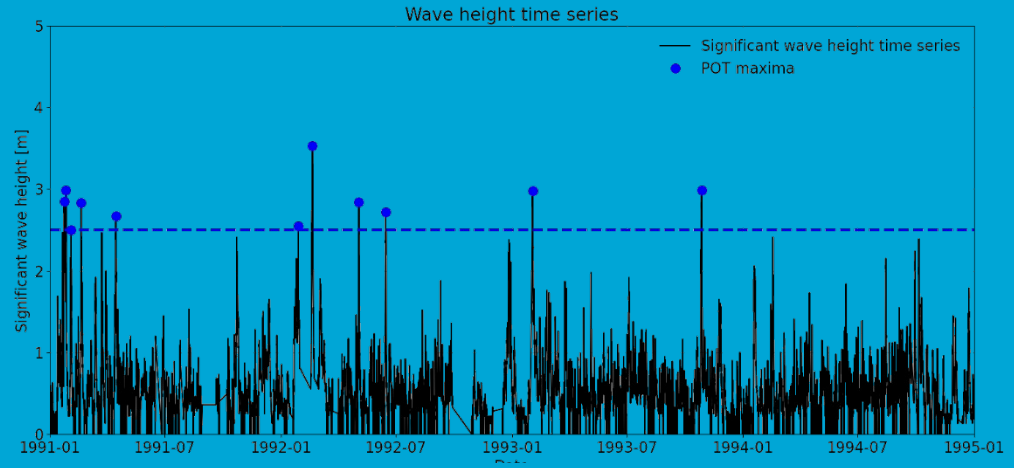




EVA: Design requirements.

Binomial distribution

Patricia Mares Nasarre



TU Delft

Learning objectives

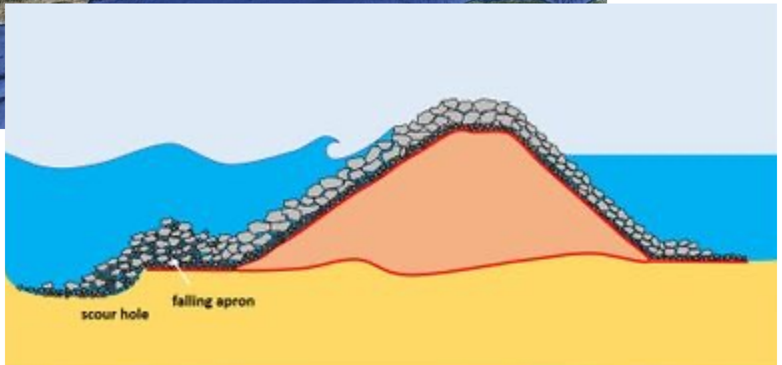
1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**

Bernoulli trial
Binomial distribution

3. Apply **extreme value analysis** to datasets:
    a. Block maxima - GEV
    b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

TU Delft

Example case: intervention in the Mediterranean coast



- It may be a coastal structure, a water intake, the restoration of a sandy beach, between others.
- Here: **design a mound breakwater**
- Mound breakwater must resist wave storms → H_s
- ***But which one?***

Design requirements

Regulations and recommendations → Exceedance probability or **return period**

Country	Standard	T_R **(years)**	DL (years)	$p_{f,DL}$ (-)
England	BS 6349-1-1:2013	**50-100***	50-100	0.05*
Japan	TS Ports-2009	**50-100**	50	0.40-0.64
Spain	ROM 0.0-01/1.0-09	**113-4,975**	25-50	0.01-0.2

*Not well defined

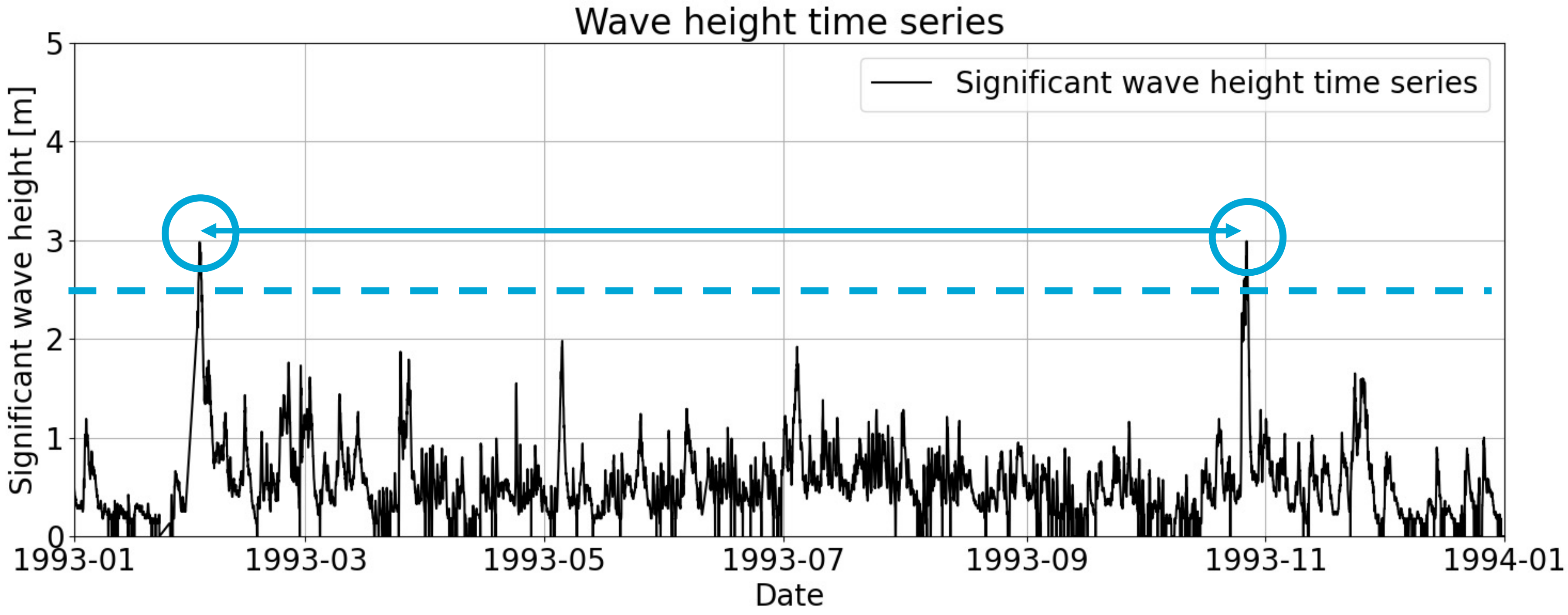
But what is return period?

Return Period

The Return Period (T_R) is the expected time between exceedances.
"In other words, we have to make, on average, $1/p_{f,y}$ trials in order that the event happens once" (Gumbel) or **wait $1/p_{f,y}$ years before the next occurrence**, being $p_{f,y}$ the exceedance probability.

Assumption of stationarity:
Every year the probability of the event being higher/lower than the threshold is always the same

$$T_R(t) = \frac{1}{p_{f,y}}$$



Design requirements

Regulations and recommendations → Exceedance probability or **return period**

But also Design Life and the probability of failure during the design life ($p_{f,DL}$)

Country	Standard	T_R (years)	DL (years)	$p_{f,DL}$ (-)
England	BS 6349-1-1:2013	50-100*	50-100	0.05*
Japan	TS Ports-2009	50-100	50	0.40-0.64
Spain	ROM 0.0-01/1.0-09	113-4,975	25-50	0.01-0.2

*Not well defined

Back to basics – Bernoulli process

Extremes can be assimilated as a Bernoulli process

Bernoulli process	Extremes
Two possible outcomes: success or failure	✓ Each observation can be an over or below
Outcomes are mutually exclusive and collectively exhaustive	✓ over vs. below the design value
Constant probability of success	✓ stationarity
Independence between trials	✓ Hypothesis of EVA *iid* events

TU Delft

Back to basics – Binomial distribution

Extremes can be assimilated as a Bernoulli process

Number of exceedances (succeses) in a given number of trials follows a Binomial distribution

$$p_X(x) = P[X = x|n,p] = \binom{n}{x} p^x (1-p)^{n-x} \qquad for\ x = 0, 1, \dots, n; p \in [0,1]$$

$$p_X(x) = P[X = x|n,p] = 0 \qquad otherwise$$

where

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

TUDelft

Design requirements

Regulations and recommendations → Exceedance probability or **return period**

But also Design Life and the probability of failure during the design life (p_{DL})

Country	Standard	RT (years)	DL (years)	$p_{f,DL}$ (-)
England	BS 6349-1-1:2013	50-100*	50-100	0.05*
Japan	TS Ports-2009	50-100	50	0.40-0.64
Spain	ROM 0.0-01/1.0-09	113-4,975	25-50	0.01-0.2

*Not well defined

Any idea now?

TUDelft

Design requirements – Binomial distribution

- $\boxed{p_{f,DL} - p_{f,y} - DL - T_R}$ $T_R = 1/p_{f,y}$
- Each year is a trial ⇨ Success (excess the design value) or failure (no excess)?
- The number of exceedances (successes) in a given number of years (trials) ~ Binomial

M
O
D
E
L

- $p_{f,DL}$ is the probability of an excess at least once in the DL
- $p_{f,DL} = 1 - \underbrace{\text{probability of no excess}}$

$$p_X(0) = P[X = 0|DL, p_{f,y}] = \binom{DL}{0} {p_{f,y}}^0 (1 - p_{f,y})^{DL-0}$$

- $p_{f,DL} = 1 - (1 - p_{f,y})^{DL}$

$$\boxed{T_R = \frac{1}{p_{f,y}} = \frac{1}{1 - (1 - p_{f,DL})^{1/DL}}}$$

TU Delft

Design requirements – Binomial distribution

$$\boxed{T_R = \frac{1}{\mathrm{p_{f,y}}} = \frac{1}{1-(1-\mathrm{p_{f,DL}})^{1/DL}}}$$

- DL = 20 years
- $\mathrm{p_{f,DL}}$ = 0.20

$$T_R = \frac{1}{\mathrm{p_{f,y}}} = \frac{1}{1-(1-0.2)^{1/20}} \approx 90\ years$$

$$\mathrm{p_{f,y}} \approx 0.011$$

Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**

Bernoulli trial
Binomial distribution

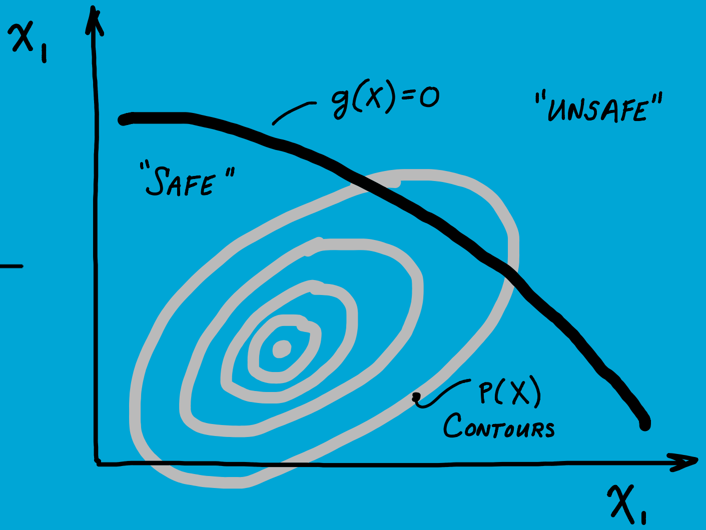
3. Apply **extreme value analysis** to datasets:
    a. Block maxima - GEV
    b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

