R-LADIES LANCASTER, NOVEMBER 2023

ESTIMATING EXTREME SEA LEVELS WITH R

Eleanor D'Arcy
Lancaster University

About me

- PhD in Statistics at STOR-i Centre for Doctoral Training
- PhD in partnership with EDF UK Natural Hazards & Marine Environment R&D team
- Supervisors: Jonathan Tawn and Dafni Sifnioti
- Collaborators include Jenny Samson (Environment Agency), Jo Williams (NOC), Tom Howard (Met Office) and Ivan Haigh (University of Southampton)
- Founded diverSTOR-i network - focus on safe space for coding

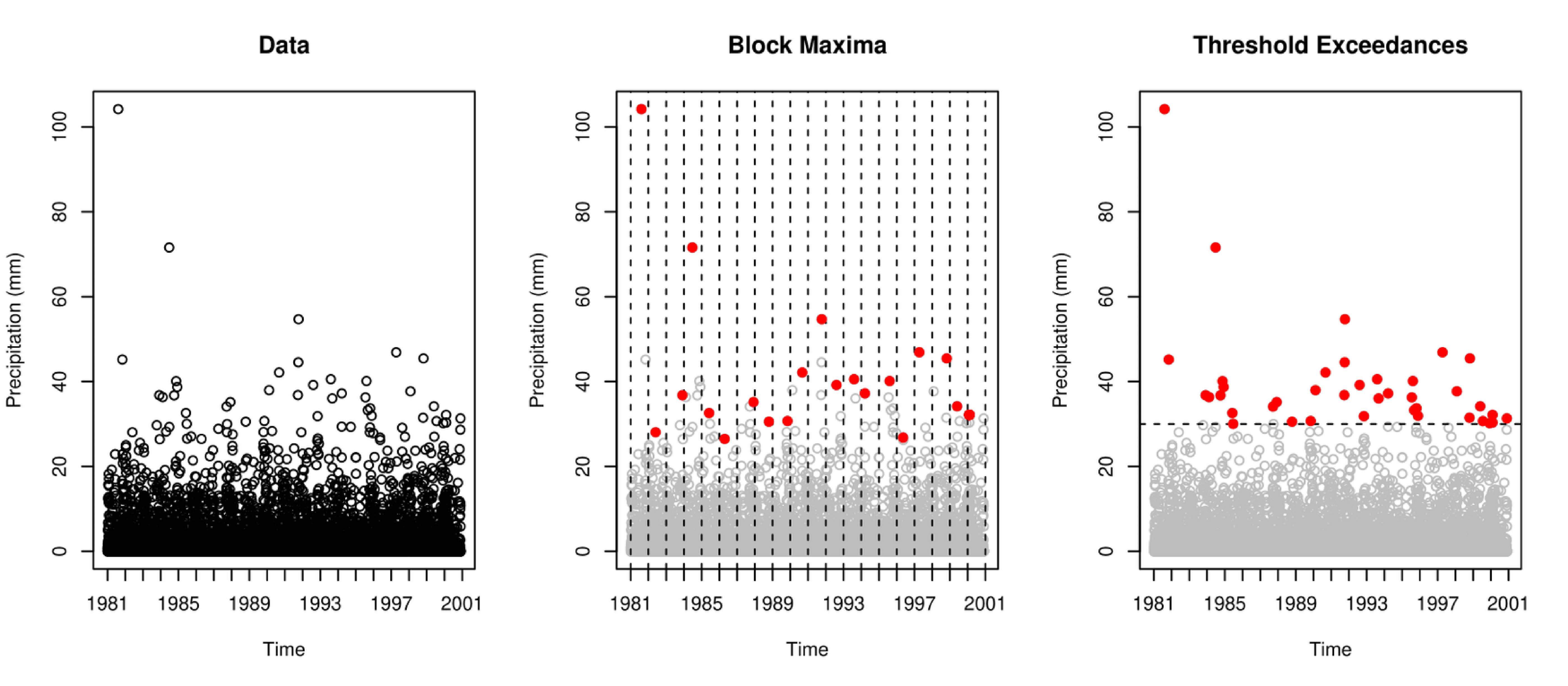
Agenda

R-LADIES LANCASTER, NOVEMBER 2023

EXTREME VALUE STATISTICS

Eleanor D’Arcy
Lancaster University

What are extremes?



Modelling extremes

1. Block maxima -> Generalised Extreme Value distribution (GEV)
2. Threshold Exceedances -> Generalised Pareto Distribution (GPD)
    1. Scale: σ
    2. Shape: ξ
    3. Rate: λ

Return levels

Level we expect to be exceeded once a year with probability p

- Correspond to return period $1/p$
- E.g. $p = 0.1$ corresponds to 10 years
- Require statistical models for extrapolation

Applications

Accounting for Climate Change in Extreme Sea Level Estimation

by Eleanor D'Arcy [1,*], Jonathan A. Tawn [1] and Dafni E. Sifnioti [2]

Modelling Extremes of Spatial Aggregates of Precipitation using Conditional Methods

Jordan Richards[A,B], Jonathan A. Tawn[A], and Simon Brown[C]

A temperature dependent extreme value analysis of UK surface ozone, 1980–2019

Lily Gouldsbrough [a], Ryan Hossaini [a b], Emma Eastoe [c], Paul J. Young [a b]

Inference for extreme spatial temperature events in a changing climate with application to Ireland

Dáire Healy[1], Jonathan Tawn[2], Peter Thorne[3], and Andrew Parnell[1]

On spatial conditional extremes for ocean storm severity

R. Shooter[1] | E. Ross[2] | J. Tawn[3] | P. Jonathan[3,4]

A marginal modelling approach for predicting wildfire extremes across the contiguous United States

Eleanor D'Arcy[1] · Callum J. R. Murphy-Barltrop[1] · Rob Shooter[2] · Emma S. Simpson[3]

INFERENCE FOR EXTREME EARTHQUAKE MAGNITUDES ACCOUNTING FOR A TIME-VARYING MEASUREMENT PROCESS

By Zak Varty[1], Jonathan A. Tawn[1,*] Peter M. Atkinson[1,†] and Stijn Bierman[2,‡]

R-LADIES LANCASTER, NOVEMBER 2023

EXTREMES IN R

Eleanor D’Arcy
Lancaster University

Packages

- evd
- extRemes
- fExtremes
- ismev
- mev
- POT

For installing R packages:

```
> install.packages('evd')
> library(evd)
```

Check out the reference manual on CRAN

Belzile, L. R., Dutang, C., Northrop, P. J., & Opitz, T. (2023). *A modeler's guide to extreme value software.* Extremes, 1-44.

CRAN task view: *https://cran.r-project.org/web/views/ExtremeValue.html*

```
> ctv::install.views("ExtremeValue", coreOnly = TRUE)
```

Data

Sea level observations at Heysham (m) recorded twice daily from 1964-2016.

Data is pubicly available on BODC:
https://www.bodc.ac.uk/data/hosted_data_systems/sea_level/uk_tide_gauge_network/

```
> load('Heysham.RData')
> head(Heysham)
```

```
             datetime sealevel
1 1964-01-01 12:00:00  9.86981
2 1964-01-02 01:00:00  9.80881
3 1964-01-02 13:00:00  9.99181
4 1964-01-03 02:00:00  9.29081
5 1964-01-03 14:00:00  9.56481
6 1964-01-04 02:00:00  8.83381
```

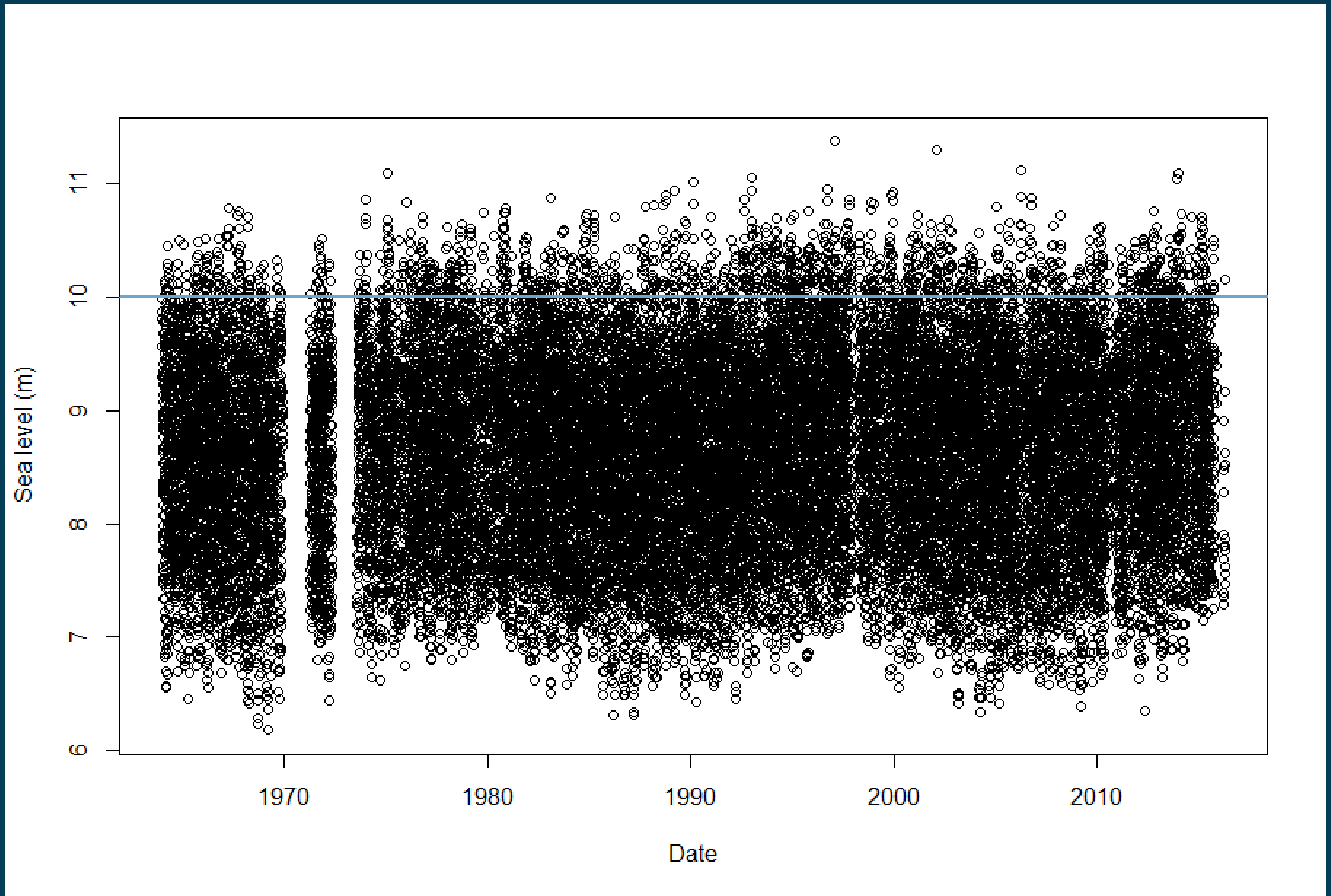
```
> dim(Heysham)
[1] 30827     2
```

```
> summary(Heysham)
```

```
    datetime                          sealevel
 Min.   :1964-01-01 12:00:00   Min.   : 6.174
 1st Qu.:1979-05-08 15:00:00   1st Qu.: 8.009
 Median :1991-04-03 14:00:00   Median : 8.680
 Mean   :1990-10-31 12:40:52   Mean   : 8.650
 3rd Qu.:2003-03-23 20:30:00   3rd Qu.: 9.280
 Max.   :2016-04-16 19:00:00   Max.   :11.374
```