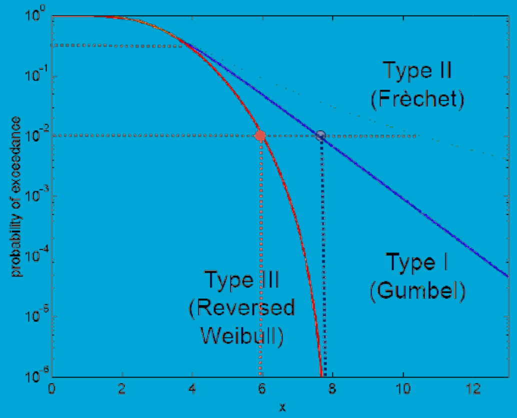
TU Delft

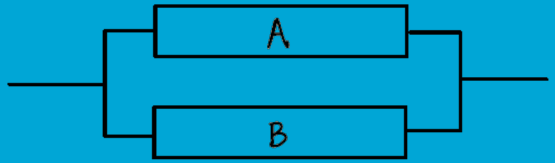
CIEM42X0 Probabilistic Design

Hydraulic and Offshore Structures (HOS) Track

Civil Engineering MSc Program

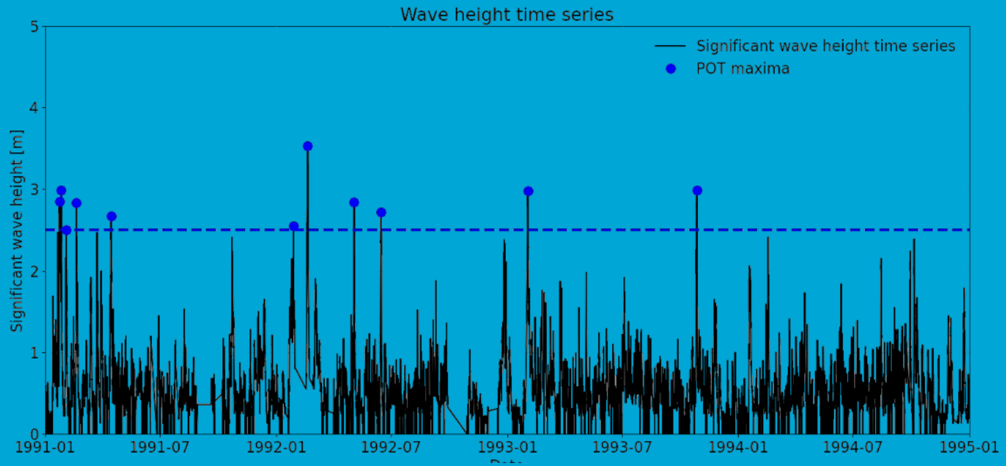






EVA: Block Maxima and GEV

Patricia Mares Nasarre

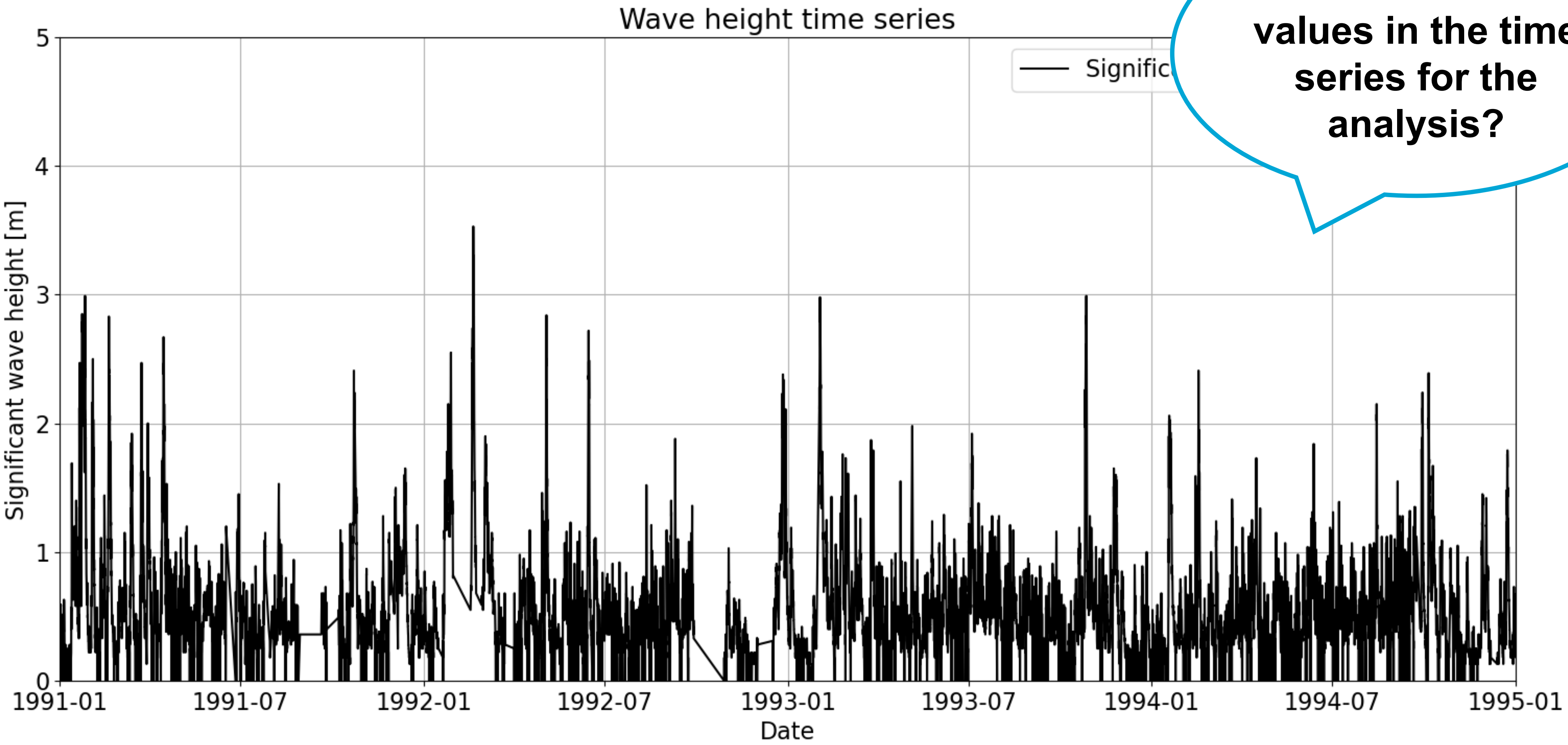


TU Delft

Learning objectives

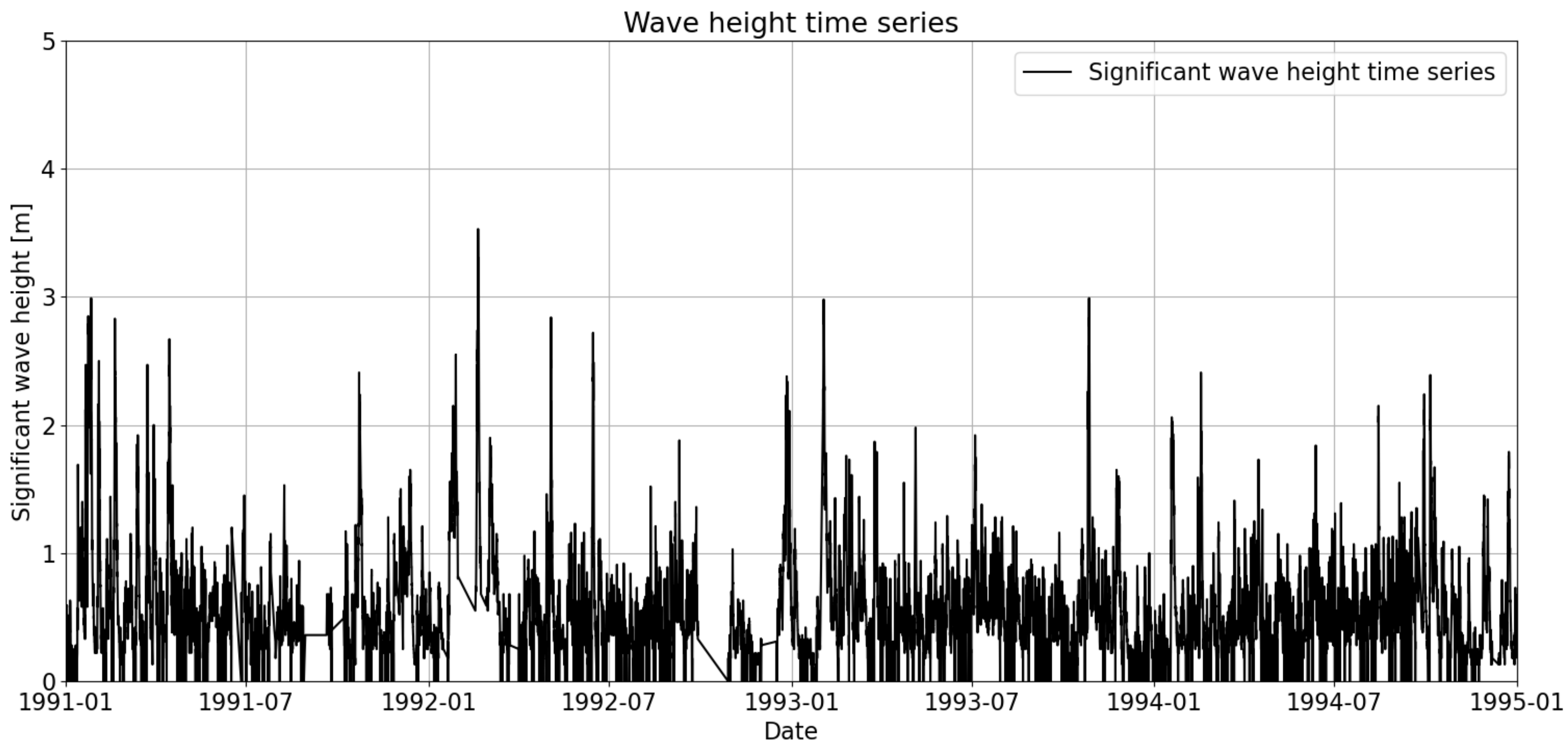
1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**
3. Apply **extreme value analysis** to datasets:
   a. Block maxima - GEV
   b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

Time series



Can I use all the values in the time series for the analysis?

We need to sample extreme values!

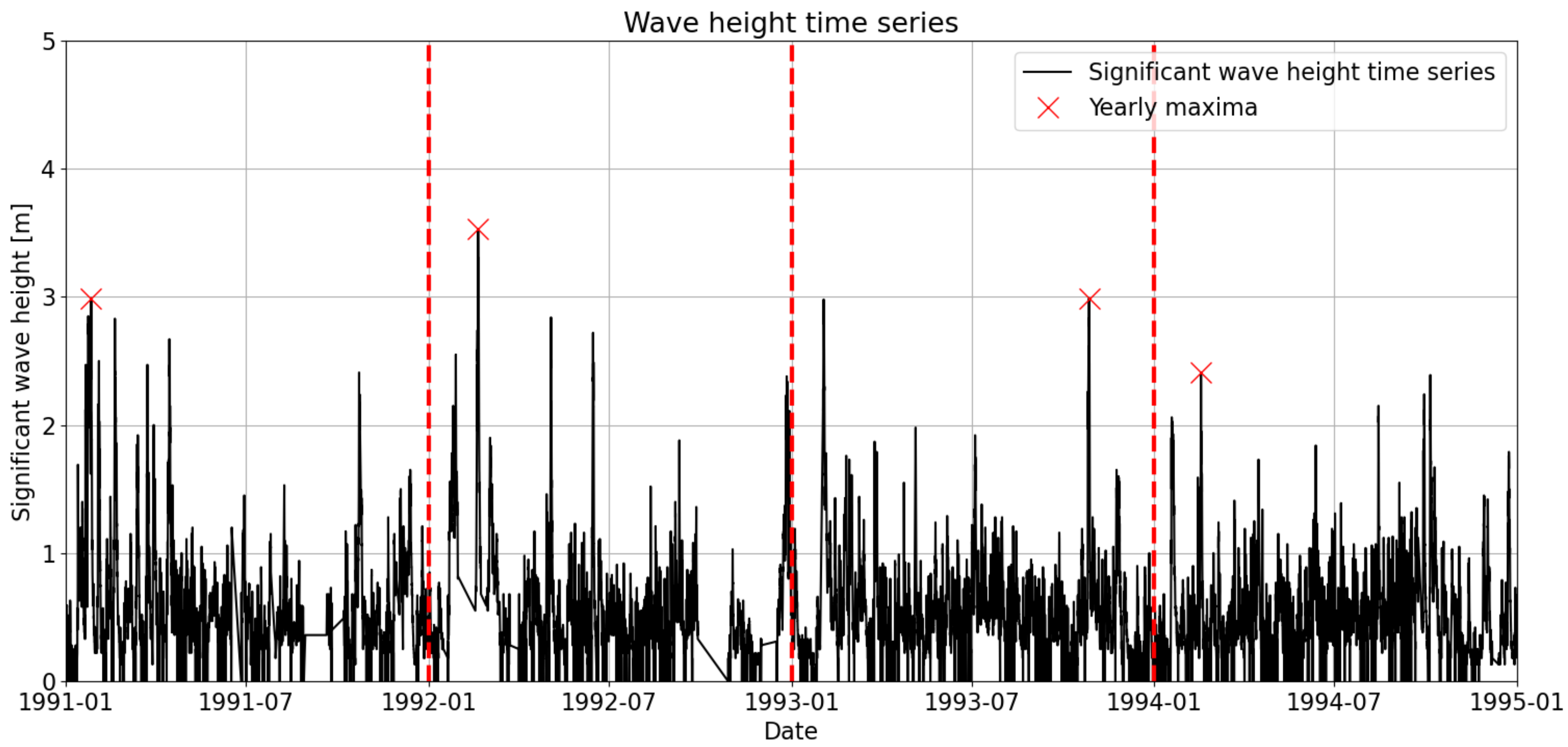


Two techniques:

1. **Block Maxima**
2. **Peak Over Threshold (POT)**

TU Delft

Sampling extremes: Block Maxima



1. **Block Maxima** (typically block=1year)
   - Maximum value within the block
   - Number of selected events=number of blocks recorded (e.g.: number of years)
   - Easy to implement

Generalized Extreme Value Distribution

- We are interested in modelling the maximum of the sequence $X = X_1, \dots, X_n$ of *iid* random variables, $M_n = \max(X_1, \dots, X_n)$, where *n* is the number of observations in a given block.

- We can prove that for large *n*, **those maxima tend to the Generalized Extreme Value (GEV) family of distributions, regardless the distribution of *X*.**

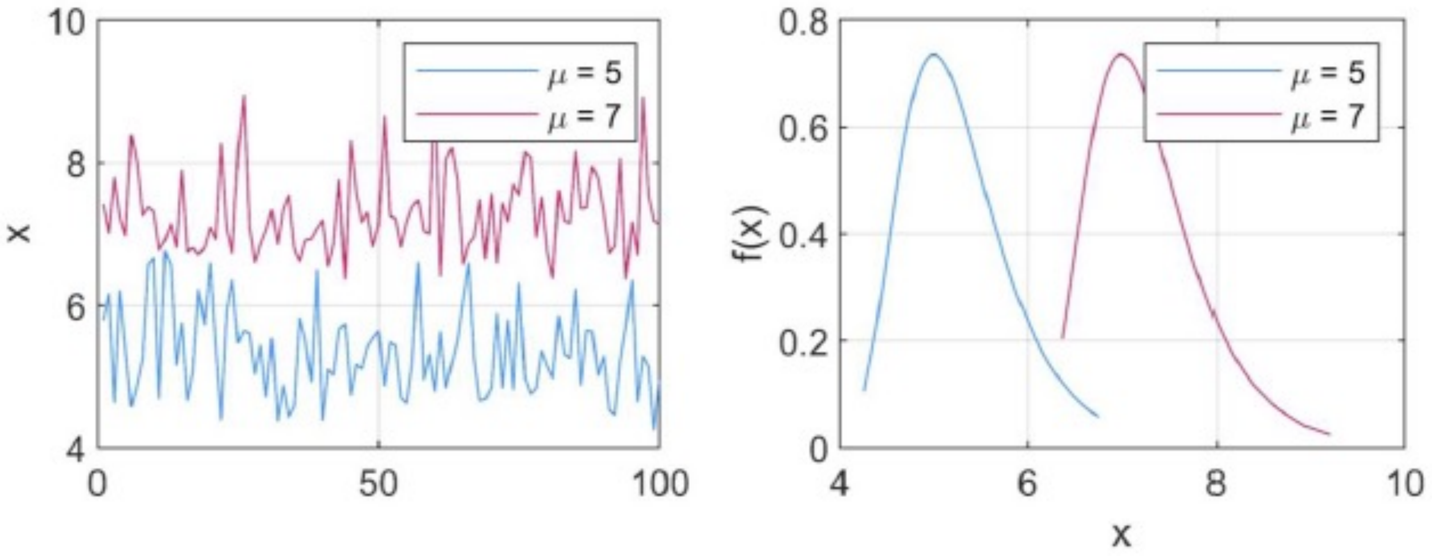
$$P[M_n \leq x] \rightarrow G(x)$$

Generalized Extreme Value Distribution

Generalized Extreme Value is defined as

$$G(x) = exp - [1 + \xi \frac{x-\mu}{\sigma}]^{-1/\xi} \qquad (1 + \xi \frac{x-\mu}{\sigma}) > 0$$

With parameters location ($-\infty < \mu < \infty$), scale ($\sigma > 0$) and shape ($-\infty < \xi < \infty$).



Location parameter (μ)

Higher μ, right displacement of the distribution, higher values.

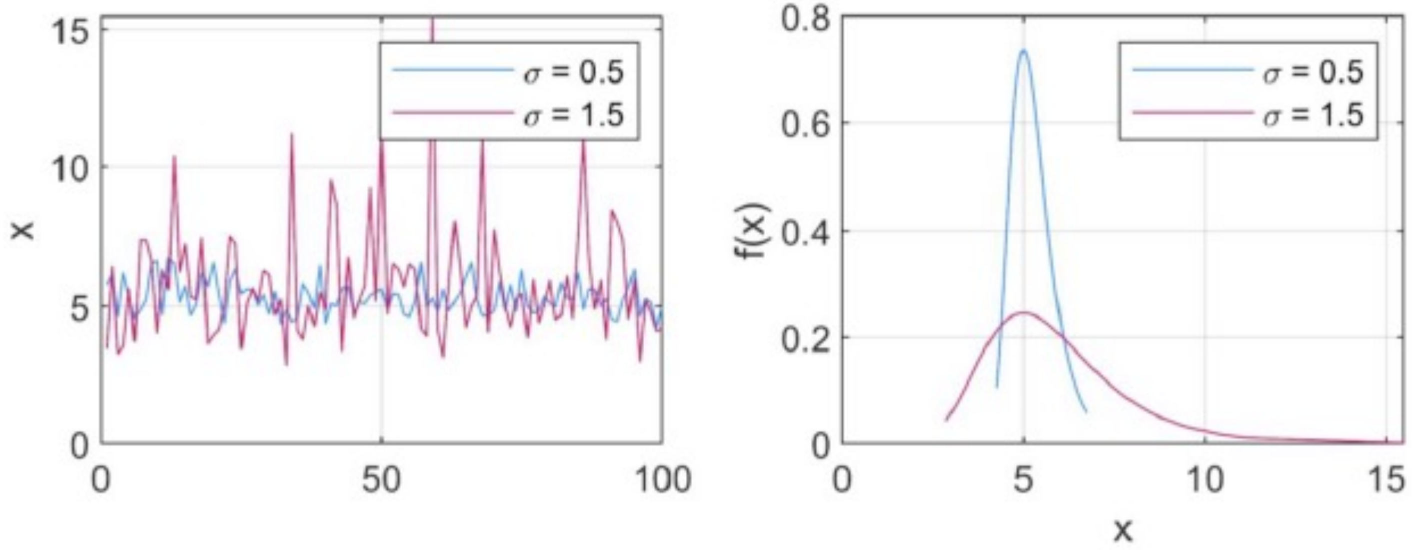
TUDelft

Generalized Extreme Value Distribution

Generalized Extreme Value is defined as

$$G(x) = exp - [1 + \xi \frac{x-\mu}{\sigma}]^{-1/\xi} \qquad (1 + \xi \frac{x-\mu}{\sigma}) > 0$$

With parameters location ($-\infty < \mu < \infty$), scale ($\sigma > 0$) and shape ($-\infty < \xi < \infty$).