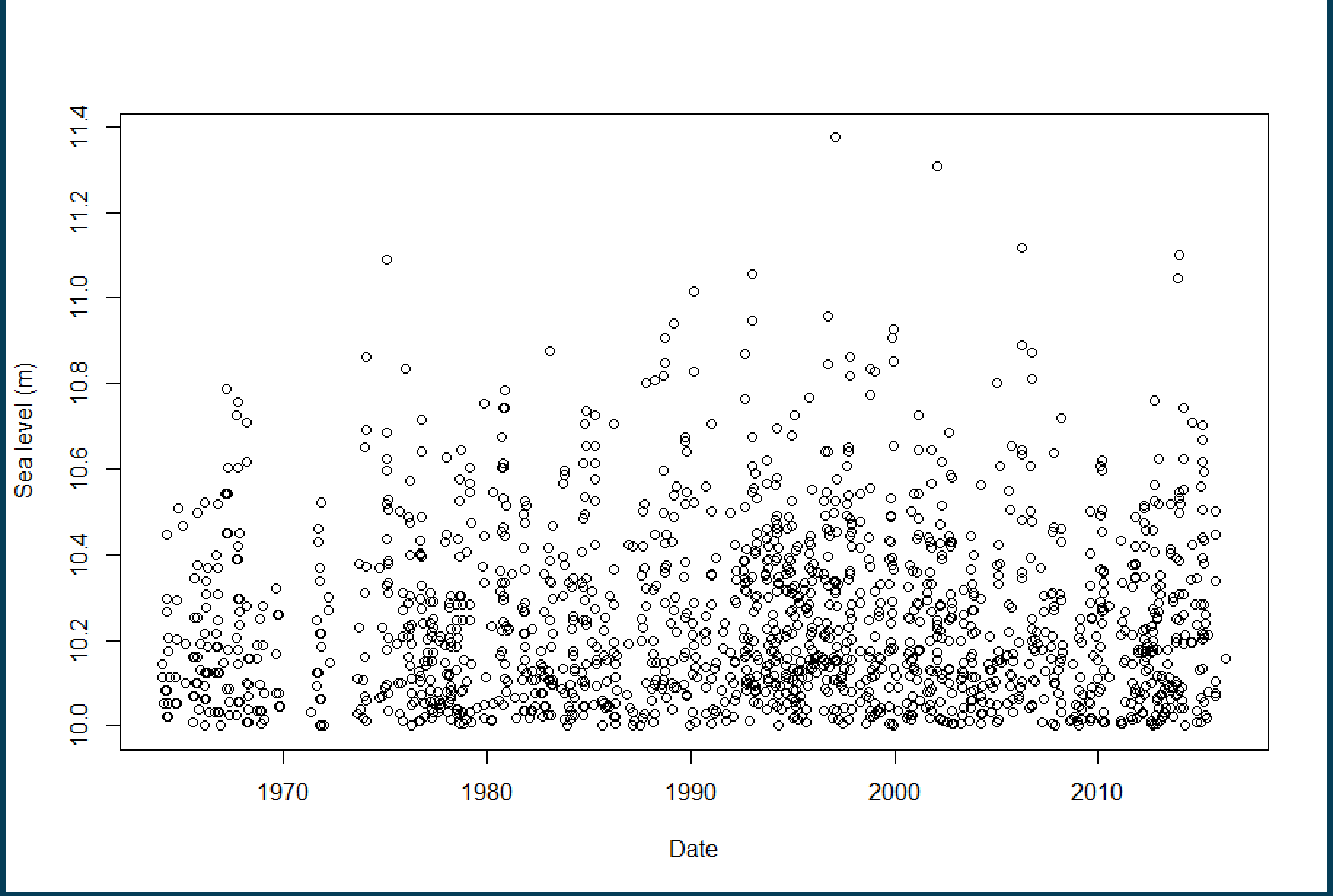
Data visualisation

```
> plot(Heysham$datetime, Heysham$sealevel, xlab='Date', ylab='Sea level (m)')
> abline(h=10)
```



Extreme observations

```
> plot(Heysham$datetime[which(Heysham$sealevel>10)],
       Heysham$sealevel[which(Heysham$sealevel>10)],
       xlab='Date', ylab='Sea level (m)')
```



Model fitting

```
> library(evd)
> model <- gpd.fit(Heysham$sealevel, 10)
```

```
$threshold
[1] 10

$nexc
[1] 1590

$conv
[1] 0

$nllh
[1] -663.1941

$mle
[1]  0.2895677 -0.1777515

$rate
[1] 0.05157816

$se
[1] 0.008554979 0.016246434
```

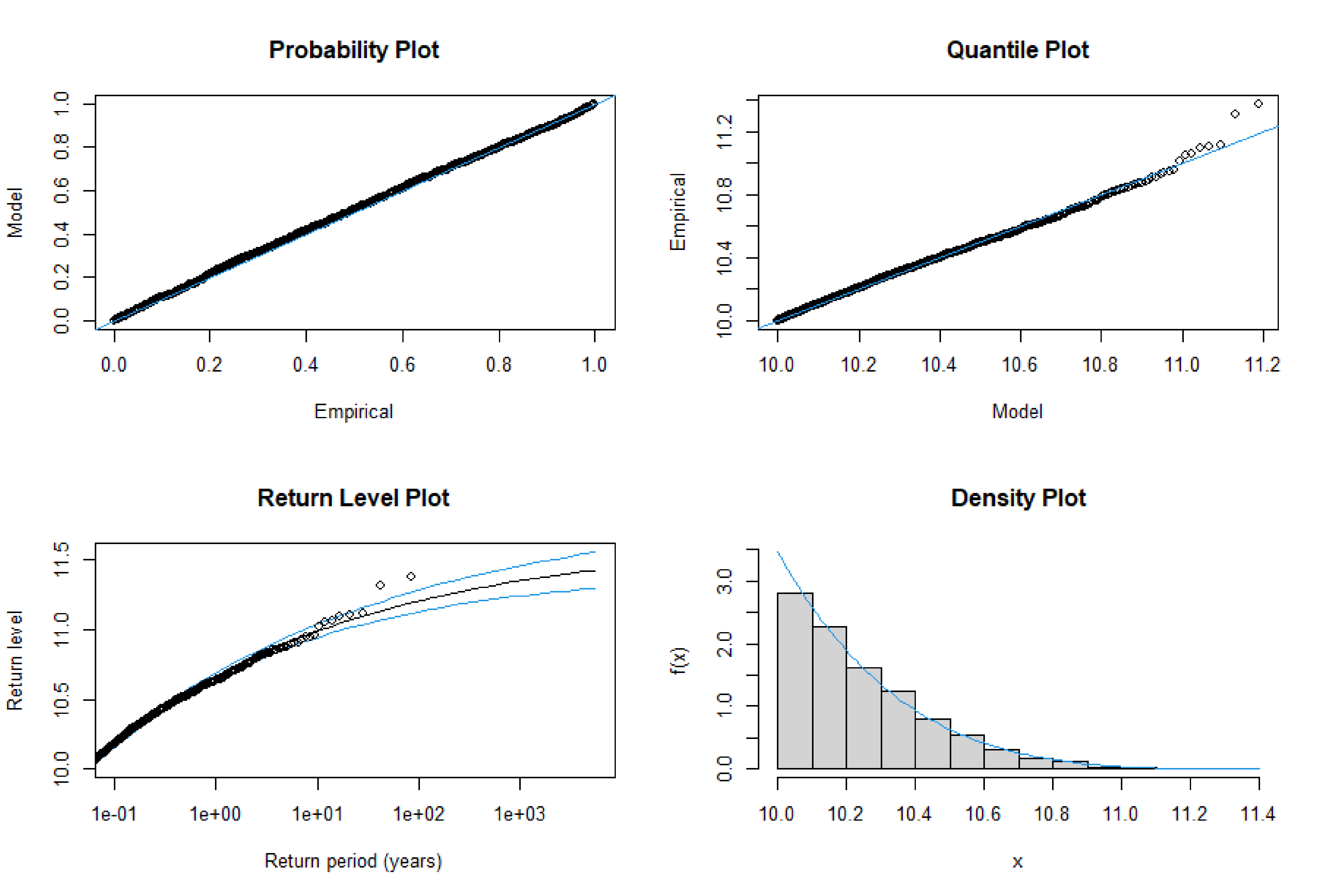
Number of exceedances

Parameter estimates

Standard error

Model diagnostics

```
> gpd.diag(model)
```



Model fitting - extRemes

```r
> library(extRemes)
> model <- fevd(Heysham$sealevel, threshold = 10, type='GP')
> model
```

```
fevd(x = Heysham$sealevel, threshold = 10, type = "GP")

[1] "Estimation Method used: MLE"


 Negative Log-Likelihood Value:  -663.1941


 Estimated parameters:
     scale      shape
 0.2895837 -0.1777924

 Standard Error Estimates:
      scale       shape
0.008554627 0.016238489
```

Return levels

```
> return.level(model, do.ci=TRUE)
fevd(x = Heysham$sealevel, threshold = 10, type = "GP")

[1] "Normal Approx."

                       95% lower CI Estimate 95% upper CI
2-year return level        10.73380 10.77441     10.81501
20-year return level       10.99302 11.06142     11.12983
100-year return level      11.10902 11.20261     11.29619
```

```
> return.level(model, return.period = c(1000, 10000), do.ci=TRUE)
fevd(x = Heysham$sealevel, threshold = 10, type = "GP")

[1] "Normal Approx."

                         95% lower CI Estimate 95% upper CI
1000-year return level       11.21151 11.34577     11.48004
10000-year return level      11.25973 11.44084     11.62196
```

R-LADIES LANCASTER, NOVEMBER 2023

EXTREME SEA LEVEL ESTIMATION

Eleanor D’Arcy
Lancaster University

Motivation

Consequences of coastal flooding:

- Loss of life
- Damage to property and infrastructure
- Coastal erosion
- Displacement of people
- Loss of habitats and ecosystems

The Office for Nuclear Regulation are interested in 10,000 year return level estimates

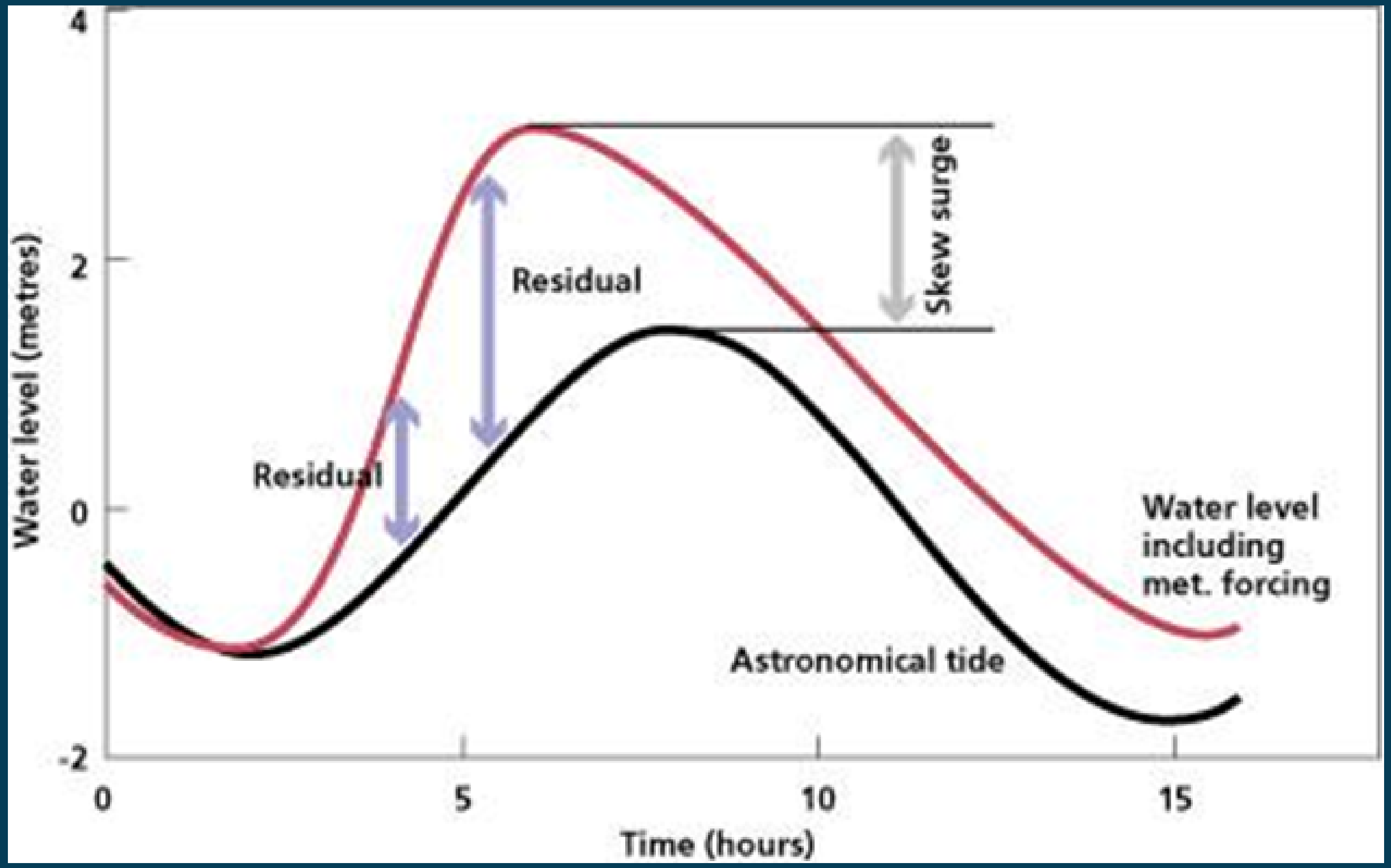
Sea level components

Tides

- Predictable rise and fall of the sea surface
- Can be well predicted far in advance

Storm surge

- Short term sea level changes caused by the weather
- Not accurately predictable



Sea Level = Mean Sea Level + Skew surge + Peak Tide

Current methodology

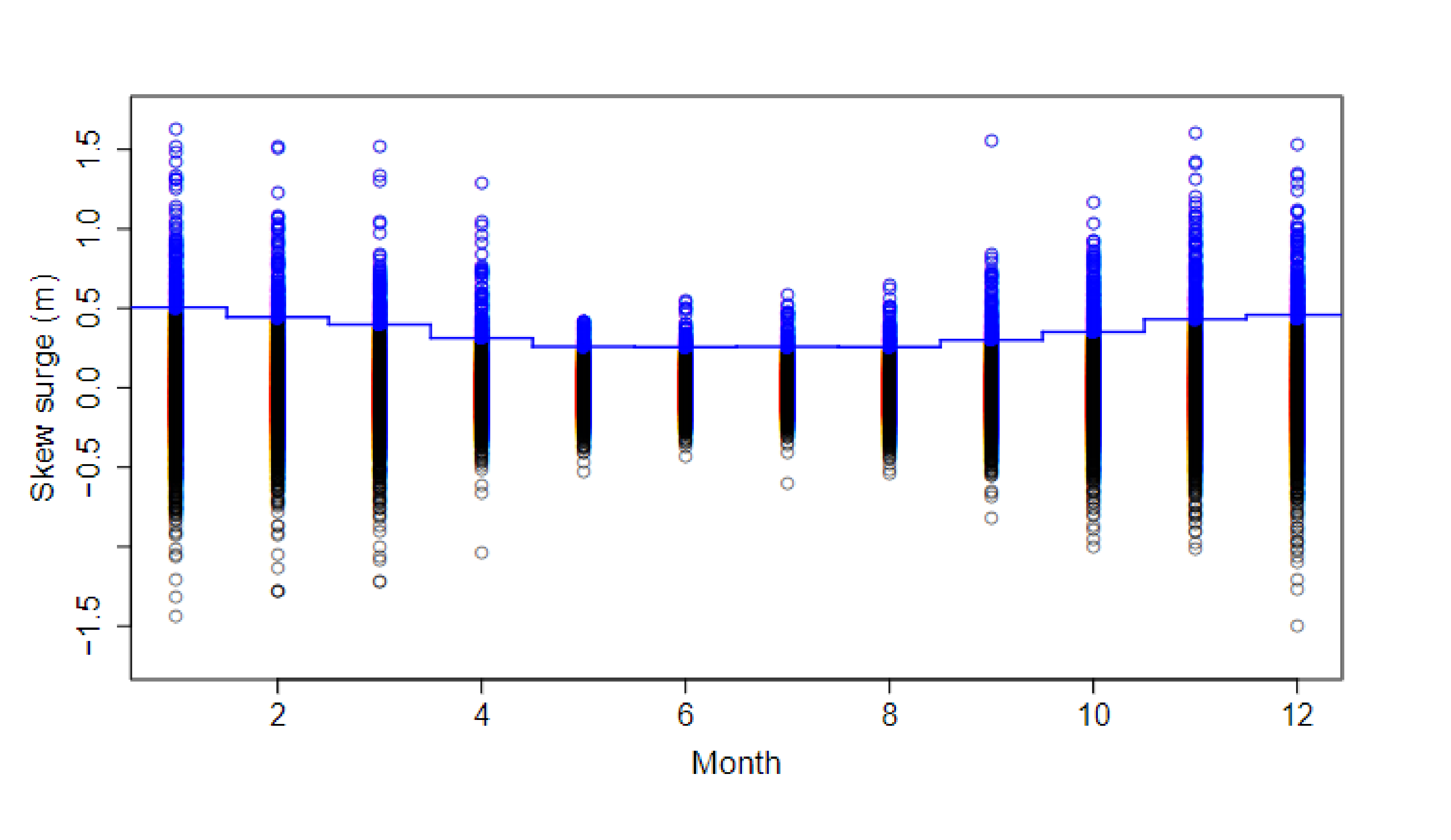
- Skew surges are independent and identically distributed
  - Model extremes using a GPD
- Tides are stationary
- Skew surge and peak tide are independent of each other

Batstone, C. et. al. (2013) *A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines.* Ocean Engineering, 71, pp.28-39.

Aim: Correct these and estimate extreme sea levels

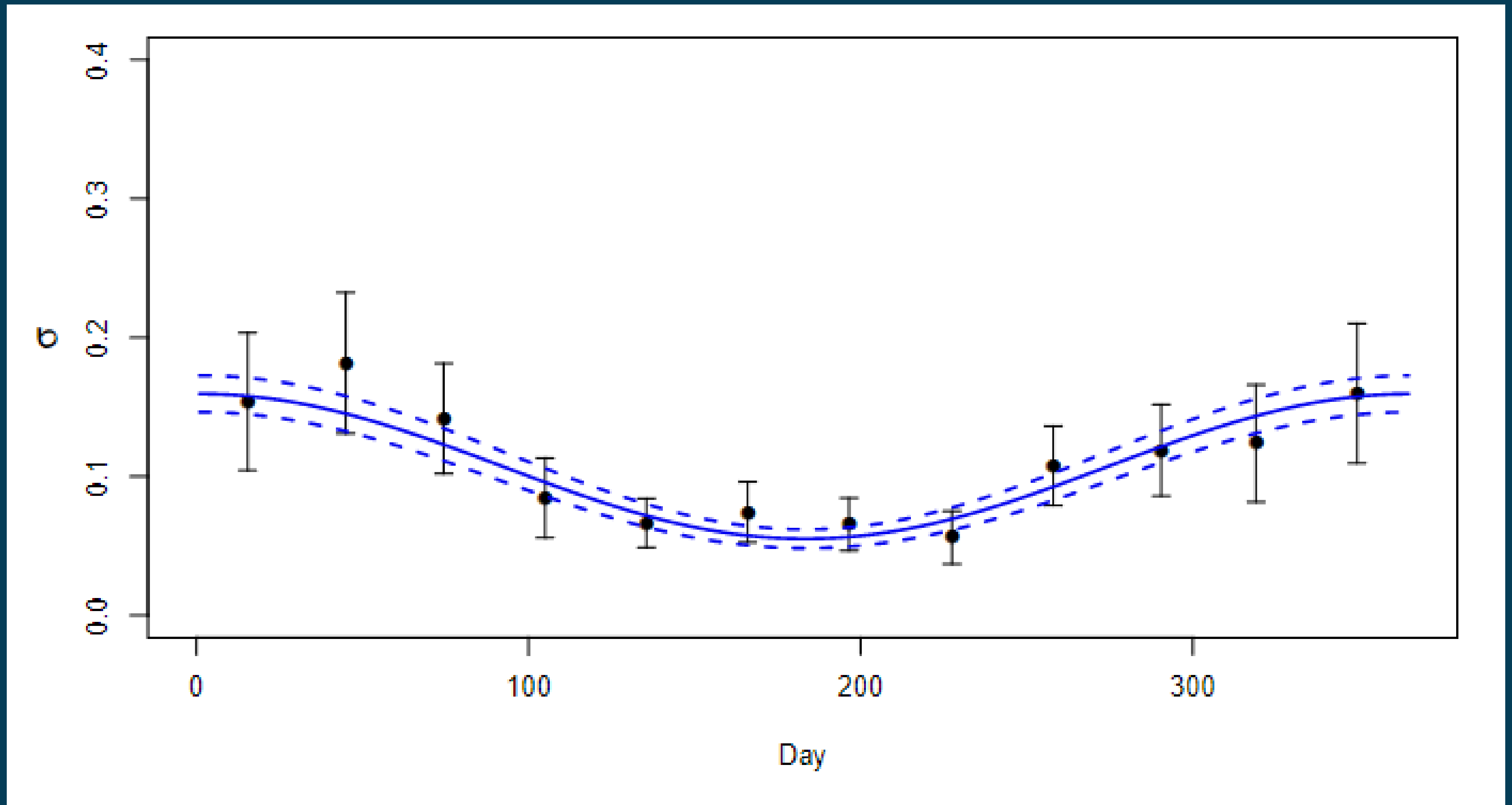
1: Accounting for Seasonality

Skew surge seasonality



- Monthly threshold
- For non-extremes, we use the monthly empirical distribution
- For extremes, we use the non-stationary GPD