Scale parameter (σ)

Higher $\boldsymbol{\sigma}$, wider distribution.

TU Delft

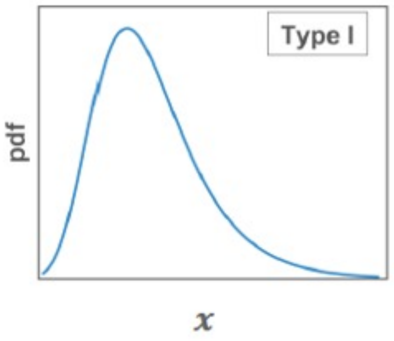
Generalized Extreme Value Distribution

Generalized Extreme Value is defined as

$$G(x) = exp-[1+\xi\frac{x-\mu}{\sigma}]^{-1/\xi} \qquad (1+\xi\frac{x-\mu}{\sigma}) > 0$$

With parameters location ($-\infty < \mu < \infty$), scale ($\sigma > 0$) and shape ($-\infty < \xi < \infty$).

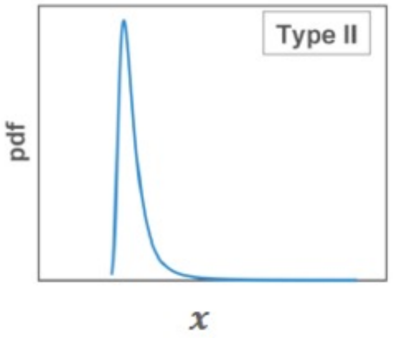
Gumbel



$\xi \to 0$

Exponential decay

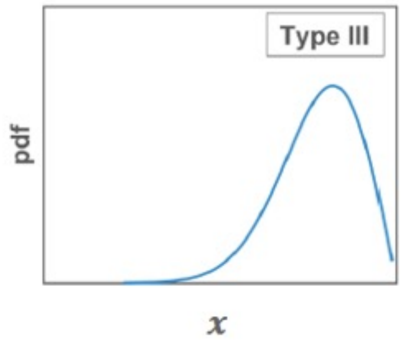
Fréchet



$\xi > 0$

Polynomial decay

Reverse Weibull



$\xi < 0$

Upper bound $(\mu - {}^{\sigma}/_{\xi})$

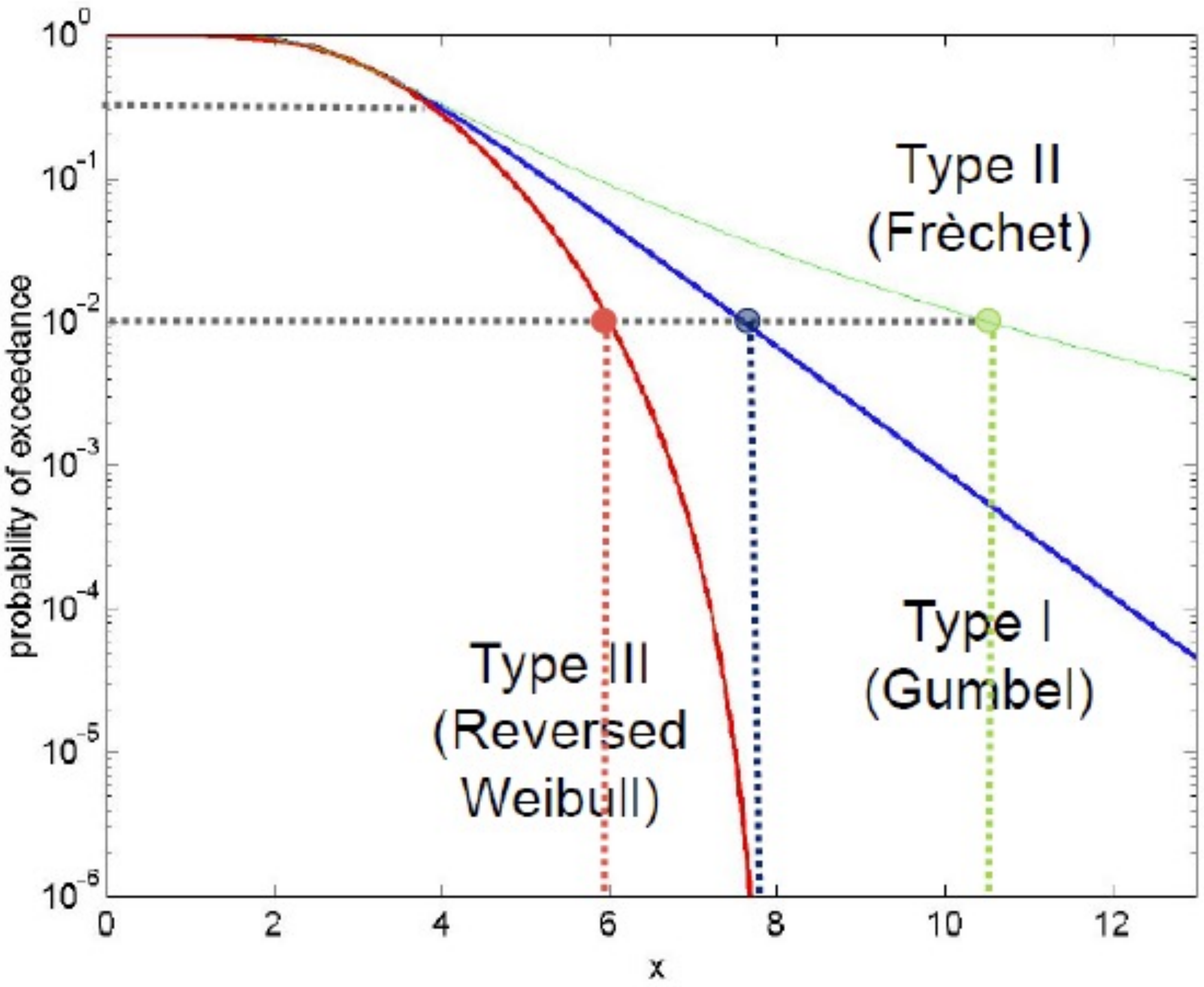
Shape parameter (ξ)

Determines the tail of the distribution.

Generalized Extreme Value Distribution

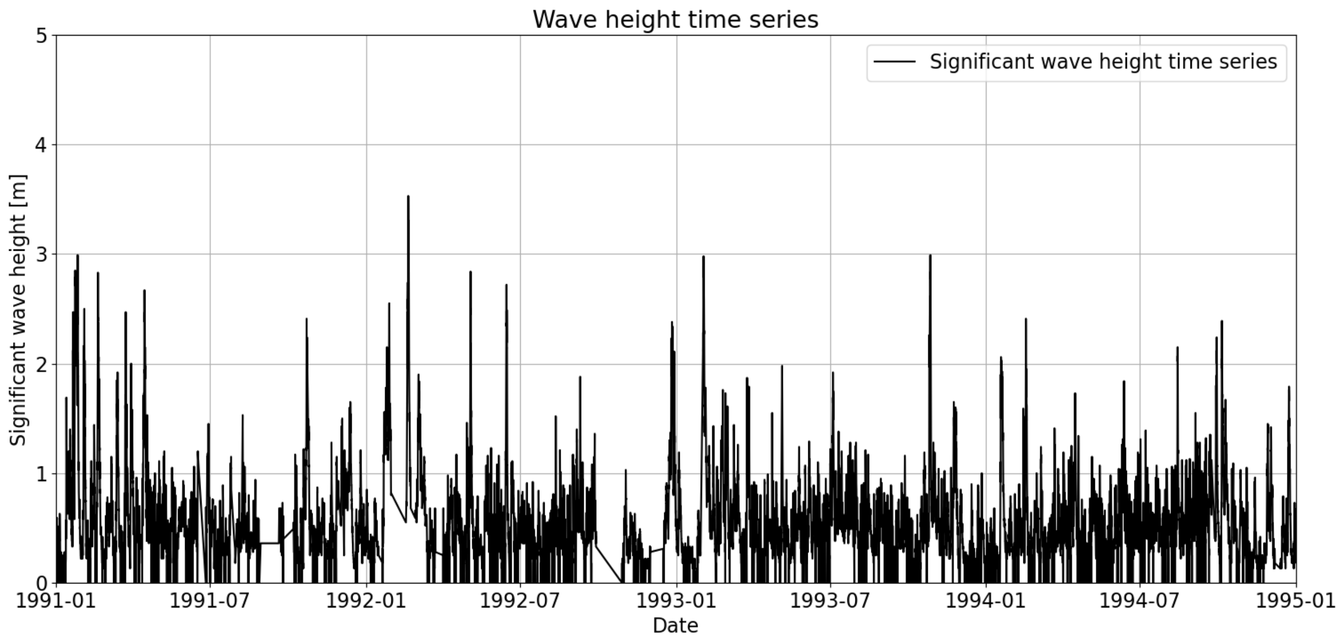
Plotting the tails…

- **Gumbel:** light tail
- **Fréchet:** heavy tail
- **Reversed Weibull:** bounded at $x = \mu - \frac{\sigma}{\xi}$



TU Delft

Let's apply it



- **Load: significant wave height (T_R=90 years)**
- 20 years of hourly measurements → **20 yearly maxima samples**

read observations

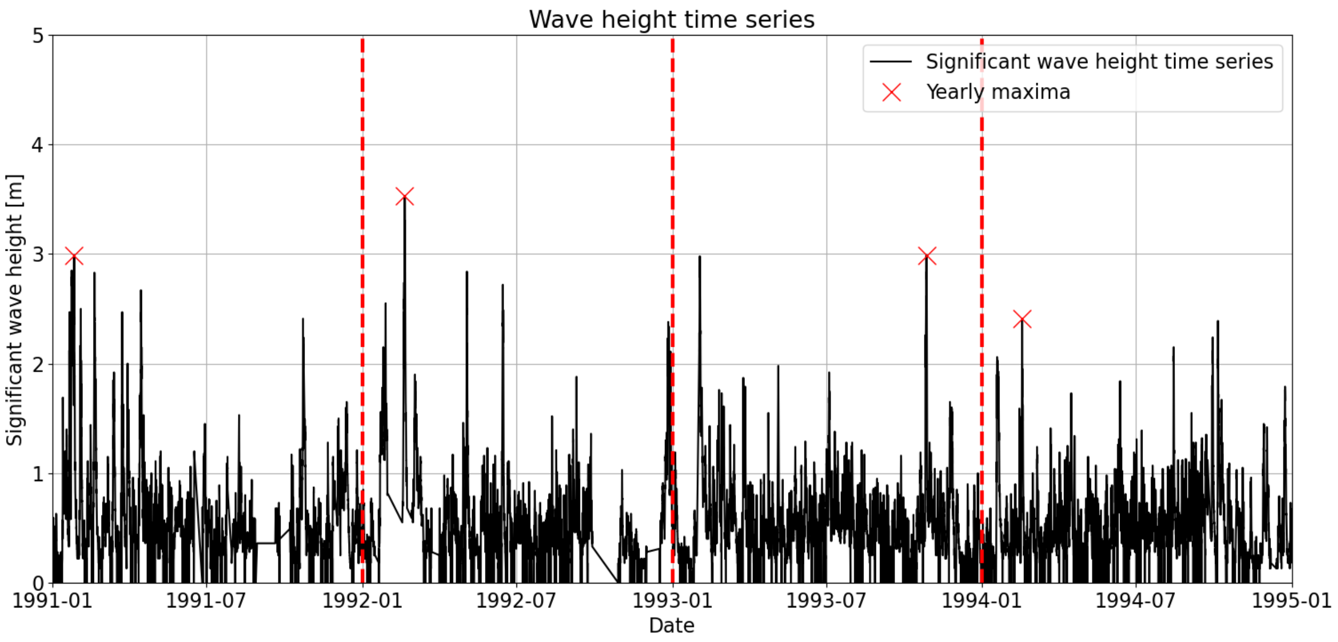
for each year i:
obs_max[i] = max(observations in year i)
end

fit GEV(obs_max)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

inverse GEV to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**
- 20 years of hourly measurements → **20 yearly maxima samples**

read observations

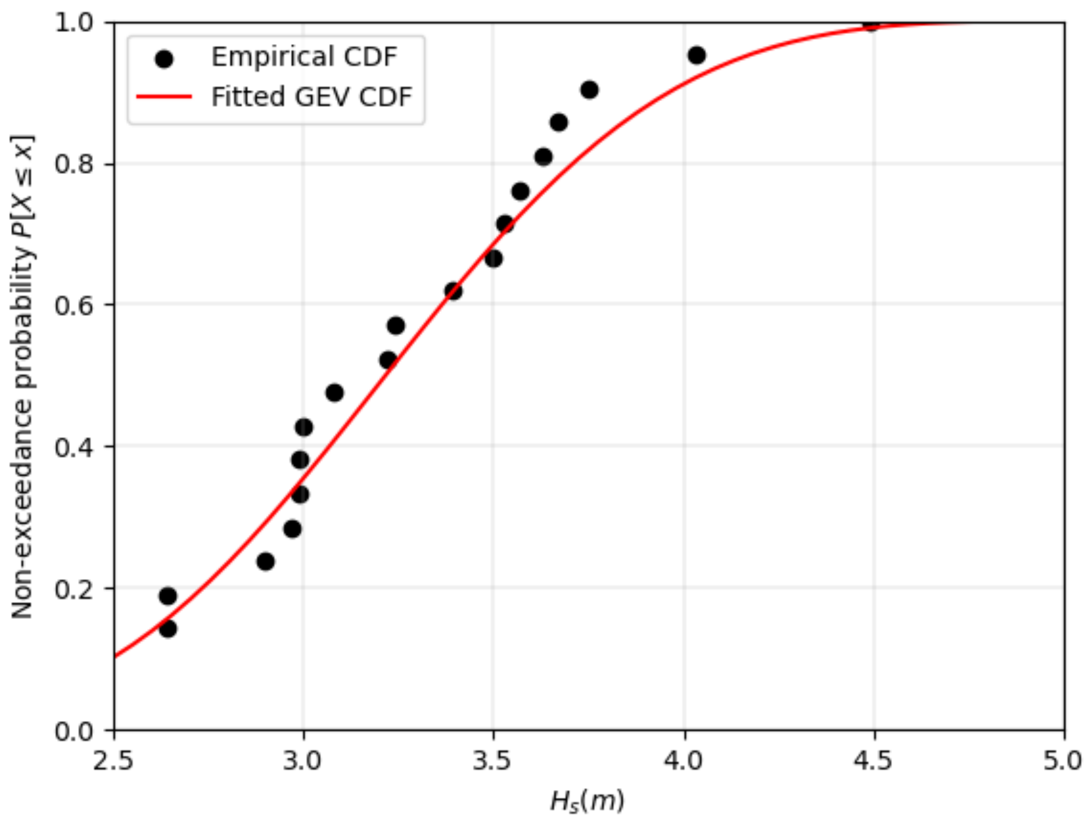
for each year i:
obs_max[i] = max(observations in year i)
end

fit GEV(obs_max)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

inverse GEV to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**
- 20 years of hourly measurements → **20 yearly maxima samples**