Skew surge modelling



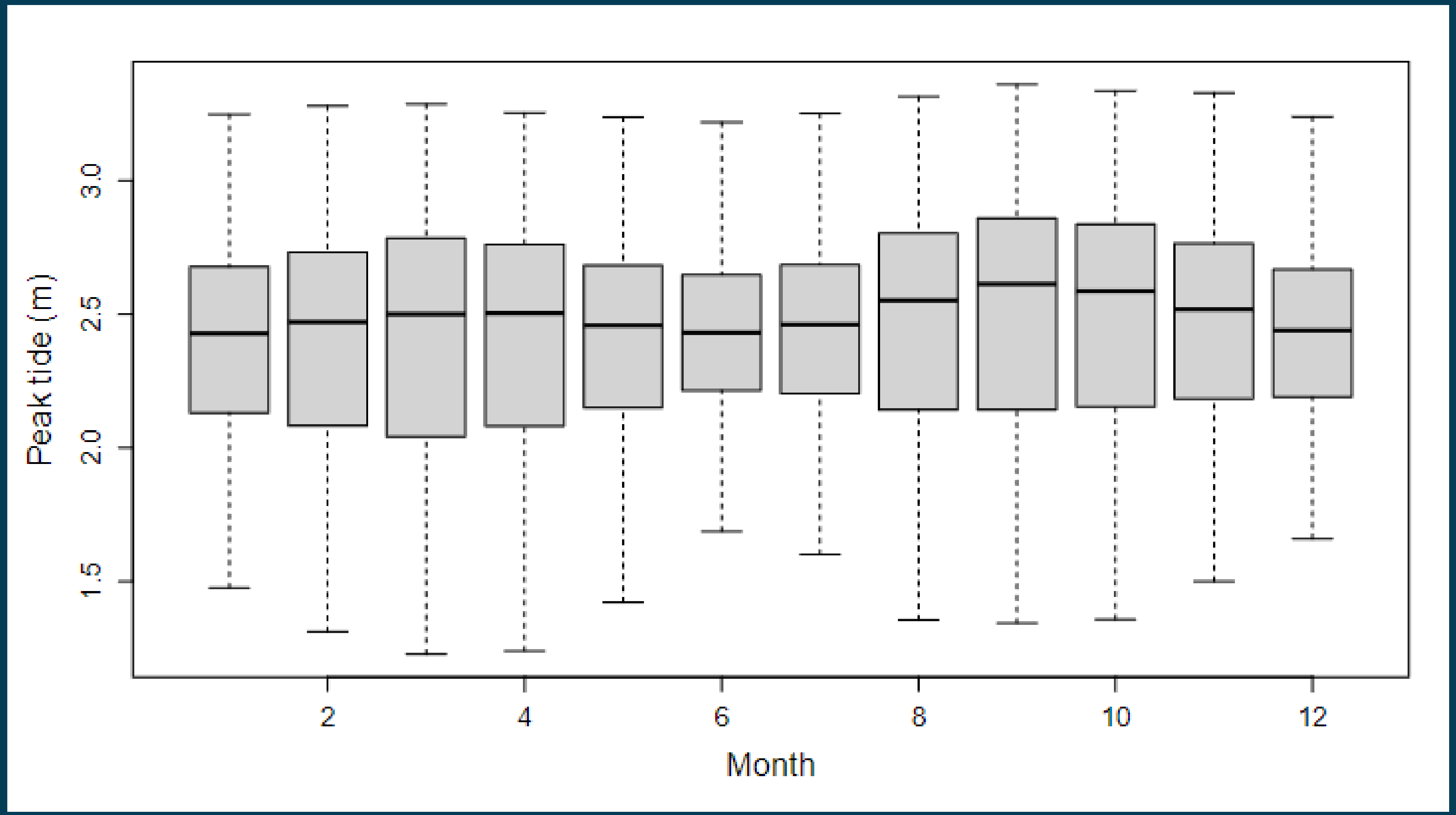
$$Y \sim GPD(\sigma_d, \xi, \lambda_d)$$

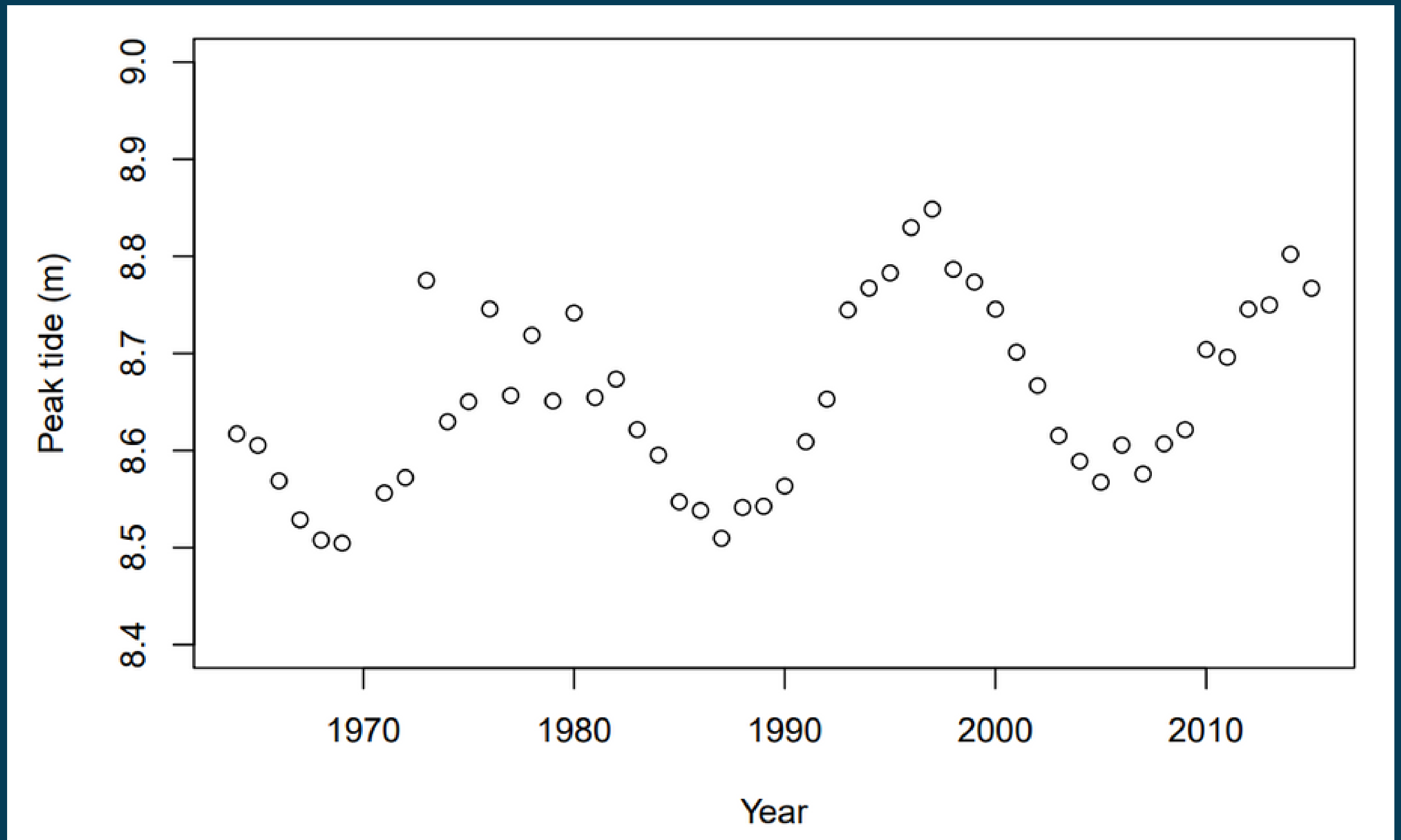
daily covariate

$$\sigma_d = \alpha_\sigma + \beta_\sigma \sin\left(\frac{2\pi}{365}(d - \phi_\sigma)\right)$$

parameters

Peak tide seasonality





Note: Tides are deterministic so don’t require modelling

2: Skew surge-peak tide dependence

SS-PT dependence

- Can be assumed as independent at most sites
- Extreme skew surges tend to occur on lower peak tides
- For extreme skew surge,

$$Y \sim GPD(\sigma_{d,t},\ \xi,\lambda_{d,t})$$

- Relationship changes throughout the year

SS-PT dependence

tide

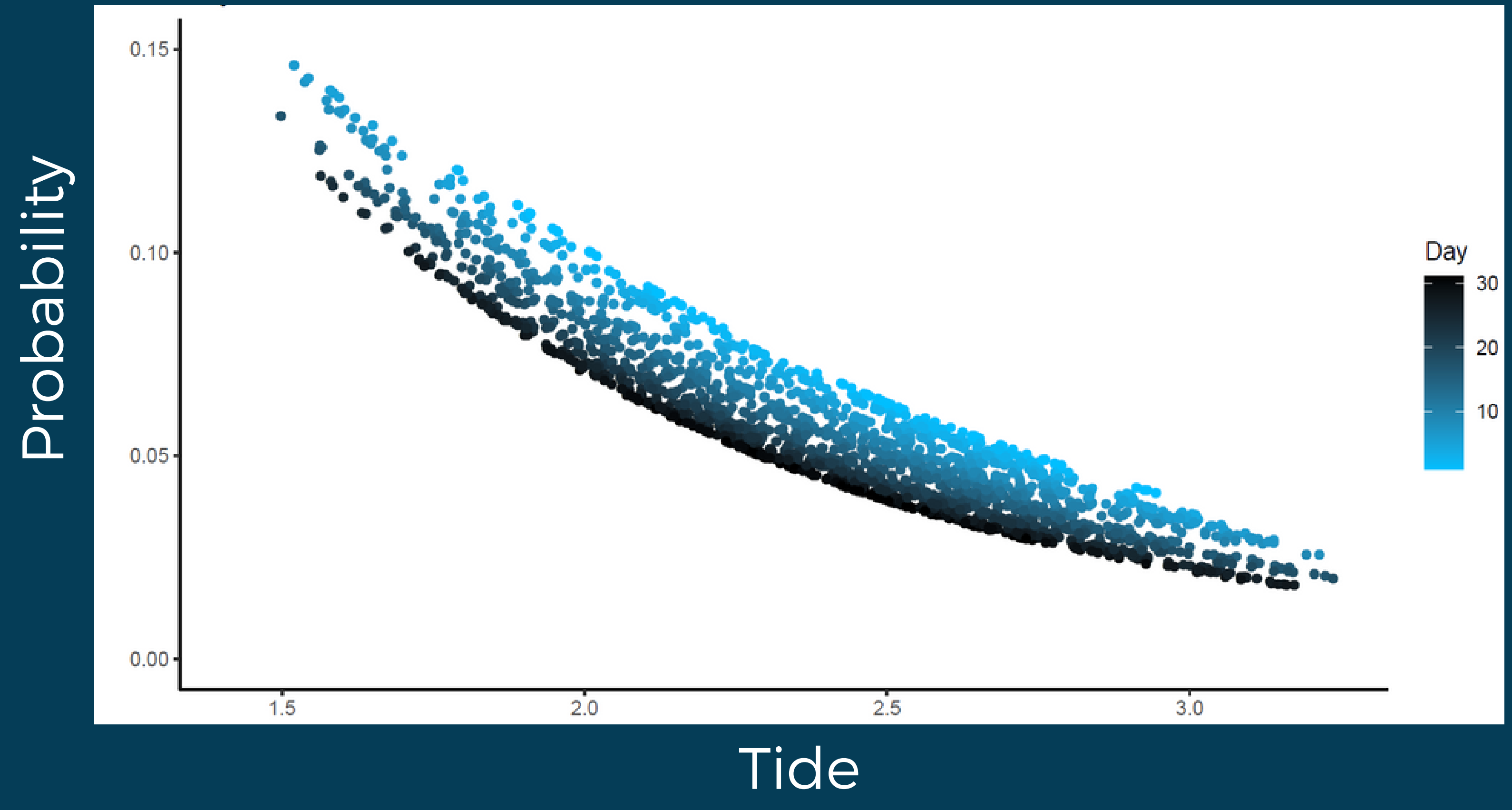
parameters

link function

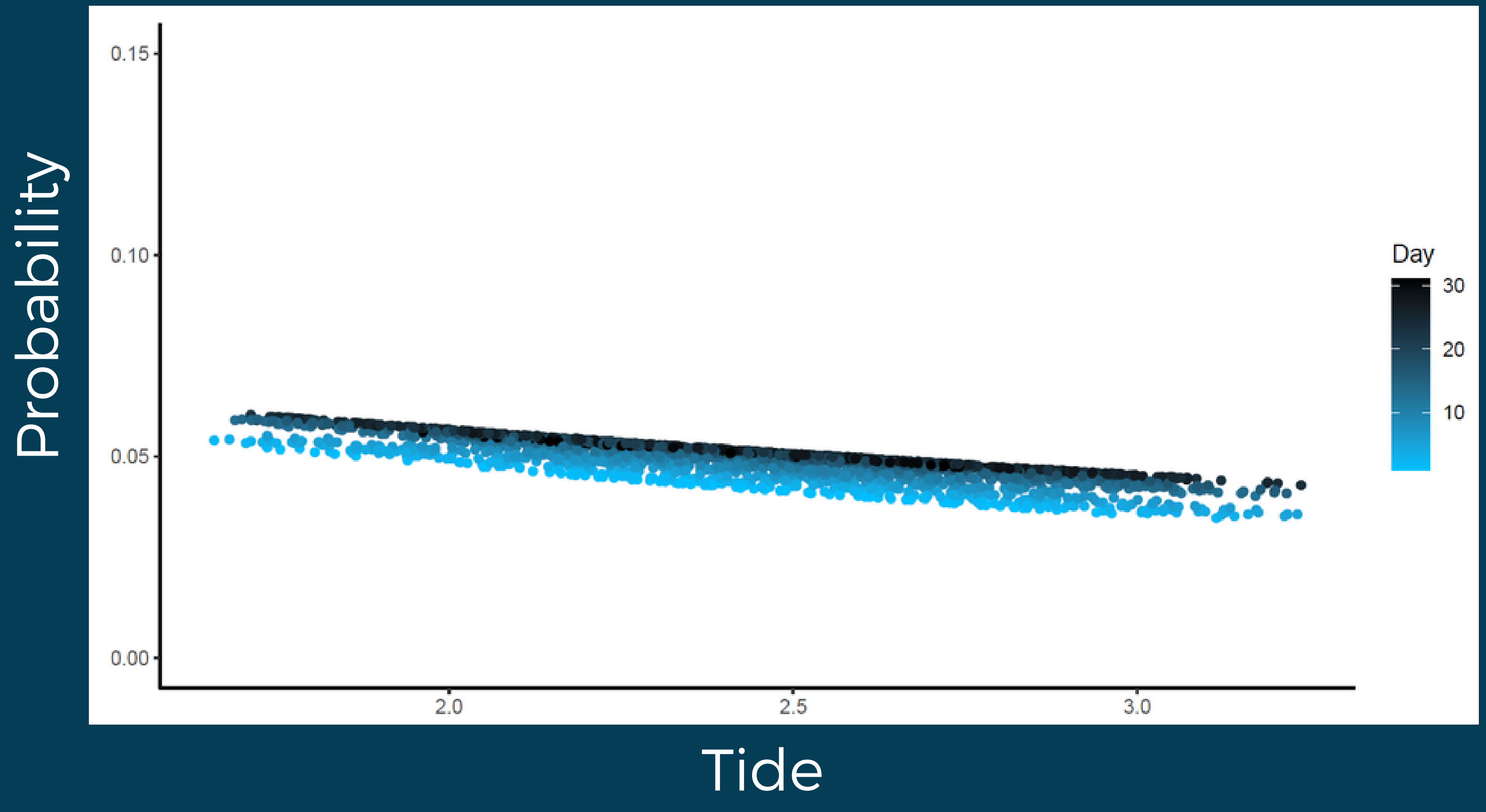
$$g(\lambda_{d,x}) = g(\lambda_d)+\tilde{x}\left[\alpha_\lambda+\beta_\lambda \sin\left(\frac{2\pi}{365}(d-\phi_\lambda)\right)\right]$$