read observations

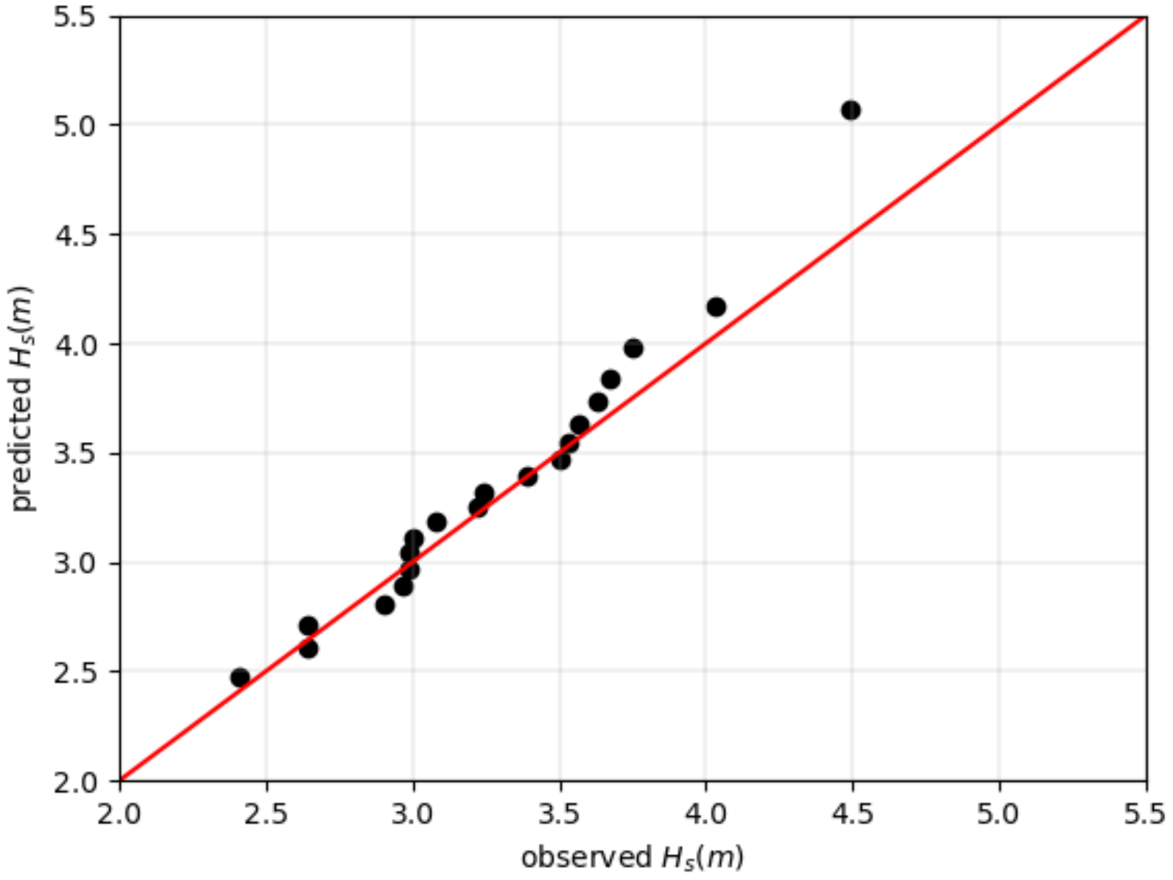
for each year i:
obs_max[i] = max(observations in year i)
end

fit GEV(obs_max)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

inverse GEV to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**
- 20 years of hourly measurements → **20 yearly maxima samples**

read observations

for each year i:
obs_max[i] = max(observations in year i)
end

fit GEV(obs_max)

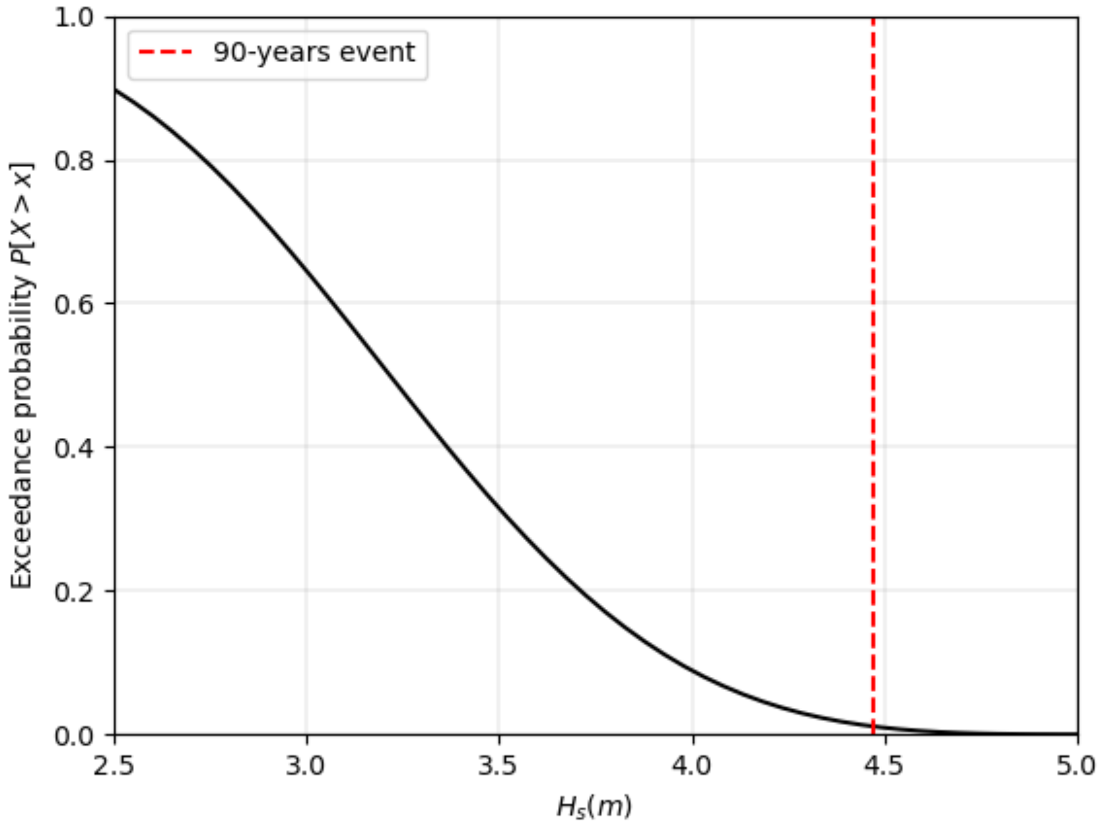
check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

inverse GEV to determine the design event

TUDelft

Let's apply it

$$z_p = G^{-1}(1-p_{f,y}) = \begin{cases} \mu - \frac{\sigma}{\xi}[1-\{-log(1-p_{f,y})\}^{-\xi}] & for\ \xi \neq 0 \\ \mu - \sigma log\{1-p_{f,y}\} & for\ \xi = 0 \end{cases}$$



- **Load: significant wave height (T_R=90 years)**
- 20 years of hourly measurements → **20 yearly maxima samples**

read observations

for each year i:
obs_max[i] = max(observations in year i)
end

fit GEV(obs_max)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

inverse GEV to determine the design event

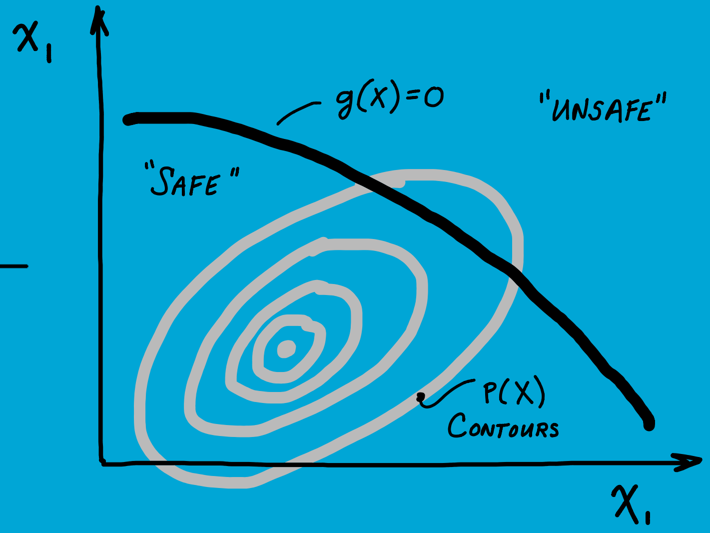
Learning objectives

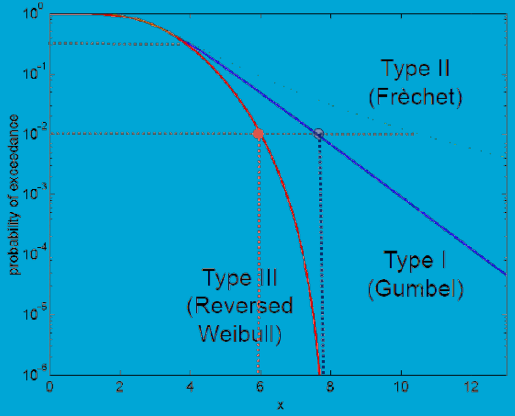
1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**
3. Apply **extreme value analysis** to datasets:
   a. Block maxima - GEV
   b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

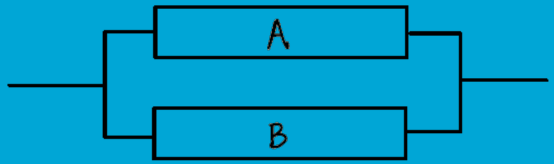
CIEM42X0 Probabilistic Design

Hydraulic and Offshore Structures (HOS) Track

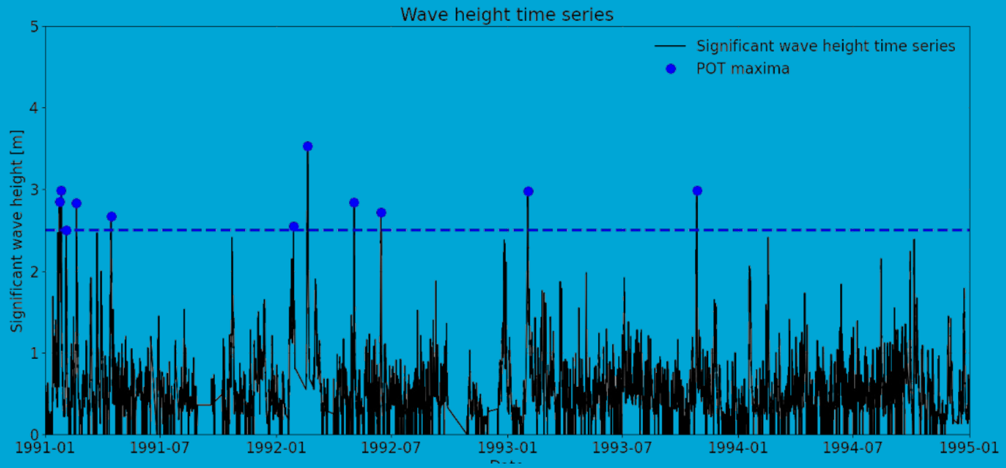
Civil Engineering MSc Program







EVA: POT and GPD (I)



Patricia Mares Nasarre

TU Delft

Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context

2. Interpret and apply the concept of **return period and design life**

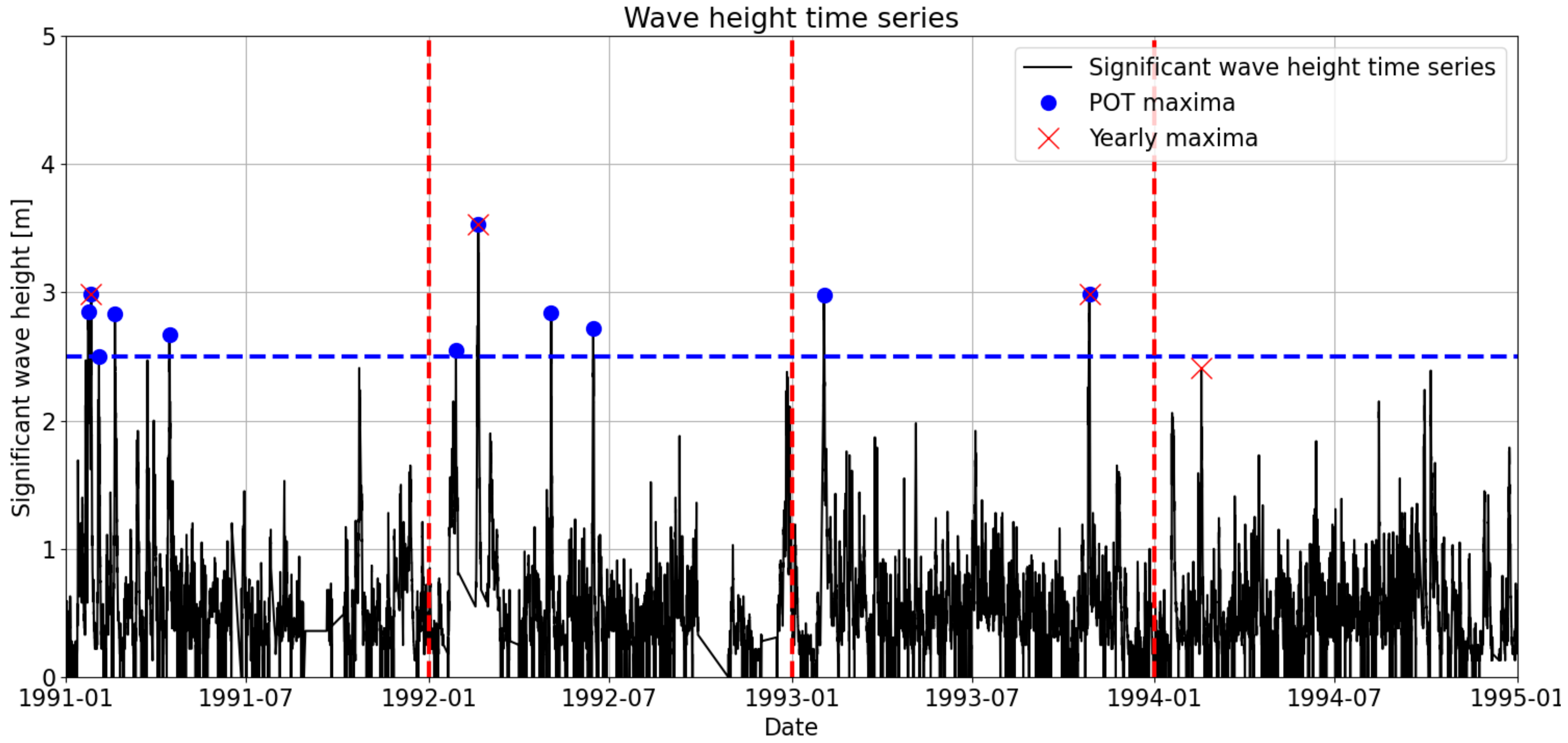
3. Apply **extreme value analysis** to datasets:

   a. Block maxima - GEV

   b. Peak over threshold (POT) – GPD

4. Apply techniques to **support the threshold selection** in POT

Sampling extremes: Peak Over Threshold (POT)

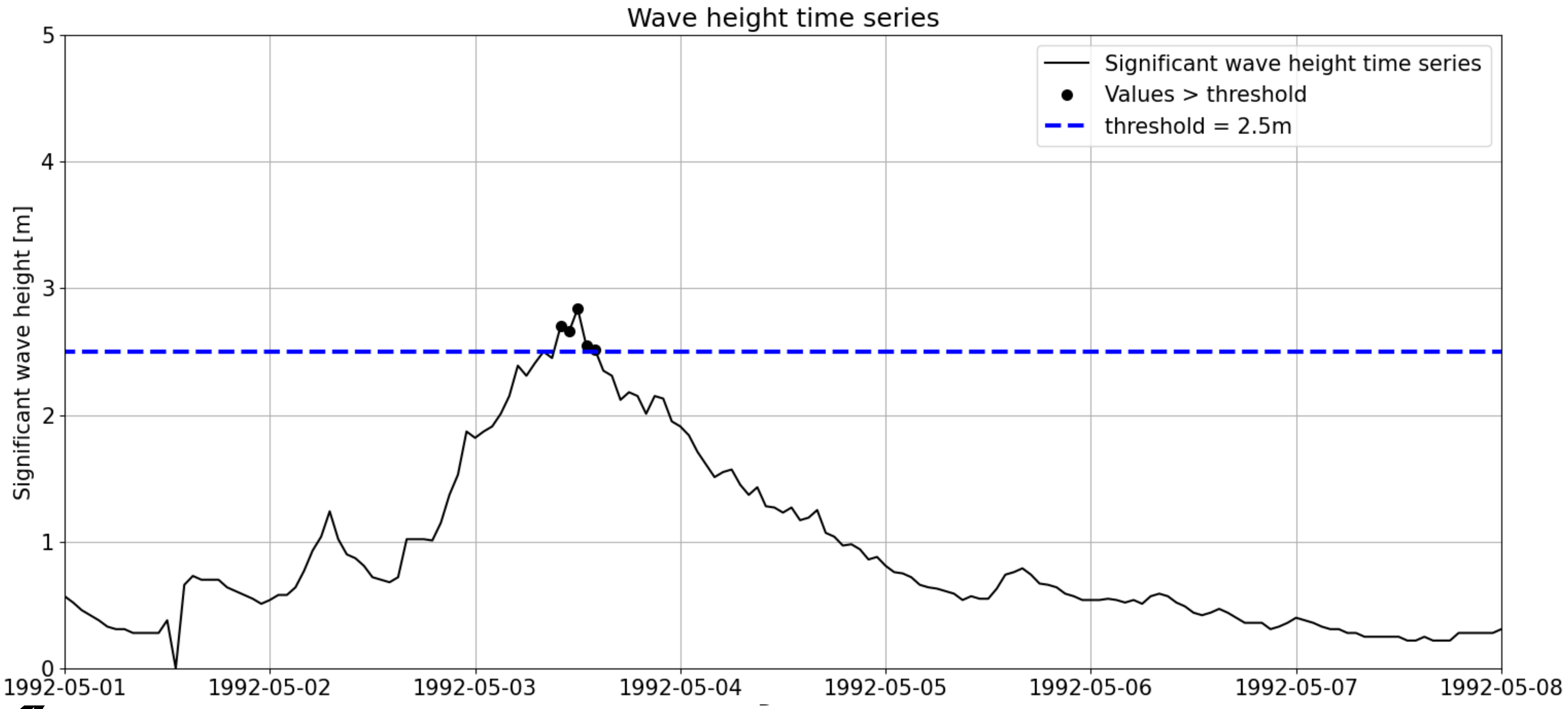


2. **Peak Over Threshold (POT)**

- Usually, higher number of extremes identified
- Additional parameters:
    - Threshold (*th*)
    - Declustering time (*dl*)

Choosing POT parameters

Basic assumption of EVA: extremes are *iid* ⟹ *th* and *dl* should be chosen so the identified extreme events are independent.