day in year

May



December

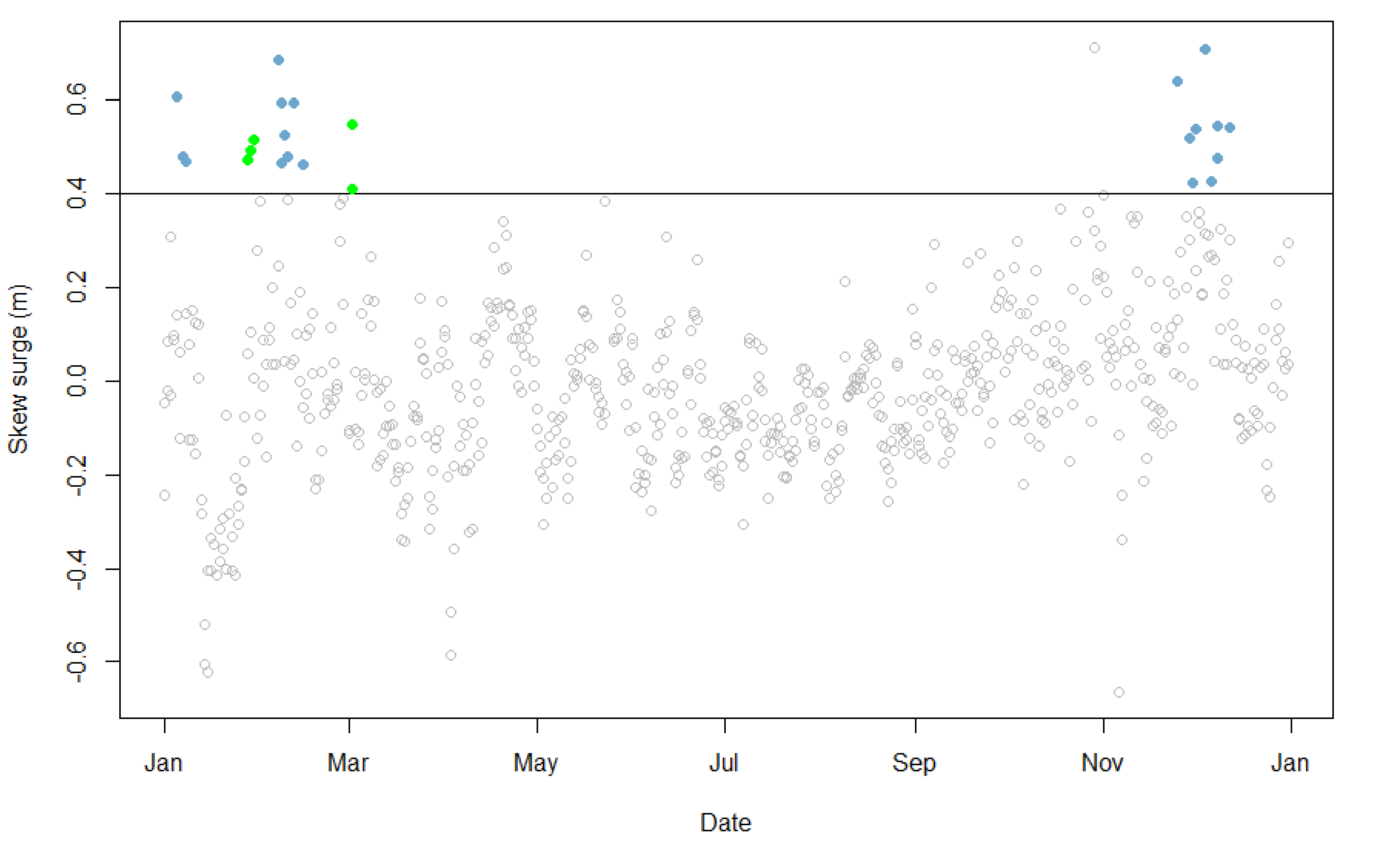


3: Skew surge temporal dependence

Clusters of extremes

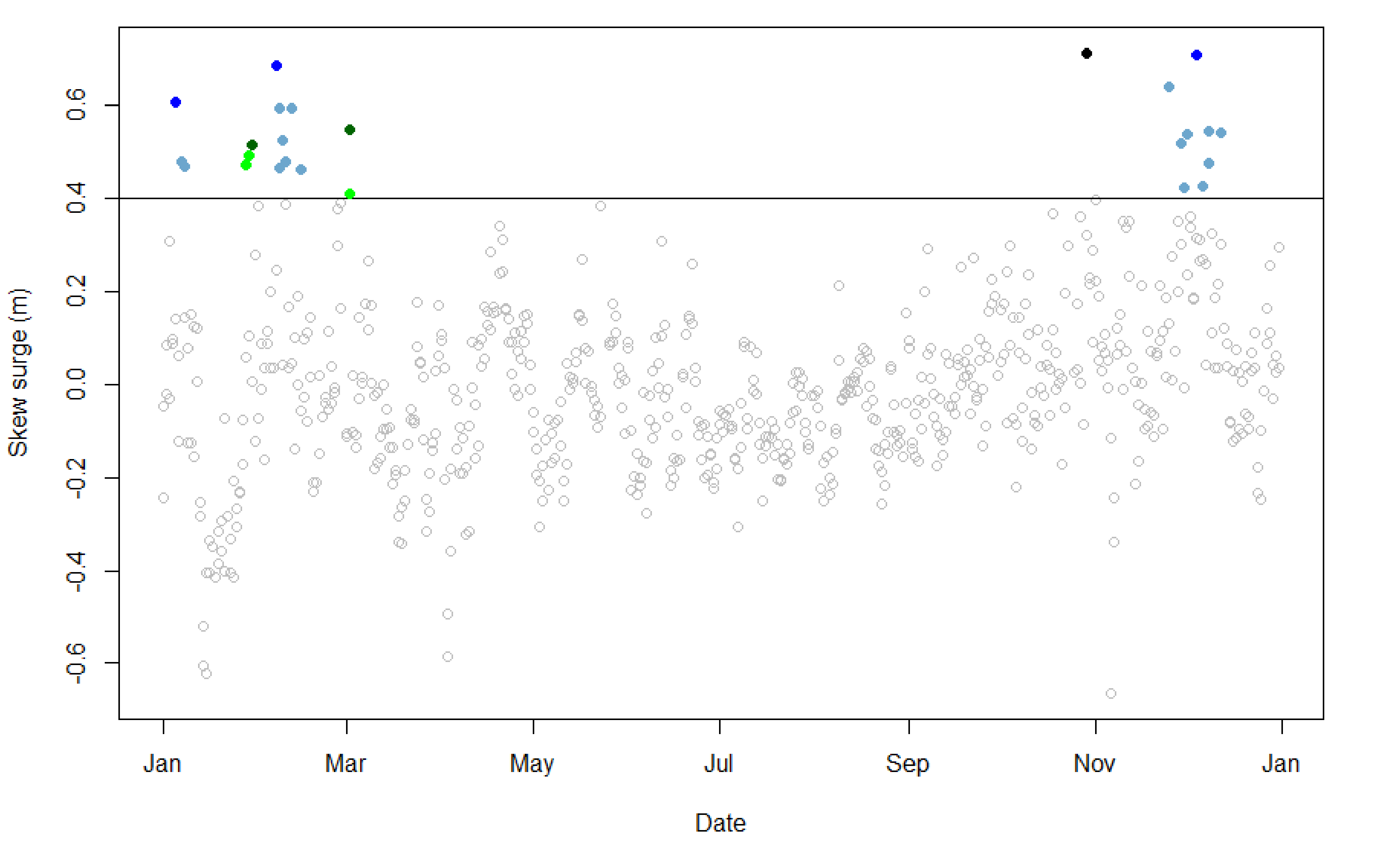
Clusters of extreme skew surge occur when a storm lasts multiple days.

Aim: Identify individual storm events



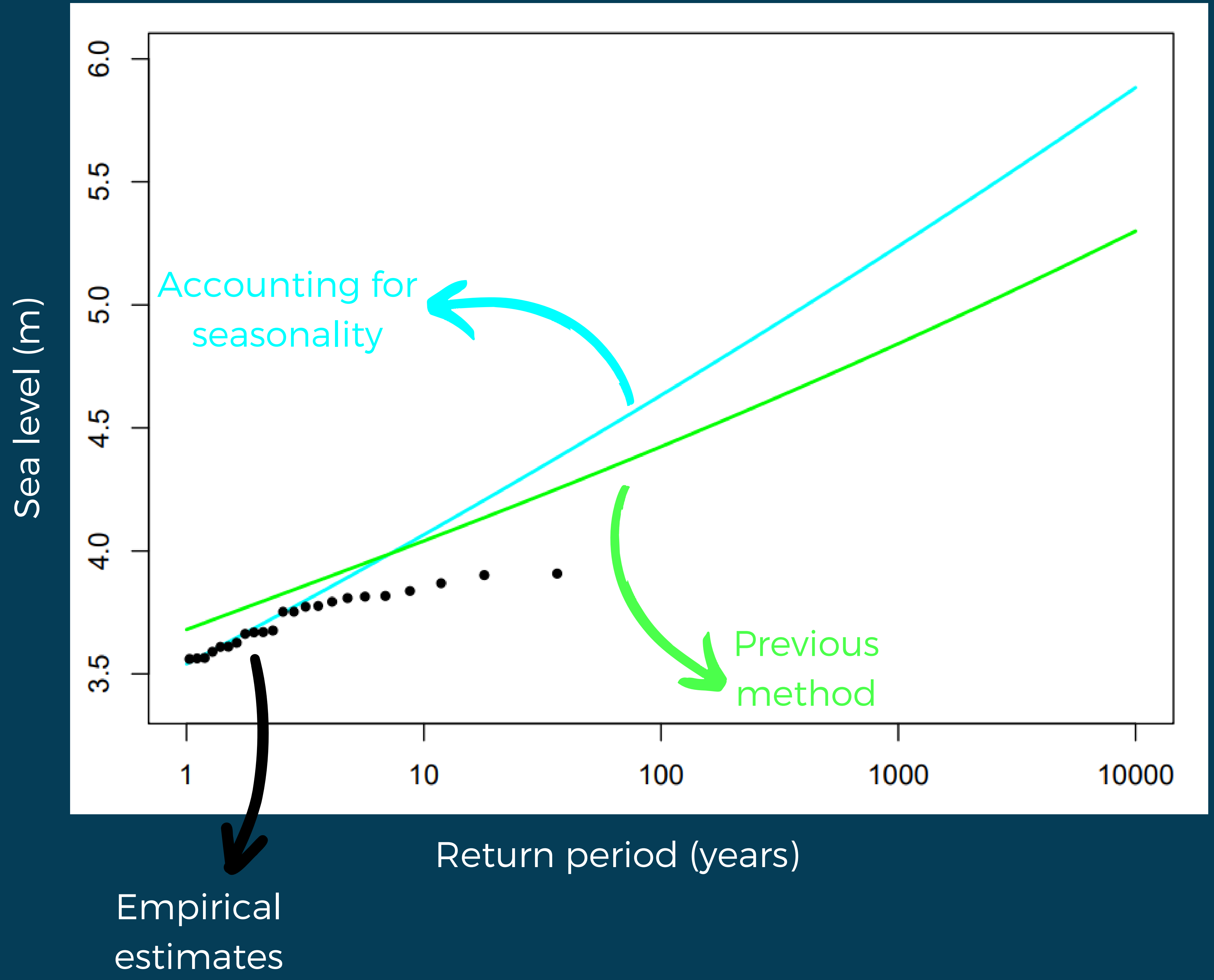
Clusters of extremes

Decluster the data by modelling cluster maxima

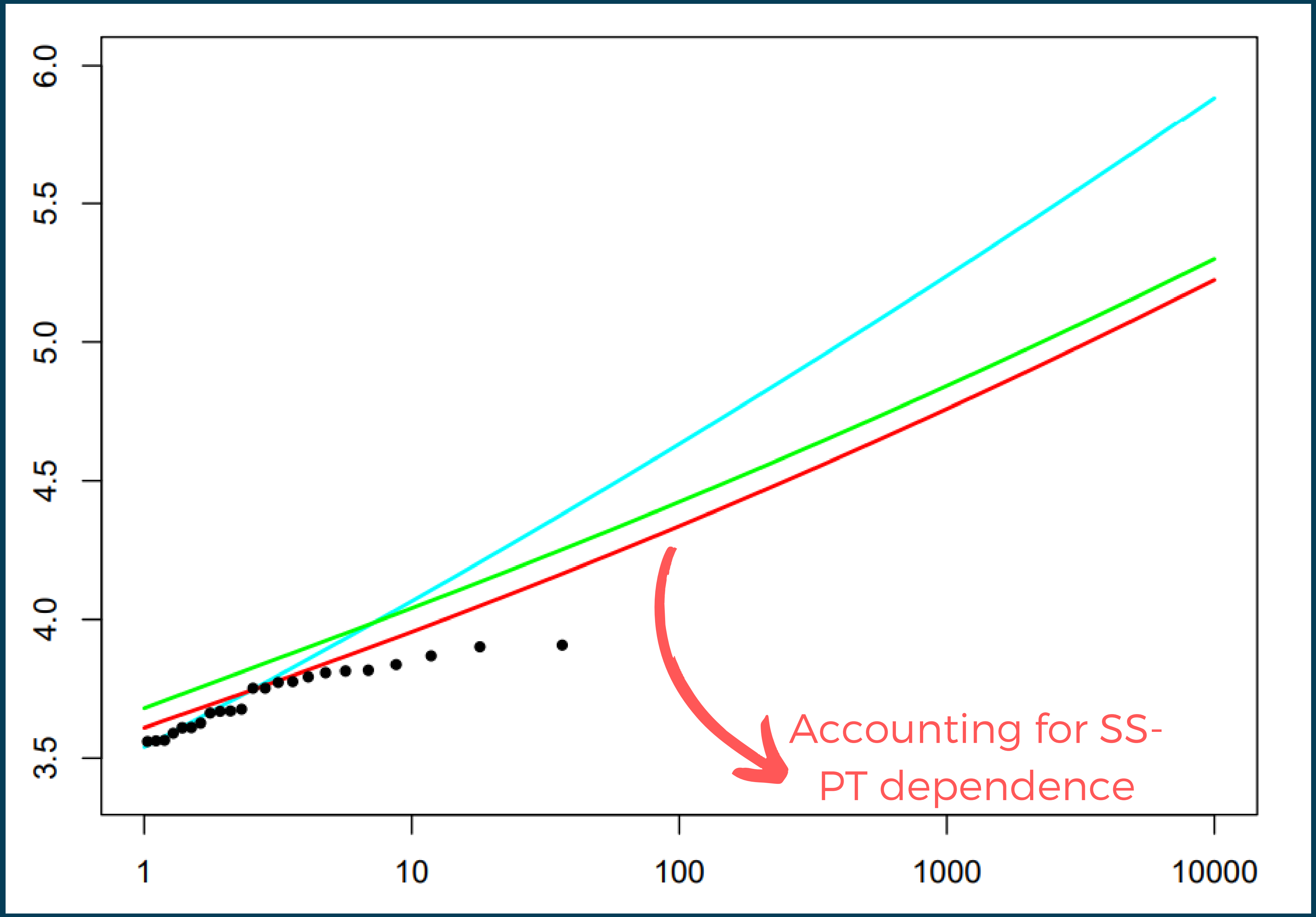


Results

Annual return levels



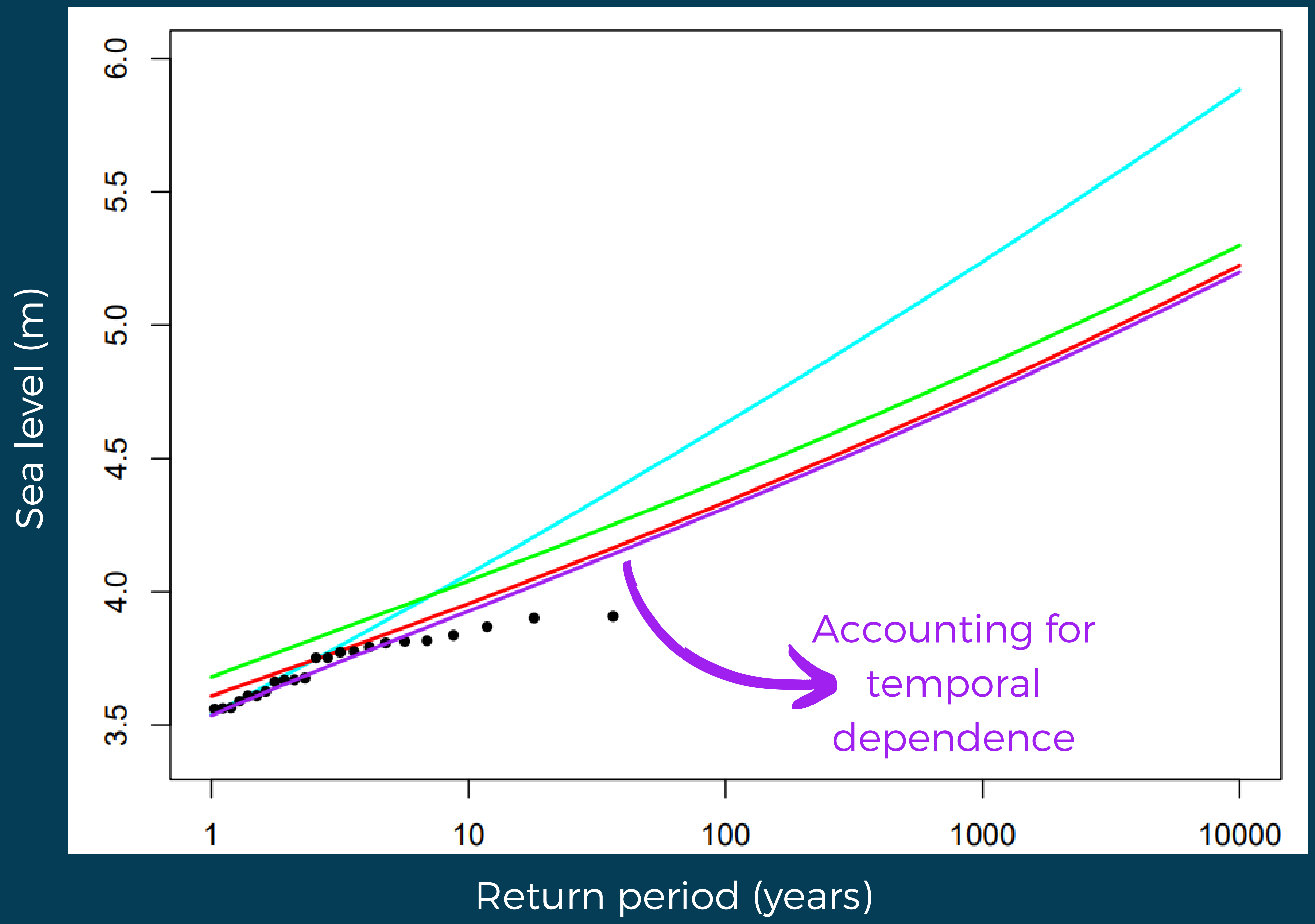
Annual return levels



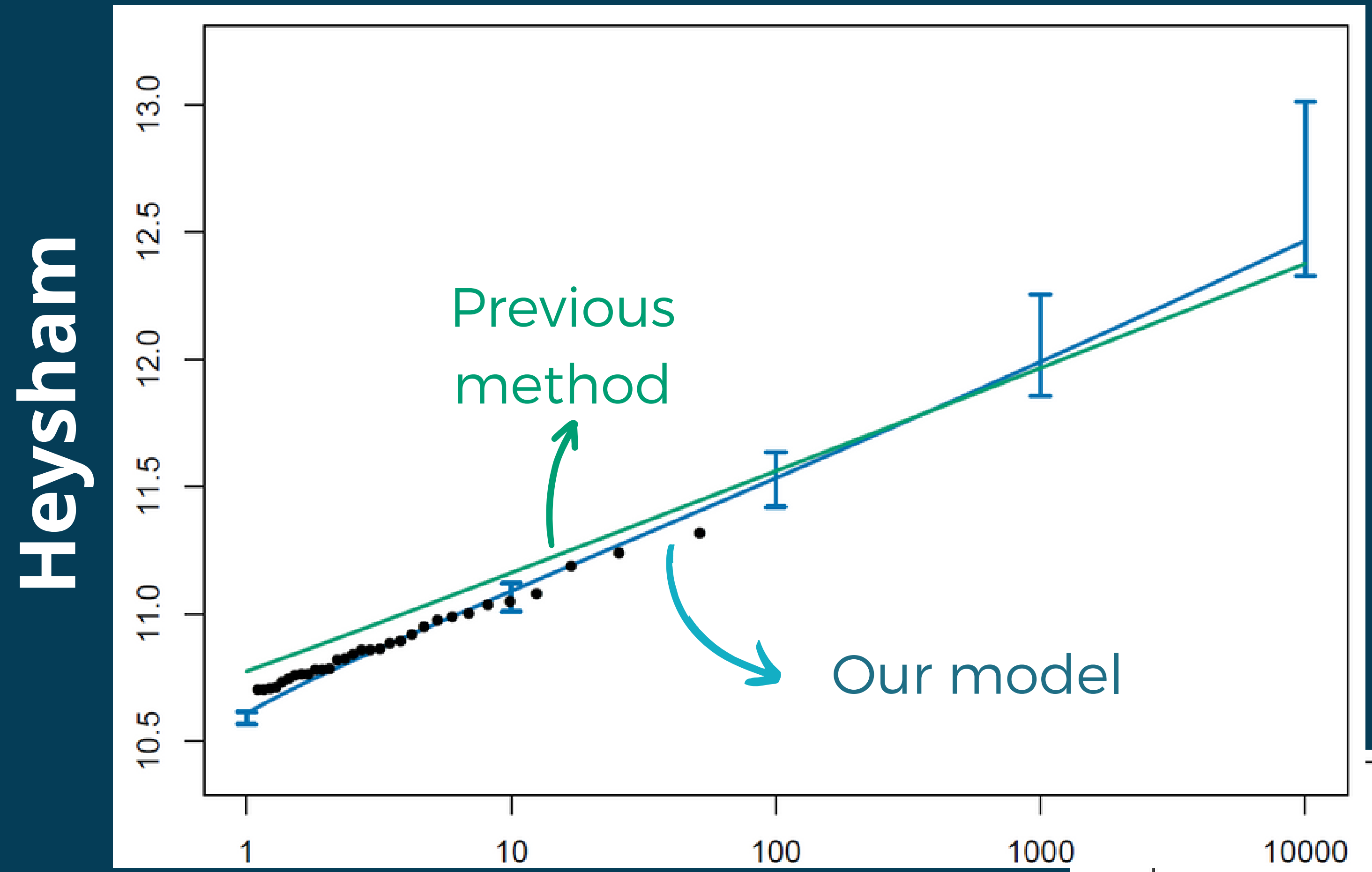
Sea level (m)

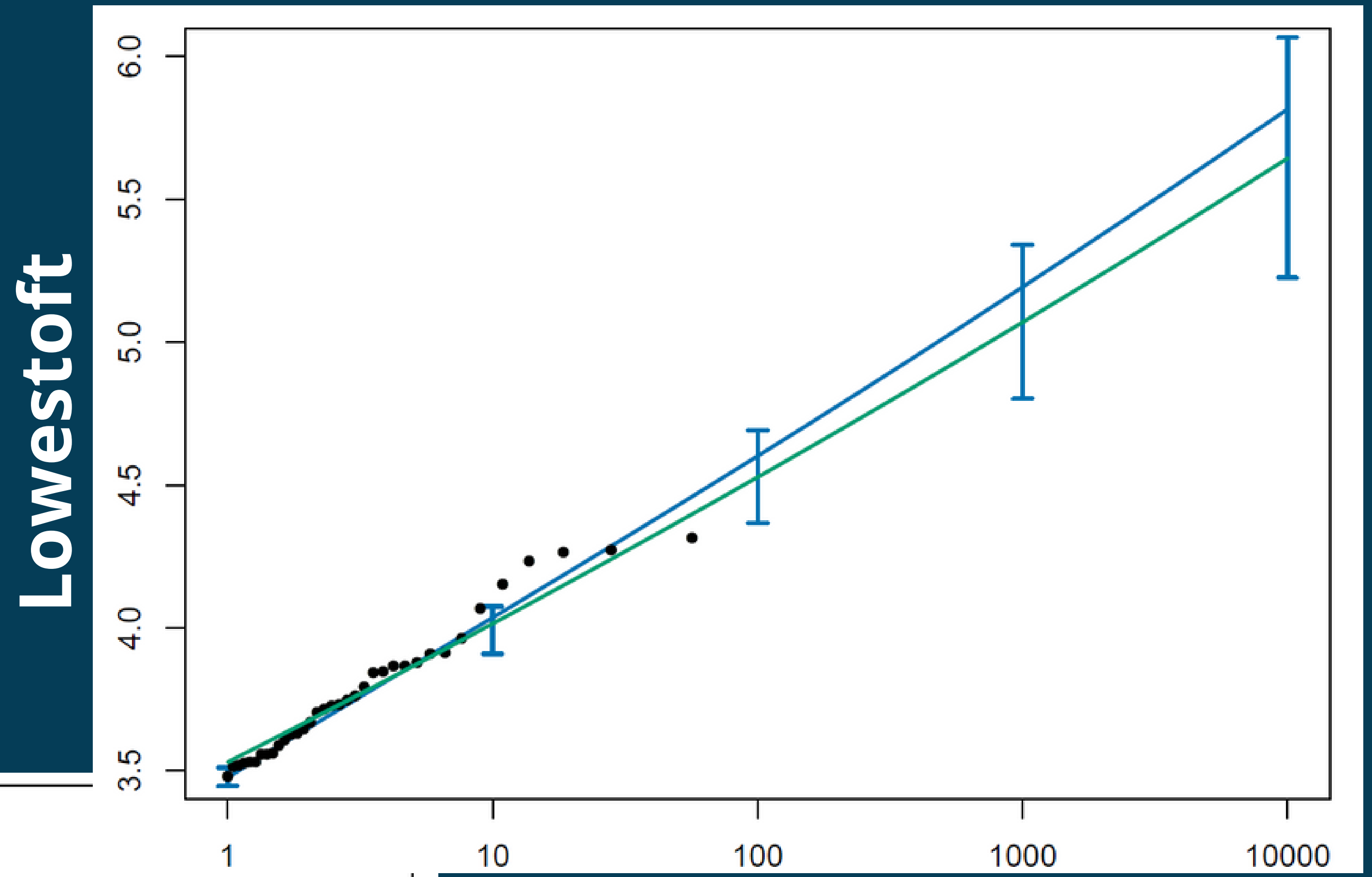
Return period (years)