Extremes cluster in time!

If *dl* is big enough, we ensure that extremes do not belong to the same storm.

dl → *th,* physical phenomena (local conditions)

TU Delft

Generalized Pareto Distribution

- The maximum of the sequence $X = X_1, \dots, X_n$ of *iid* random variables, $M_n = \max(X_1, \dots, X_n)$, where *n* is the number of observations in a given block, follows **the Generalized Extreme Value (GEV) family of distributions, regardless the distribution of *X*** for large *n*.

$$P[M_n \leq x] \rightarrow G(x)$$

- If that is true, **the distribution of the excesses can be approximated by a Generalized Pareto distribution**.

$$F_{th} = P[X - th \leq x | X > th] \rightarrow H(y)$$

- where the excesses are defined as *Y*=*X*−*th* for *X*>*th*

TU Delft

Generalized Pareto Distribution

Generalized Pareto distribution of the excesses is defined as

$$H(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{y}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

where $y \geq 0$ if $\xi \geq 0$, and $0 \leq y \leq -\frac{\sigma_{th}}{\xi}$ if $\xi < 0$.

These are conditional probabilities to *X>th*. As function of the random variable *X* and the threshold *th*

$$P[X < x | X > th] = \begin{cases} 1 - \left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{x-th}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

TU Delft

Generalized Pareto Distribution

$$P[X < x | X > th] = \begin{cases} 1 - \left(1 + \frac{\xi(x - th)}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{x - th}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

With parameters threshold (*th*>0), pareto's scale ($\sigma_{th} > 0$) and shape ($-\infty < \xi < \infty$).

Relationship with GEV's parameters

- Shape parameter is the same
- σ_{th} defined based on GEV's parameters as

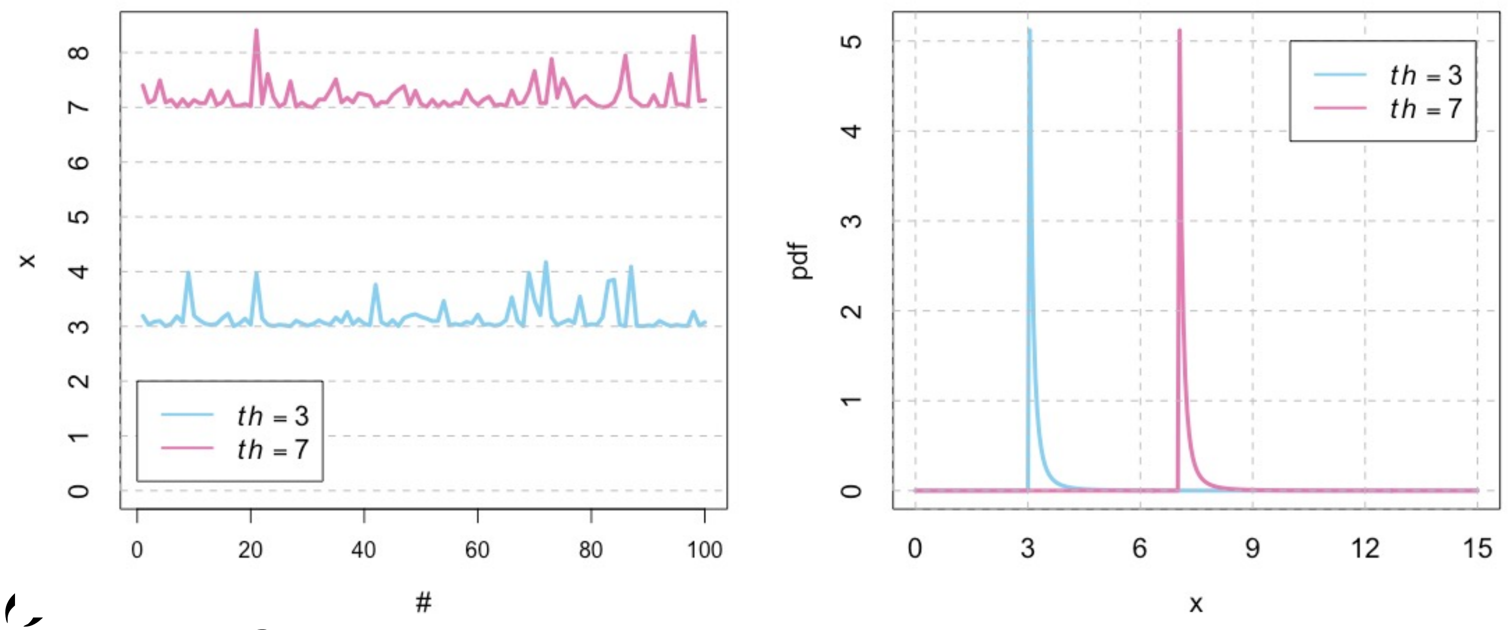
$$\sigma_{th} = \sigma + \xi(th - \mu)$$

TU Delft

Generalized Pareto Distribution

$$P[X < x | X > th] = \begin{cases} 1 - \left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{x-th}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

With parameters threshold (*th*>0), pareto's scale ($\sigma_{th} > 0$) and shape ($-\infty < \xi < \infty$).



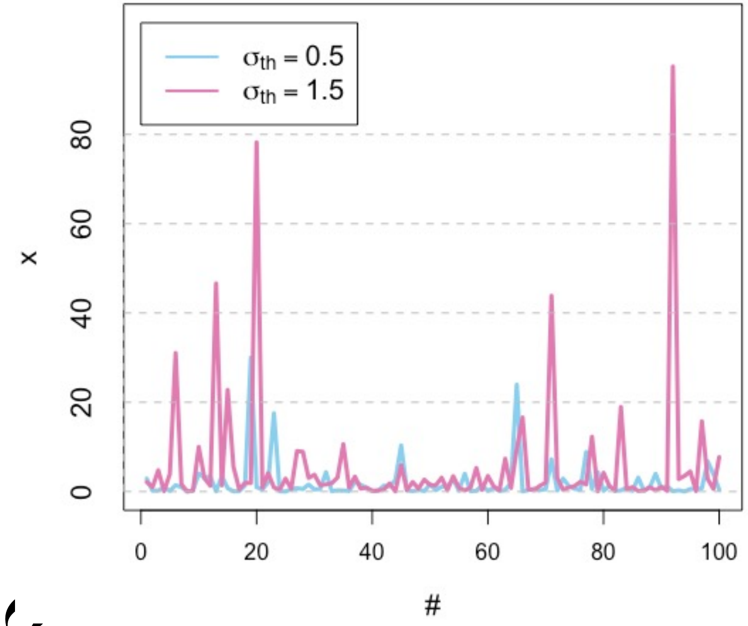
Threshold (*th*)

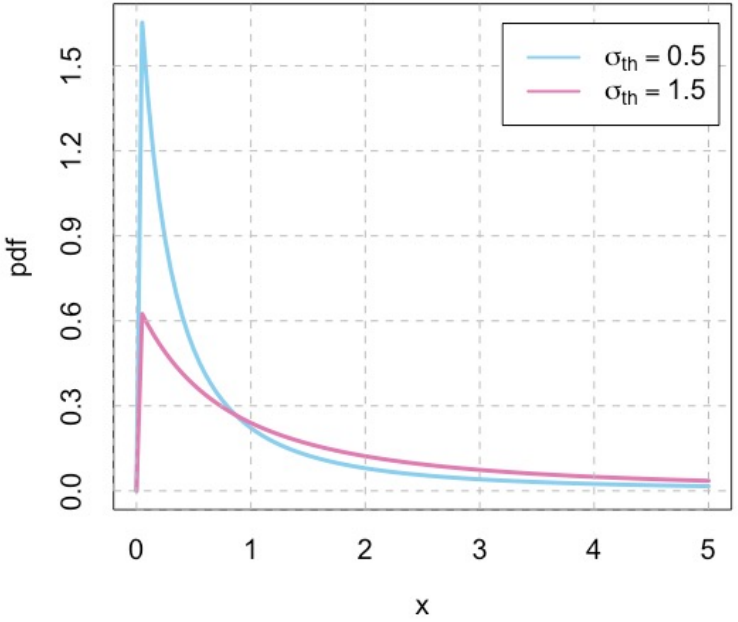
Acts like a location parameter.

Generalized Pareto Distribution

$$P[X < x|X > th] = \begin{cases} 1 - \left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{x-th}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

With parameters threshold (*th*>0), pareto's scale ($\sigma_{th} > 0$) and shape ($-\infty < \xi < \infty$).





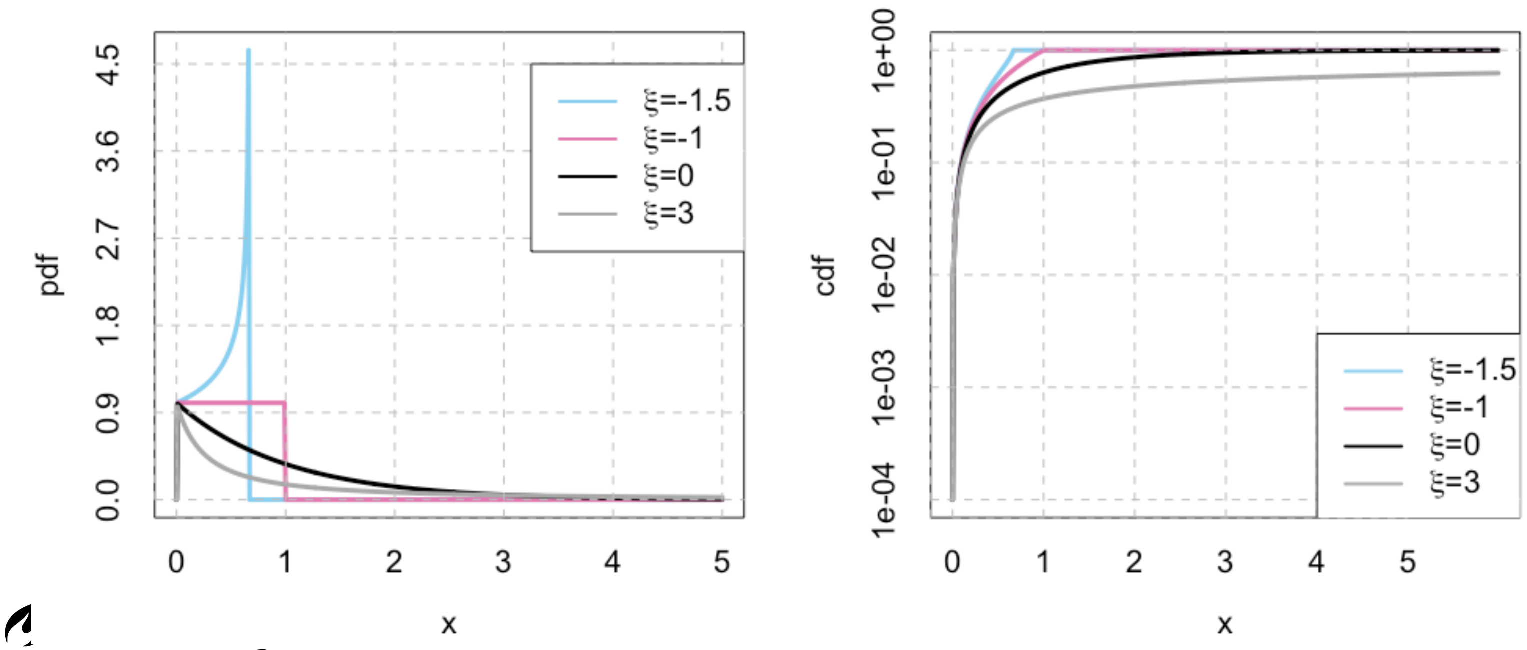
Scale parameter ($\boldsymbol{\sigma_{th}}$)

Higher $\boldsymbol{\sigma_{th}}$, wider distribution.

Generalized Pareto Distribution

$$P[X < x|X > th] = \begin{cases} 1 - \left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi} & for\ \xi \neq 0 \\ 1 - exp\left(-\frac{x-th}{\sigma_{th}}\right) & for\ \xi = 0 \end{cases}$$

With parameters threshold (*th*>0), pareto's scale ($\sigma_{th} > 0$) and shape ($-\infty < \xi < \infty$).



Shape parameter (ξ)

- ξ<0: upper bound
- ξ>0: heavy tail
- ξ=0 & *th* = 0: Exponential
- ξ=-1: Uniform

TU Delft

GPD: m return levels

We are interested in the N-year return level x_N of the studied variable, which is expected to be exceeded once every N years.

We have already fitted a GPD with ξ>0 as

$$P[X > x | X > th] = \left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi}$$

Which is a conditional probability! Accounting for the probability of observing and excess (ζ_{th})

$$P[X > x] = P[X > th]\; P[X > x | X > th] = \zeta_{th}\left(1 + \frac{\xi(x-th)}{\sigma_{th}}\right)^{-1/\xi}$$

Then, the return level x_m exceeded in average every m observations is computed as

$$1/m = \zeta_{th}\left[1 + \xi\frac{(x_m-th)}{\sigma_{th}}\right]^{-1/\xi} \Longrightarrow x_m = th + \frac{\sigma_{th}}{\xi}\left[(m\zeta_{th})^{\xi} - 1\right]$$

x_m is the m-observations return level

GPD: from m observations to N years

We are interested in the *N*-year return level x_N of the studied variable.