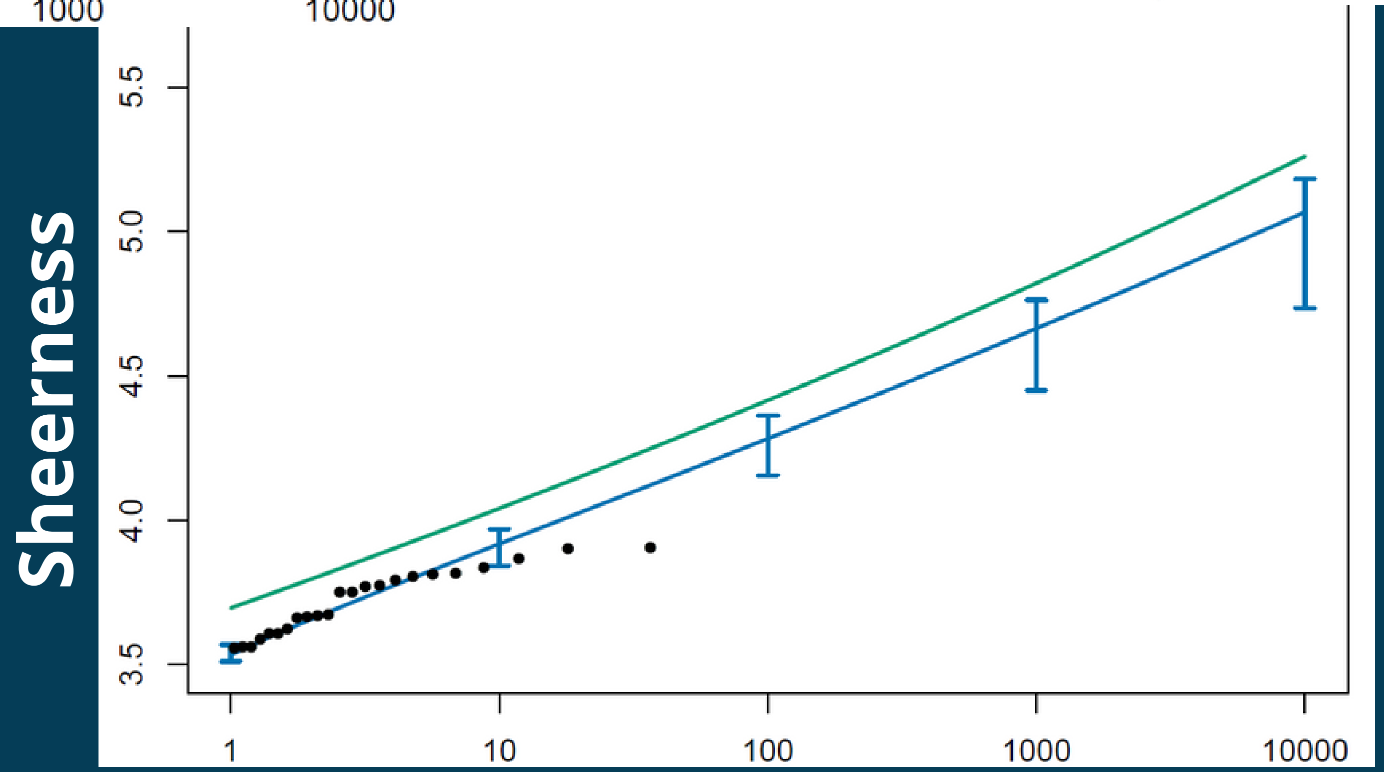
Annual return levels



Results at other sites







R-LADIES LANCASTER, NOVEMBER 2023

EXTREME SEA LEVEL ESTIMATION WITH R

Eleanor D'Arcy
Lancaster University

My Code

- *ESL_estimation.R* for return level estimation from our model,
- Exploratory analysis,
- Diagnostics,
- Return level estimation from different models,
- Data formulation
- Illustrative example at Lowestoft

eleanordarcy /
ESL_estimation

R package in development

with support from Dan Grose

Data

- Preprocessing to remove trends in mean sea level

```
> load('data_HEY.RData')
> head(data)
```

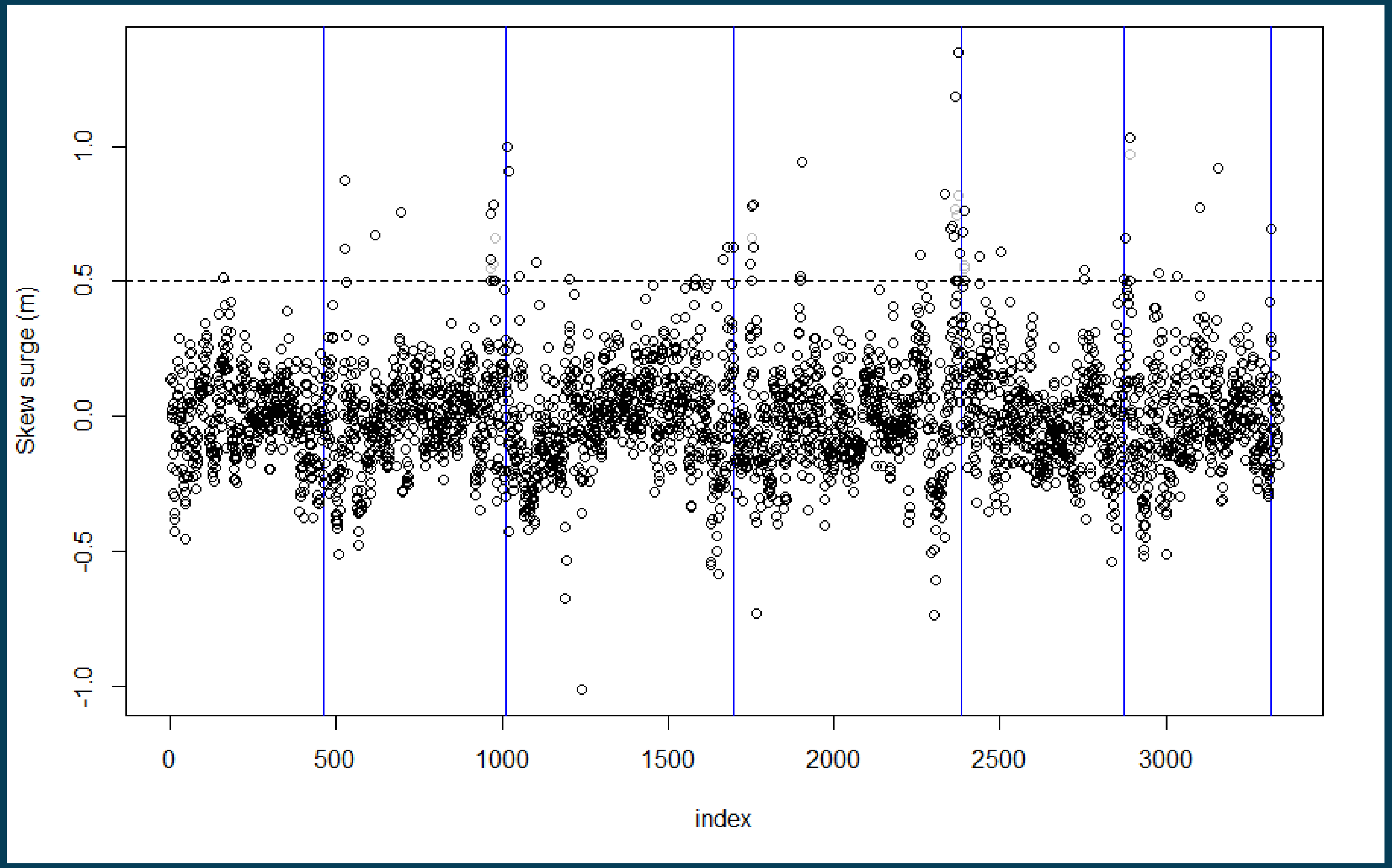
	date	day	month	year	skews	maxTide
1	1964-01-01	1	1	1964	-0.01219025	9.882
2	1964-01-02	2	1	1964	0.07980975	9.729
3	1964-01-02	2	1	1964	0.15980975	9.832
4	1964-01-03	3	1	1964	-0.04319025	9.334
5	1964-01-03	3	1	1964	0.01280975	9.552
6	1964-01-04	4	1	1964	-0.09819025	8.932

Declustering in R

```
> library(extRemes)
> extremalindex(data$skews, threshold = 0.5)
```

extremal.index	number.of.clusters	run.length
0.4264282	278.0000000	13.0000000

```
> EI <- decluster(data$skews, threshold = 0.5)
> plot(EI)
```



2010-2016 data

Model fitting

1. *GPD(σ,ξ)*

$$\sigma_{d,x} = \alpha_\sigma + \beta_\sigma \sin\left(\frac{2\pi}{365}(d - \phi_\sigma)\right) + \gamma_\sigma x$$

```
> ModS4(data = data_ex, init.par = c(0.2,0,100,0,0.01))
```

Shape

2. *GPD(λ)*

$$g(\lambda_{d,x}) = \alpha_\lambda^1 + \beta_\lambda^1 \sin\left(\frac{2\pi}{365}(d - \phi_\lambda^1)\right) + \tilde{x}\left[\alpha_\lambda^2 + \beta_\lambda^2 \sin\left(\frac{2\pi}{365}(d - \phi_\lambda^2)\right)\right]$$

```
> ModR1(data = data, init.par = c(0.05,0,100,0,0,100))
```

Return level estimation

```
> RL_est(p, interval=c(quantile(data$sealevel, 0.5), max(sea$level)+3))
```

Annual exceedance probability

- *p=0.0001* for *10,000* year return level
- Can be a vector

```
> CI_est(p, interval=c(quantile(data$sealevel, 0.5), max(sea$level)+3),
  n.boot=200, block.length=10, prob=0.95)
```

Impact

- Implementation at EDF

- Working with the Environment Agency for improving existing best ESL estimates in the Coastal Flood Boundary Report

- Thames Estuary 2100 plan

References

- Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.
- D'Arcy, E., Tawn, J. A., Joly, A., & Sifnioti, D. E. (2023). *Accounting for seasonality in extreme sea-level estimation.* The Annals of Applied Statistics, 17(4), 3500-3525.
- D'Arcy, E., Tawn, J. A., & Sifnioti, D. E. (2022). *Accounting for climate change in extreme sea level estimation.* Water, 14(19), 2956.
- GitHub repository: *github.com/eleanordarcy/ESL_estimation/*

R-LADIES LANCASTER, NOVEMBER 2023

THANK YOU
ANY QUESTIONS?

Eleanor D’Arcy
Lancaster University