The x_m return level is given by

$$x_m = th + \frac{\sigma_{th}}{\xi}\left[(m\zeta_{th})^{\xi} - 1\right]$$

To go from *m* observations to *N* years, we need to account for the number of observations each year n_y as

$$m = N \times n_y$$

Applied to the previous expression, we obtain the *N*-year return level as

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}\left[(Nn_y\zeta_{th})^{\xi} - 1\right] & for\ \xi \neq 0 \\ th + \sigma_{th} log(Nn_y\zeta_{th}) & for\ \xi = 0 \end{cases}$$

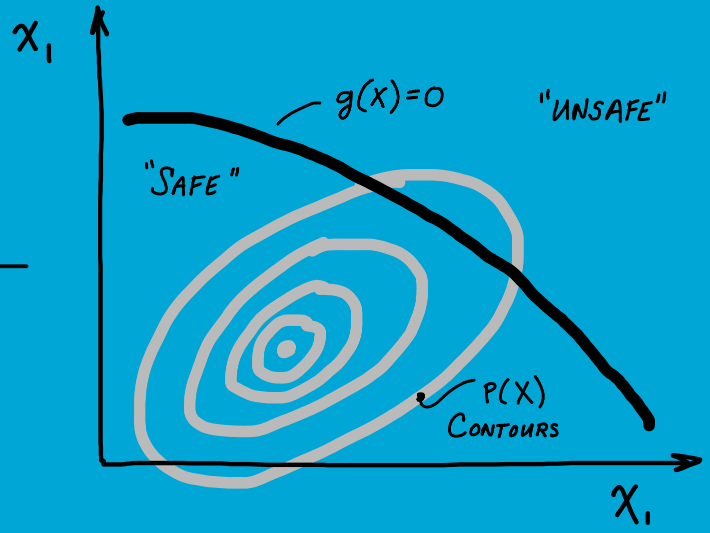
But how can I calculate ζ_{th}?

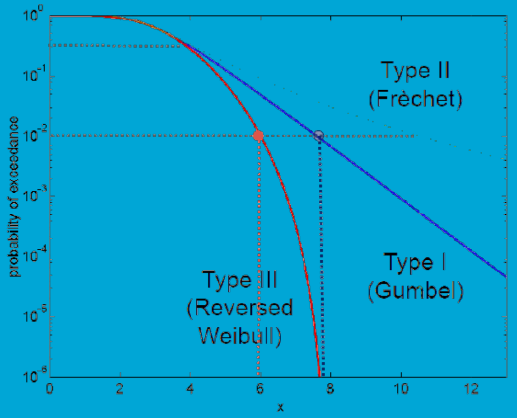
TU Delft

CIEM42X0 Probabilistic Design

Hydraulic and Offshore Structures (HOS) Track

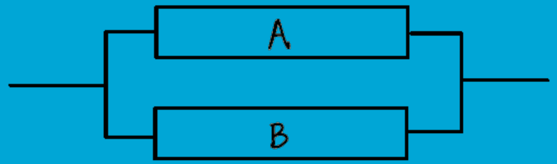
Civil Engineering MSc Program



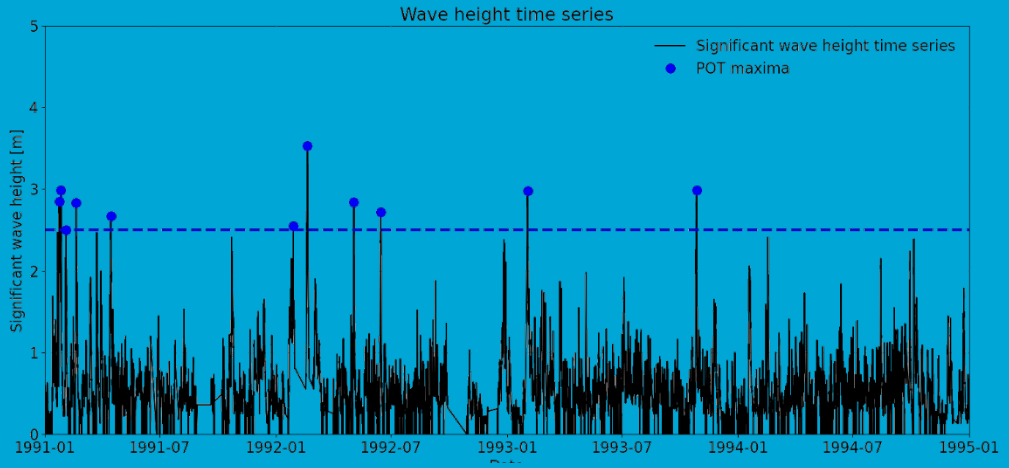


EVA: POT and GPD (II).

Poisson approximation to Binomial.



Patricia Mares Nasarre



TU Delft

Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context

2. Interpret and apply the concept of **return period and design life**

3. Apply **extreme value analysis** to datasets:

   a. Block maxima - GEV

   b. Peak over threshold (POT) – GPD

Poisson distribution

4. Apply techniques to **support the threshold selection** in POT

TU Delft

GPD: from m observations to N years

We are interested in the *N*-year return level x_N of the studied variable.

The x_m return level is given by

$$x_m = th + \frac{\sigma_{th}}{\xi}\left[(m\zeta_{th})^{\xi} - 1\right]$$

To go from *m* observations to *N* years, we need to account for the number of observations each year n_y as

$$m = N \times n_y$$

Applied to the previous expression, we obtain the *N*-year return level as

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(Nn_y\zeta_{th})^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(Nn_y\zeta_{th}) & for\ \xi = 0 \end{cases}$$

But how can I calculate ζ_{th}?

Poisson distribution

TU Delft

Intermezzo – Poisson distribution

The Binomial distribution is defined as

$$p_X(x) = P[X = x|n, p] = \binom{n}{x} p^x (1-p)^{n-x}$$

If $n \rightarrow \infty$, x and p are finite and defined and p is very small, $\lambda = np$.

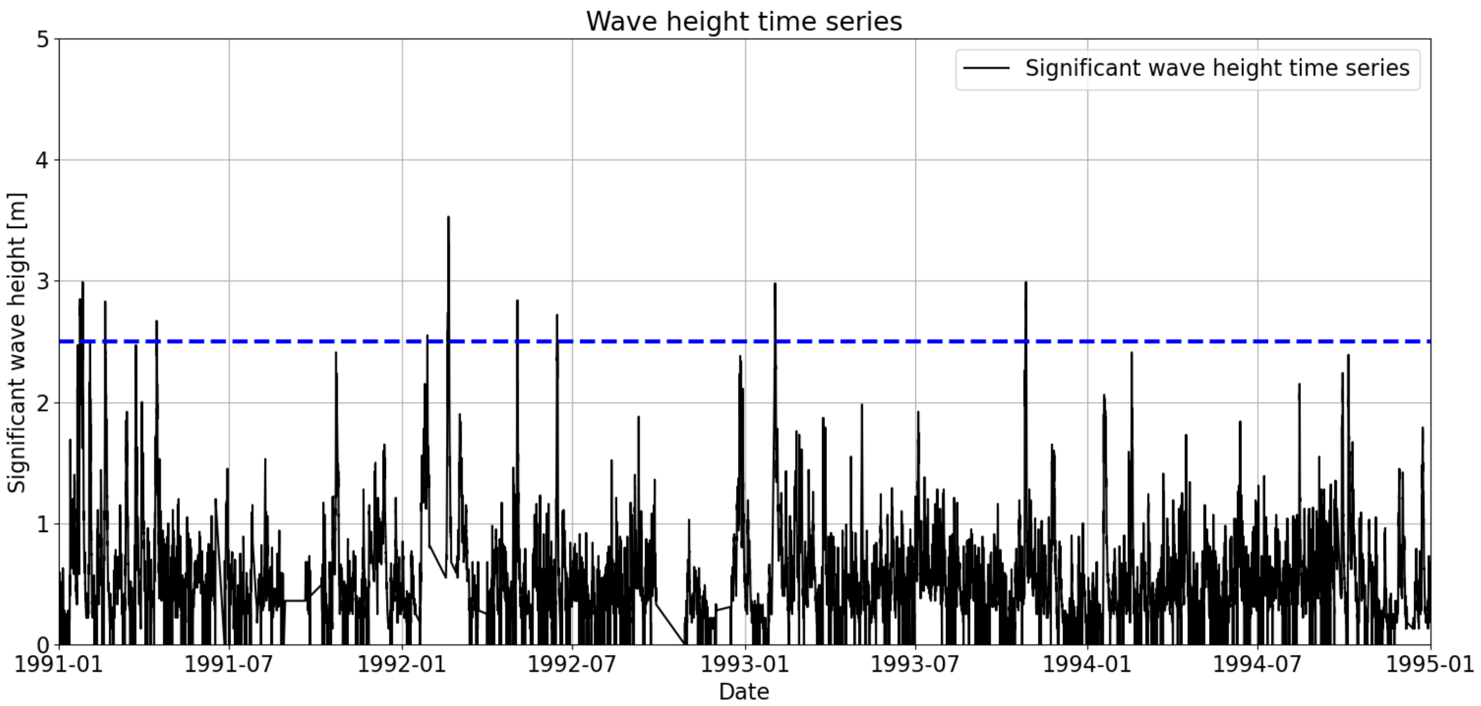
After some simplifications… **Poisson distribution**

$$p_X(x) = P[X = x|n, p] = \frac{\lambda^x e^{-\lambda}}{x!} \qquad for\ x = 0, 1, 2, \ldots\ and\ \lambda > 0$$

$$p_X(x) = P[X = x|p] = 0 \qquad otherwise$$

Binomial is based on **discrete events**, while the **Poisson** is based on **continuous events**. That is, in Poisson distribution $n \rightarrow \infty$ and p is very small, so you have an infinite number of trials with infinitesimal chance of success.

POT and Poisson



- Each hour is a trial ($n \rightarrow \infty$)
- Over or below the threshold?
- p_{above} is very small (tail of the distribution)
- Block = 1 year
- Number of excesses over the threshold ~ Poisson

Almost all the techniques to formally select the threshold and declustering time for POT are based on the assumption that the sampled extremes should follow a Poisson distribution.

TU Delft

GPD: N Return levels

The N-year return level is given by

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(N n_y \zeta_{th})^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(N n_y \zeta_{th}) & for\ \xi = 0 \end{cases}$$

Modelling the number of exceedances per year using a Poisson distribution

$$E[X] = Var[X] = \lambda \Longrightarrow \hat{\zeta}_{th} = \frac{\hat{\lambda}}{n_y}$$

where $\hat{\lambda}$ can be estimated as

$$\hat{\lambda} = \frac{n_{th}}{M}$$

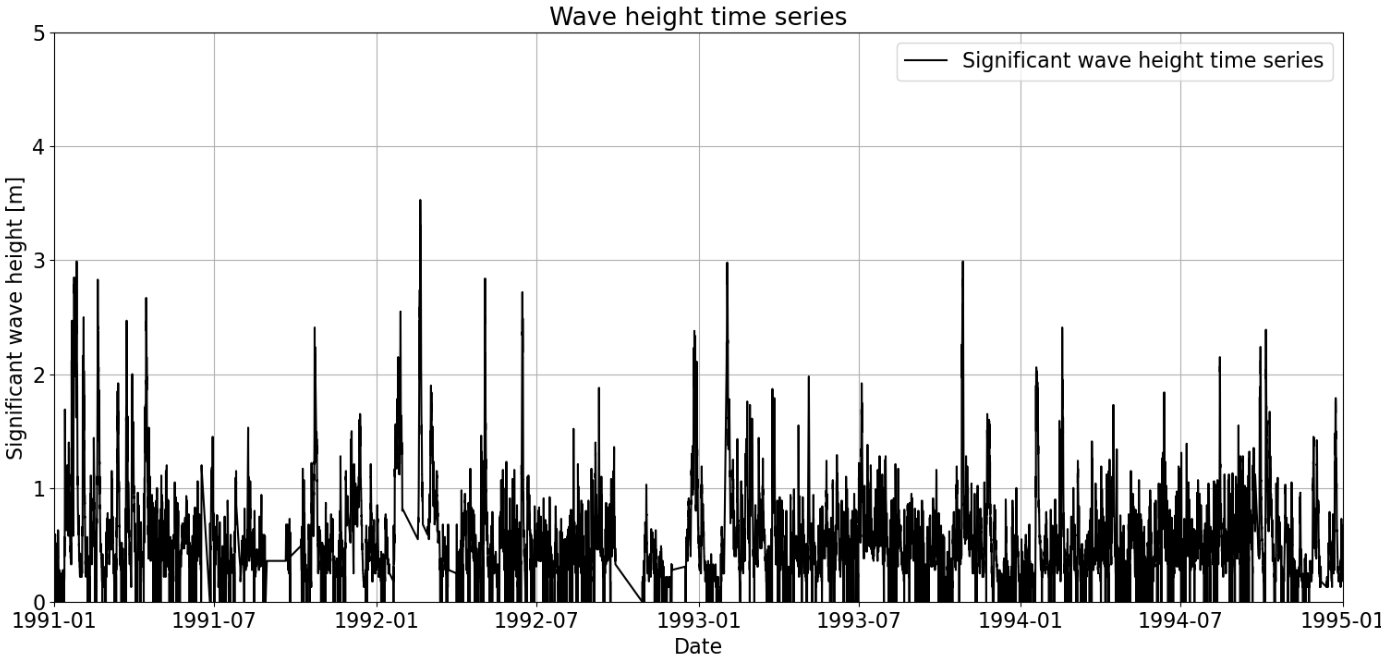
$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(\lambda N)^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(\lambda N) & for\ \xi = 0 \end{cases}$$

or

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(\frac{n_{th}}{M} N)^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(\frac{n_{th}}{M} N) & for\ \xi = 0 \end{cases}$$

TU Delft

Let's apply it



- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

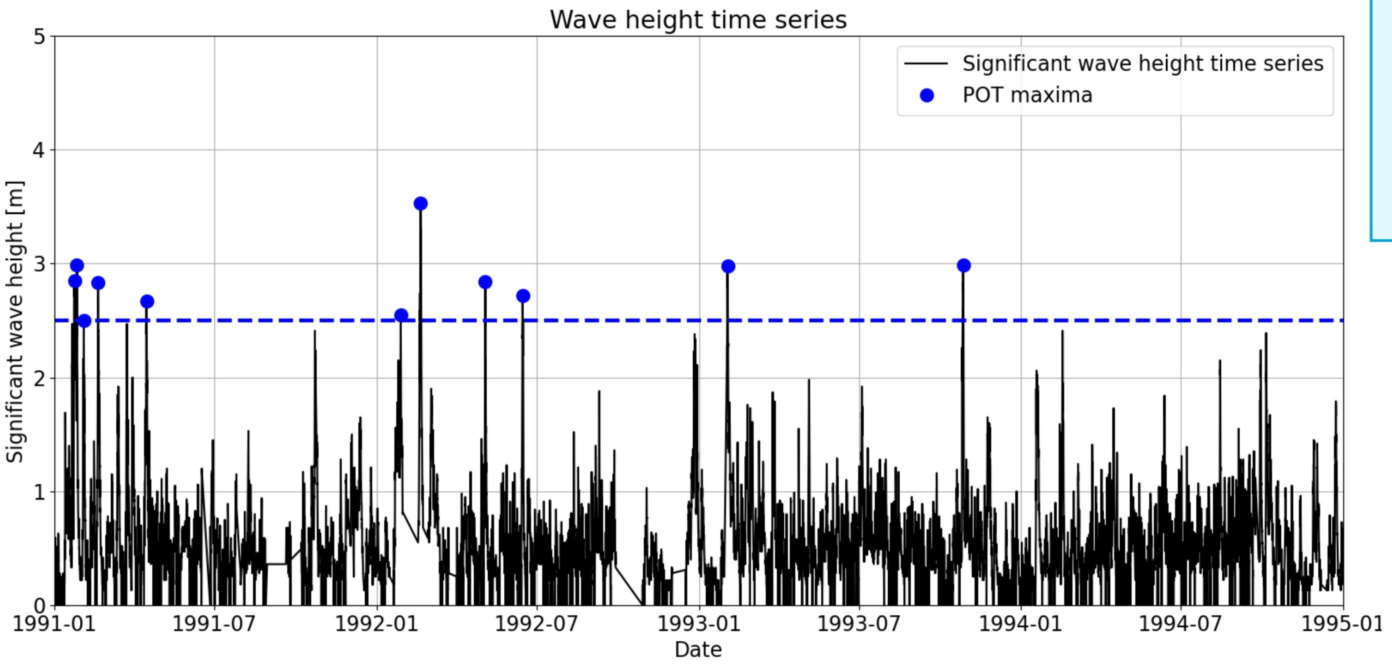
fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

inverse GPD to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

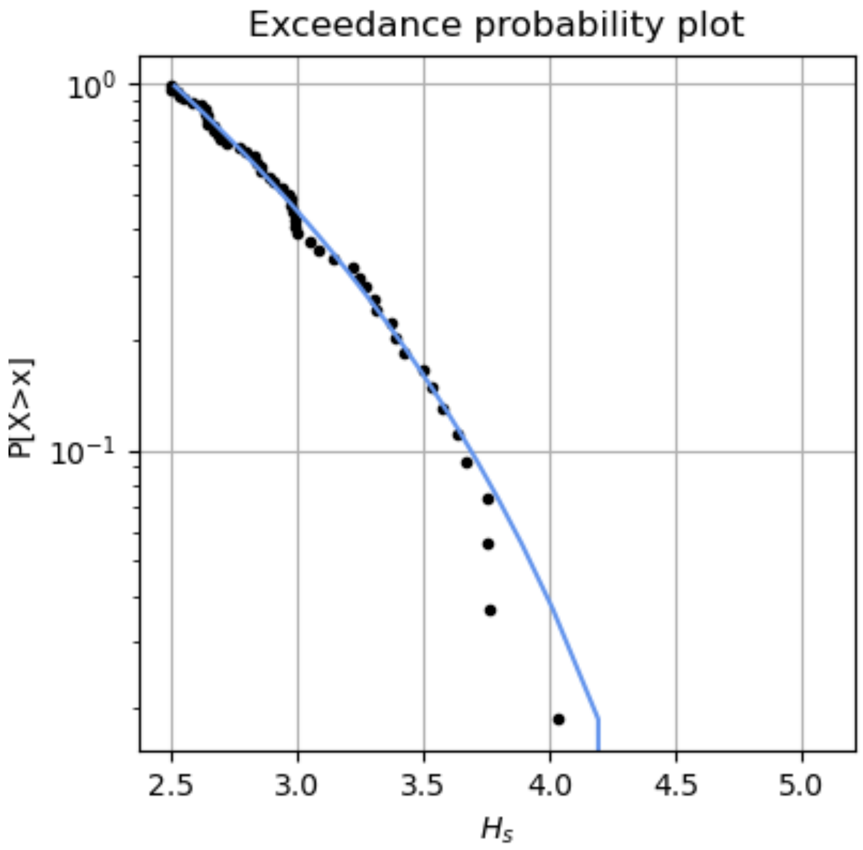
fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

inverse GPD to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

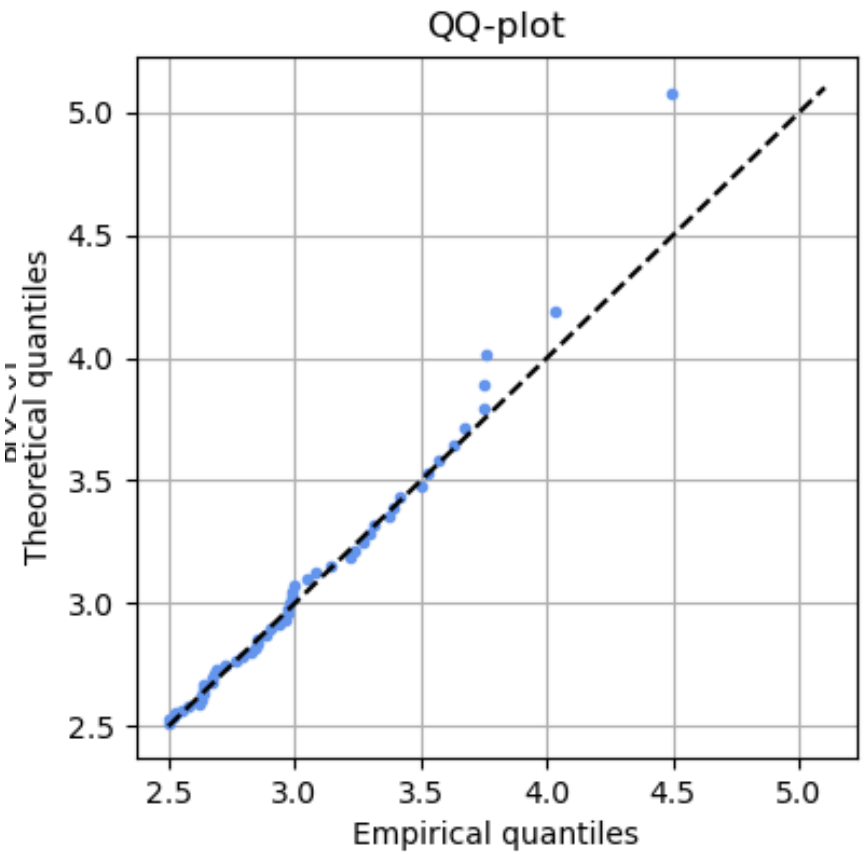
fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

inverse GPD to determine the design event

Let's apply it



- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

inverse GPD to determine the design event

Let's apply it

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(\lambda N)^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(\lambda N) & for\ \xi = 0 \end{cases}$$

T_R=90 years
M = 20 years
n_{th} = 54 events

$\Rightarrow$ $\hat{\lambda} = \frac{54}{20} = 2.7$

- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

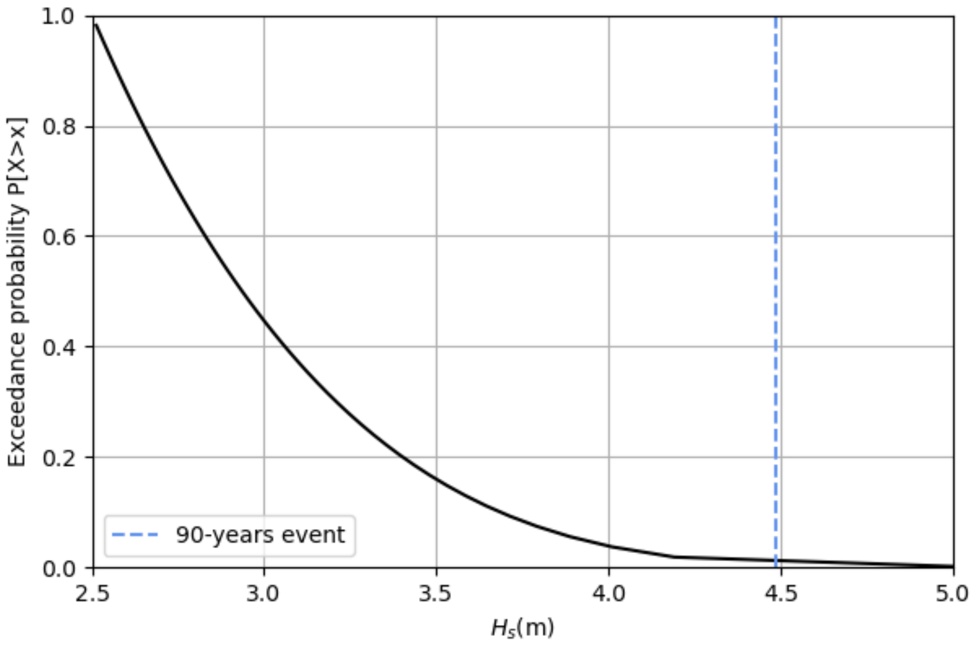
inverse GPD to determine the design event

Let's apply it

$$x_N = \begin{cases} th + \frac{\sigma_{th}}{\xi}[(\lambda N)^{\xi} - 1] & for\ \xi \neq 0 \\ th + \sigma_{th} log(\lambda N) & for\ \xi = 0 \end{cases}$$

T_R=90 years
M = 20 years
n_{th} = 54 events

$$\hat{\lambda} = \frac{54}{20} = 2.7$$



- **Load: significant wave height (T_R=90 years)**

read observations

th = 2.5
dl = 48 #in hours
excesses = find_peaks(observations, threshold = th, distance = dl) – th

fit GPD(excesses)

check fit (e.g., QQ-plot or Kolmogorov-Smirnov test)

determine lambda

inverse GPD to determine the design event

Learning objectives

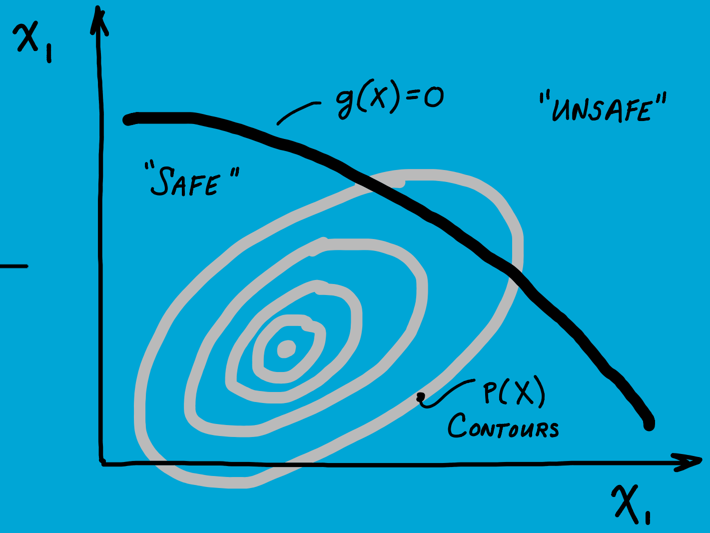
1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**
3. Apply **extreme value analysis** to datasets:
   a. Block maxima - GEV
   b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

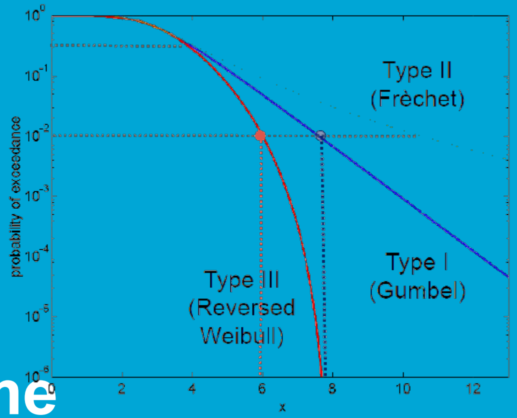
Poisson distribution

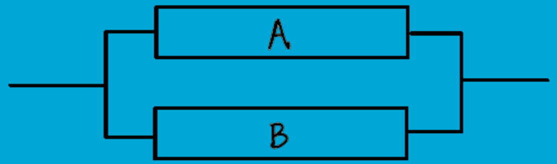
CIEM42X0 Probabilistic Design

Hydraulic and Offshore Structures (HOS) Track

Civil Engineering MSc Program



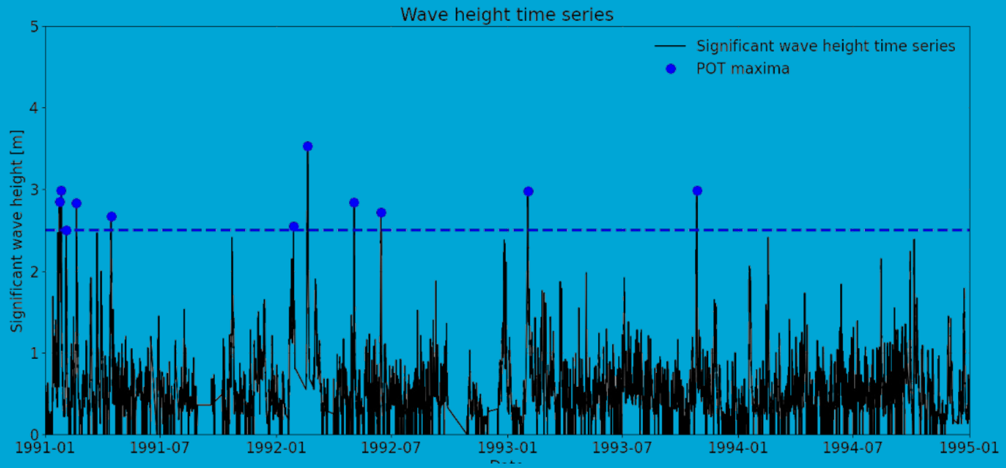




EVA: POT and GPD (III).

Threshold and declustering time selection.

Patricia Mares Nasarre

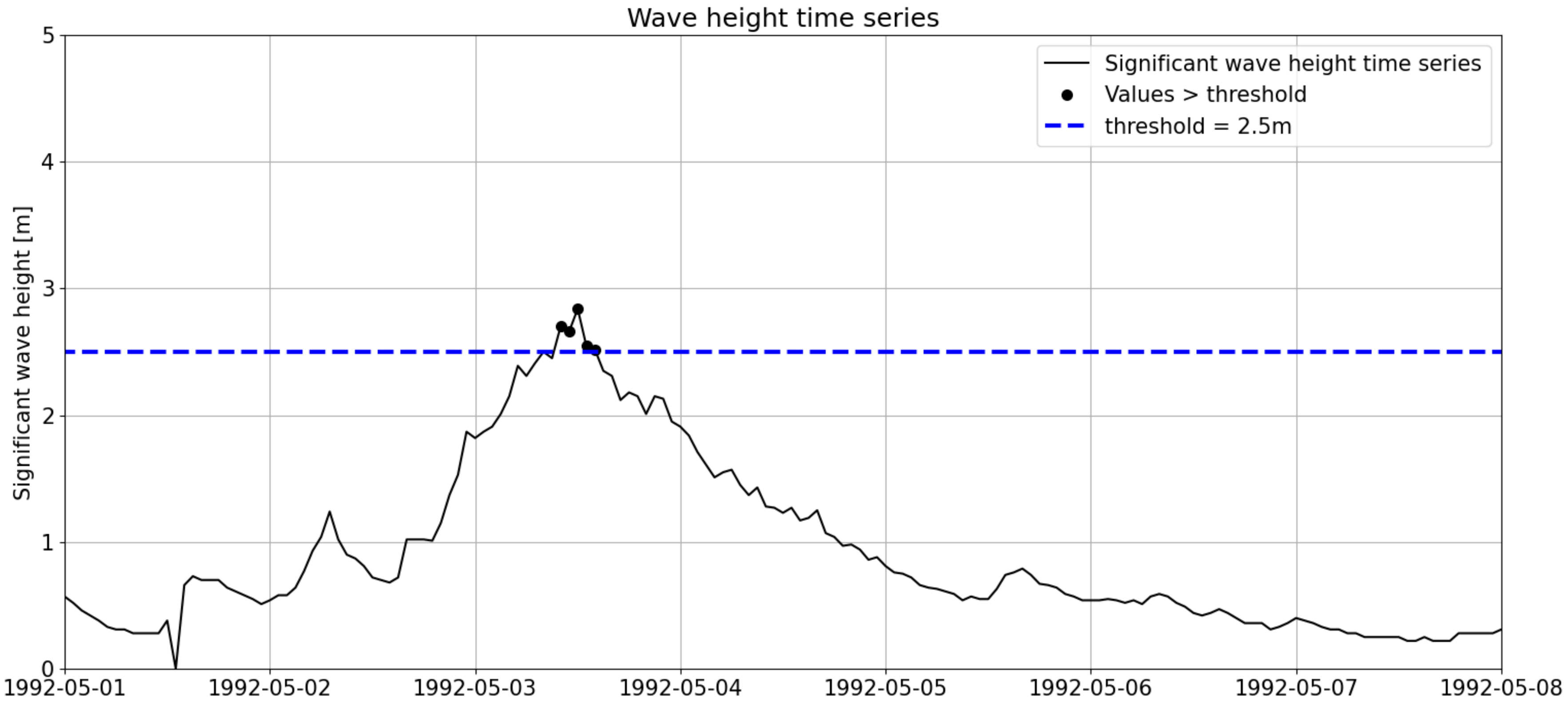


Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**
3. Apply **extreme value analysis** to datasets:
   a. Block maxima - GEV
   b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT

Choosing POT parameters

Basic assumption of EVA: extremes are *iid* ⟹ *th* and *dl* should be chosen so the identified extreme events are independent.



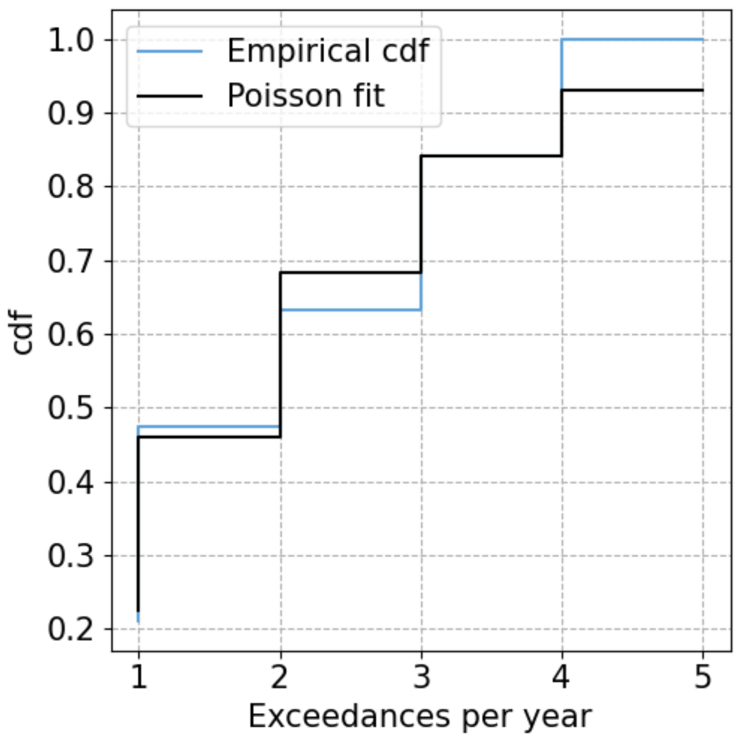
Extremes cluster in time!

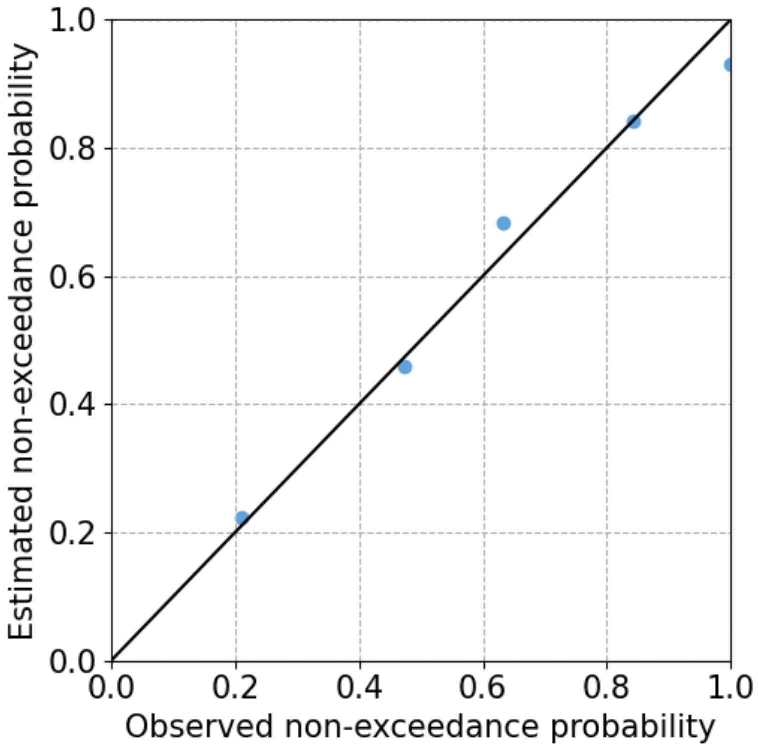
If *dl* is big enough, we ensure that extremes do not belong to the same storm.

dl → *th,* physical phenomena (local conditions)

Samples: Poisson

If the number of excesses per year follows a Poisson distribution ⟹ Sampled maxima are independent ✓





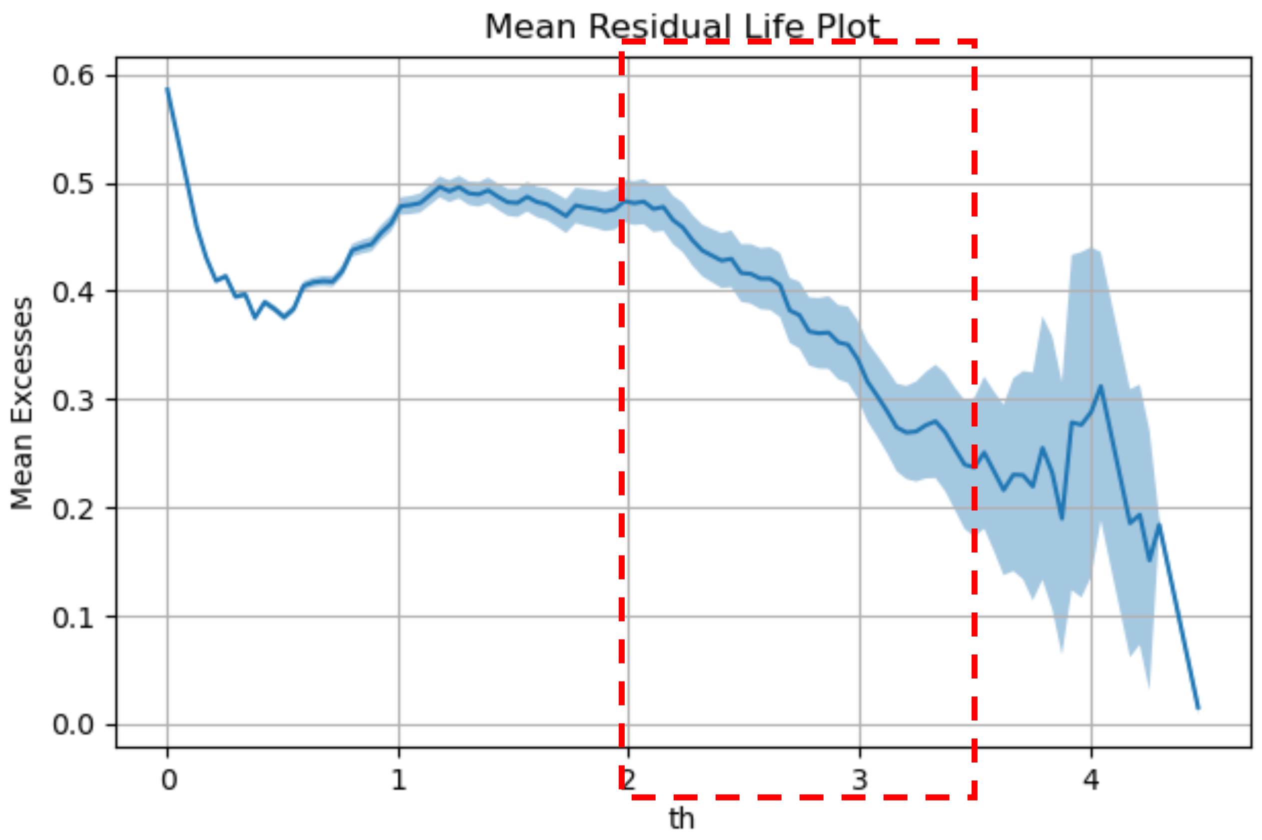
- Compute the number of excesses per year
- Empirical pmf and cdf
- Fit Poisson distribution using L-moments

$$E[X] = Var[X] = \lambda$$

- Check the fit
  - Graphically
  - Chi-squared test

Mean Residual Life (MRL) plot

MRL plot presents in the x-axis different values of *th* and, in the y-axis, the mean excess for that value of the *th*. The range of **appropriate threshold** would be that where the **mean excesses follows a linear trend**.

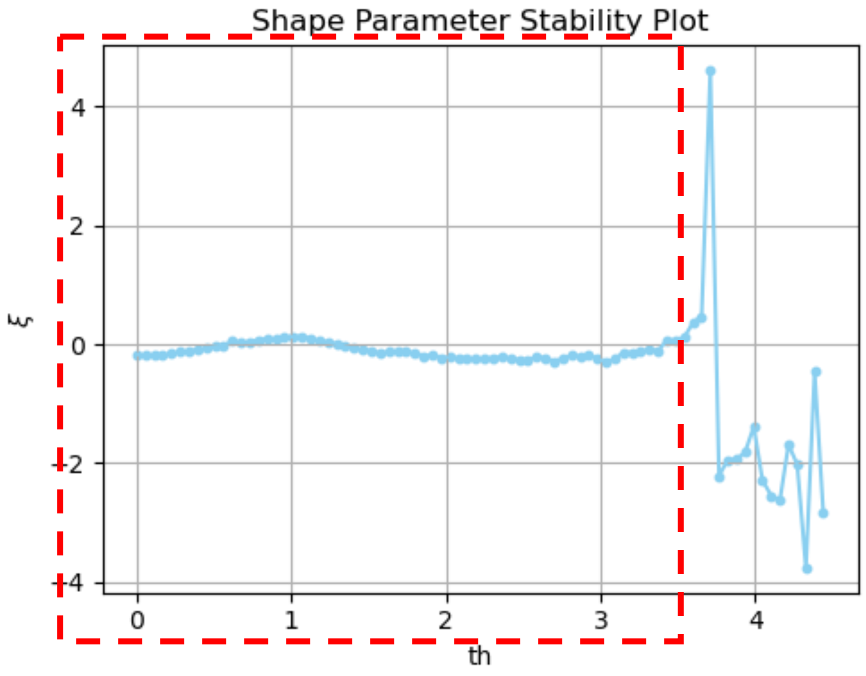


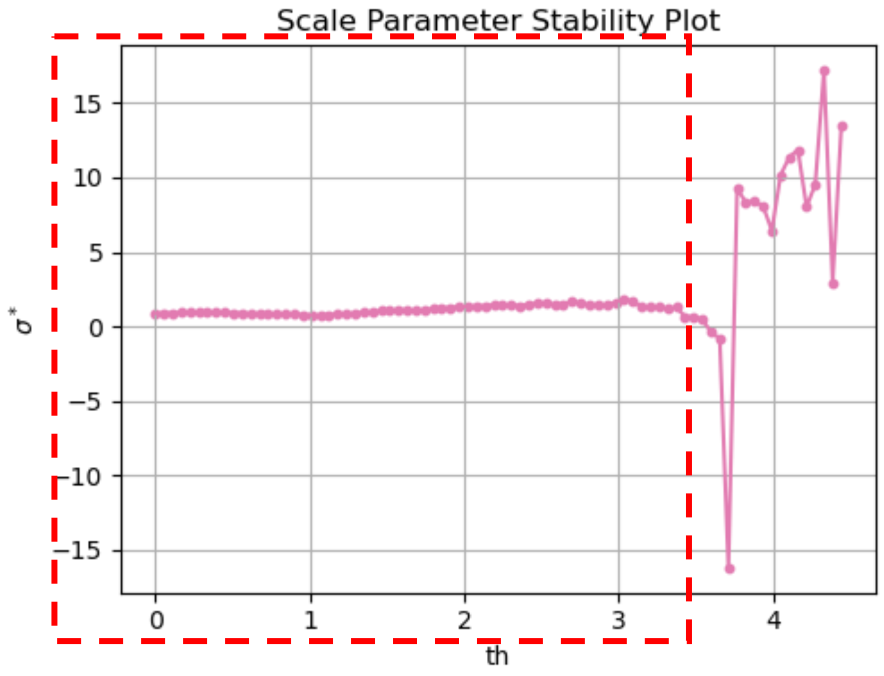
GPD parameter stability plot

GPD distribution is "threshold stable"

If the exceedances over a high threshold (*th0*) a GPD with parameters ξ and σ_{th0}, then for any other threshold (*th>th0)*, the exceedances will also follow a GPD with the same ξ and

$\sigma_{th} = \sigma_{th0} + \xi(th - th0) \Longrightarrow \sigma^* = \sigma_{th} - \xi th \Longrightarrow \sigma^* = \xi \, th0$





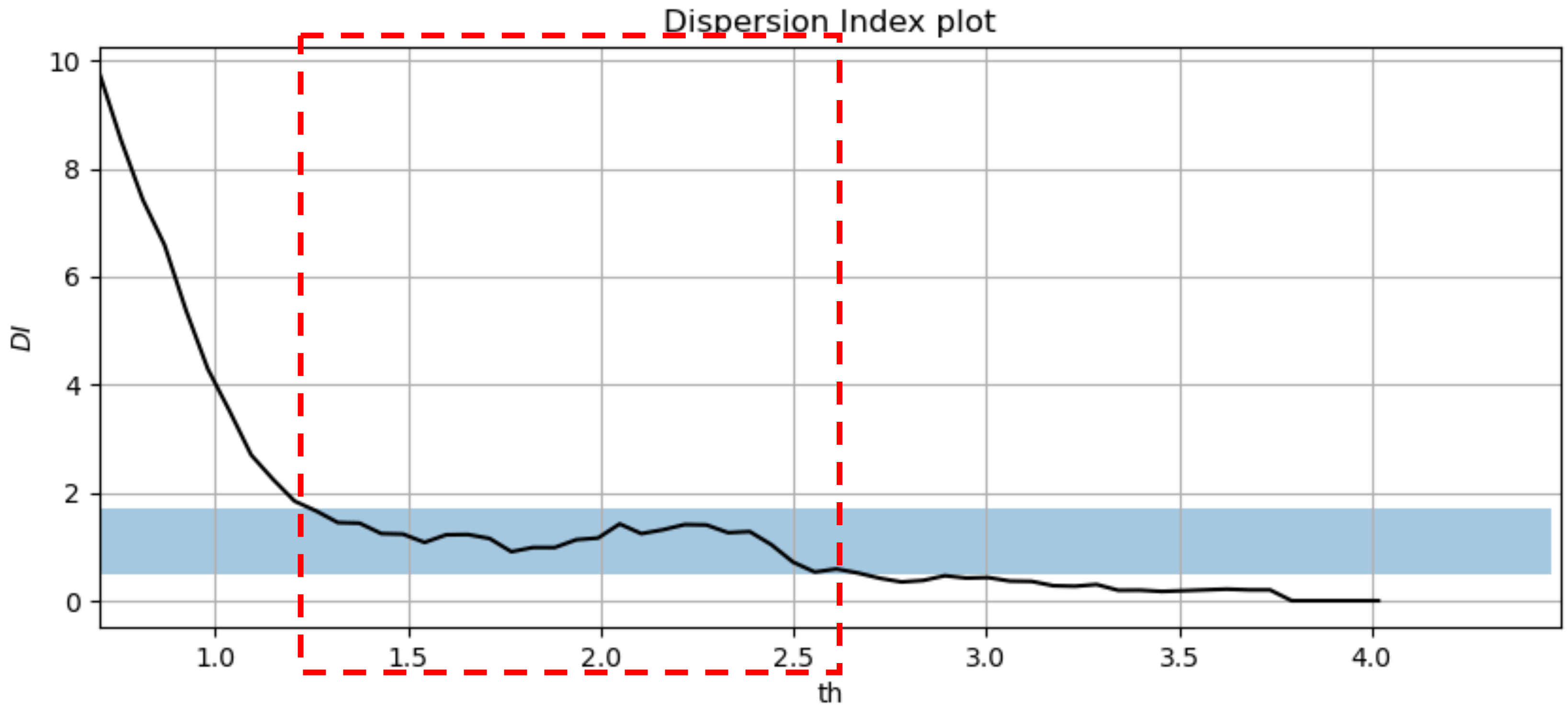
Dispersion Index (DI)

Based on Poisson process
Property of Poisson distribution: $E[X] = Var[X] = \lambda$

Dispersion Index: $DI = \frac{\sigma^2}{\mu} \approx 1$

Confidence interval for DI:

$$\left(\frac{\chi^2_{\alpha/2,M-1}}{(M/1)}, \frac{\chi^2_{1-\alpha/2,M-1}}{(M/1)} \right)$$



Learning objectives

1. Identify what is an **extreme value** and apply it within the engineering context
2. Interpret and apply the concept of **return period and design life**
3. Apply **extreme value analysis** to datasets:
   a. Block maxima - GEV
   b. Peak over threshold (POT) – GPD
4. Apply techniques to **support the threshold selection** in POT