EXTRAPOLATION OF EXTREME SEA LEVELS: INCORPORATION OF OVER-THRESHOLD-MODELING TO THE JOINT PROBABILITY METHOD

Franck Mazas, **Luc Hamm**, ARTELIA Maritime, Grenoble

Xavier Kergadallan, CETMEF, Brest

CONTEXT AND PURPOSE OF THE STUDY

WHAT IS AT STAKE

- Storm **XYNTHIA**, France, 27-28th February 2010: > 30 people drowned because of coastal flooding
- Need for accurate estimation of extreme sea levels for risk assessment in coastal areas

STATE OF THE ART Two main approaches for estimation of extreme sea levels:

- **Direct approach**: direct extrapolation on sea levels data: OK for surge-dominant areas
- **Indirect approach**: separate analysis of deterministic component (astronomical tide) and stochastic component (meteorological residual, surge) then distribution of sea levels derived by convolution of tide and surge distributions. Model often referred to as **Joint Probability Method** (**JPM**) (Pugh & Vassie 1979, Tawn & Vassie 1989)
- Within this framework, statistical extrapolation is performed on surge component only. Different methods used in literature: Annual Maxima Method, r-largest Method…

AIM OF THE STUDY Incorporation of Over-Threshold Modeling methods for surge extrapolation:

- **Peaks-Over-Threshold (POT)** approach: use of a physical threshold for identification of surge events (positive or negative storms) from time series of regular (say, hourly) auto-correlated observations then use of a statistical threshold for selection of extreme i.i.d. data (Bernardara *et al.*, 2012);
- Extrapolation of (+/-) surge peaks by **Poisson-GPD** model and computation of confidence intervals

CASE STUDY Method illustrated with data from **Brest**, France (levels reduced to local chart datum noted ZH)

Figure 1 *La-Faute-sur-Mer, France in the aftermath of Xynthia (February 2010)*

LIST OF SYMBOLS

- $S,\ T,\ Z$: surge, astronomical tide, sea level
- $P^S,\ P^Z$: surge peak, sea level peak
- n: total number of sequential data in time series
- $n_S(u_S)$ number of observed sequential surges above u_s
- d_S, d_Z: mean number of sequential data per surge / sea level event
- N^S, N^Z: number of surge / sea level events (peaks)
- u_p, u_S: physical threshold, statistical threshold
- F_S, F_Z: distribution of sequential surges / sea levels
- G_S , G_Z : conditional distribution of extreme sequential surges / sea levels above a threshold
- G_{P^S} , G_{P^Z} : conditional distribution of surge / sea level events (peaks)
- λ^Z: mean number of sea level events per year
- ν: mean number of sequential values per year

ASTRONOMICAL TIDE

DETERMINISTIC COMPONENT Entirely predictible

- Continuous time series over a saros (18.6 years) necessary and sufficient for full determination of astronomical tide distribution
- This time series can be generated if the local harmonic constants are known
- Sequential values of astronomical tide modeled by a **non-parametric kernel density estimator**

MIXTURE MODEL FOR SURGE COMPONENT

BULK DISTRIBUTION Frequent values modeled by a **non-parametric kernel density estimator**

TAIL EXTRAPOLATION FOR SURGE EVENTS

- **Physical declustering** of extreme (+/-) surge events using a **physical threshold** u_p → setting up a sample of **i.i.d. surge peaks**
- Determination of a **statistical threshold** u_s using GPD properties → setting up a sample of N extreme **surge peaks excesses**
- Fit of extreme surge peak excesses by a parametric distribution such as **GPD** or **Weibull** distribution (***Fig. 2***) (Mazas & Hamm, 2011)

FROM EXTREME SURGE EVENTS TO EXTREME SEQUENTIAL SURGES

- Necessity to account for mean duration of extreme surge events → equivalent of **extremal index** for surges
- Relation between conditional distributions of surge peaks and of sequential (hourly) surges above u_S:
  $G_S(s) = 1 - \frac{N^S}{n_S(u_S)} d_S(s)\left[1 - G_{P^S}(s)\right]$
- d_S modeled by linear regression (on log-log scale) until $d_S = 1$ (***Fig. 3***)
- Connection of tail to bulk using law of threshold exceedances: $F_S(s) = 1 + \frac{n_S(u_S)}{n}\left[G_S(s) - 1\right]$ for $s > u_s$ (***Fig. 4*** and ***5***)

CONFIDENCE INTERVALS Computed by **parametric bootstrap**

- Approximation by an equivalent parametric distribution
- Same process for equivalent CI distributions as for the tails

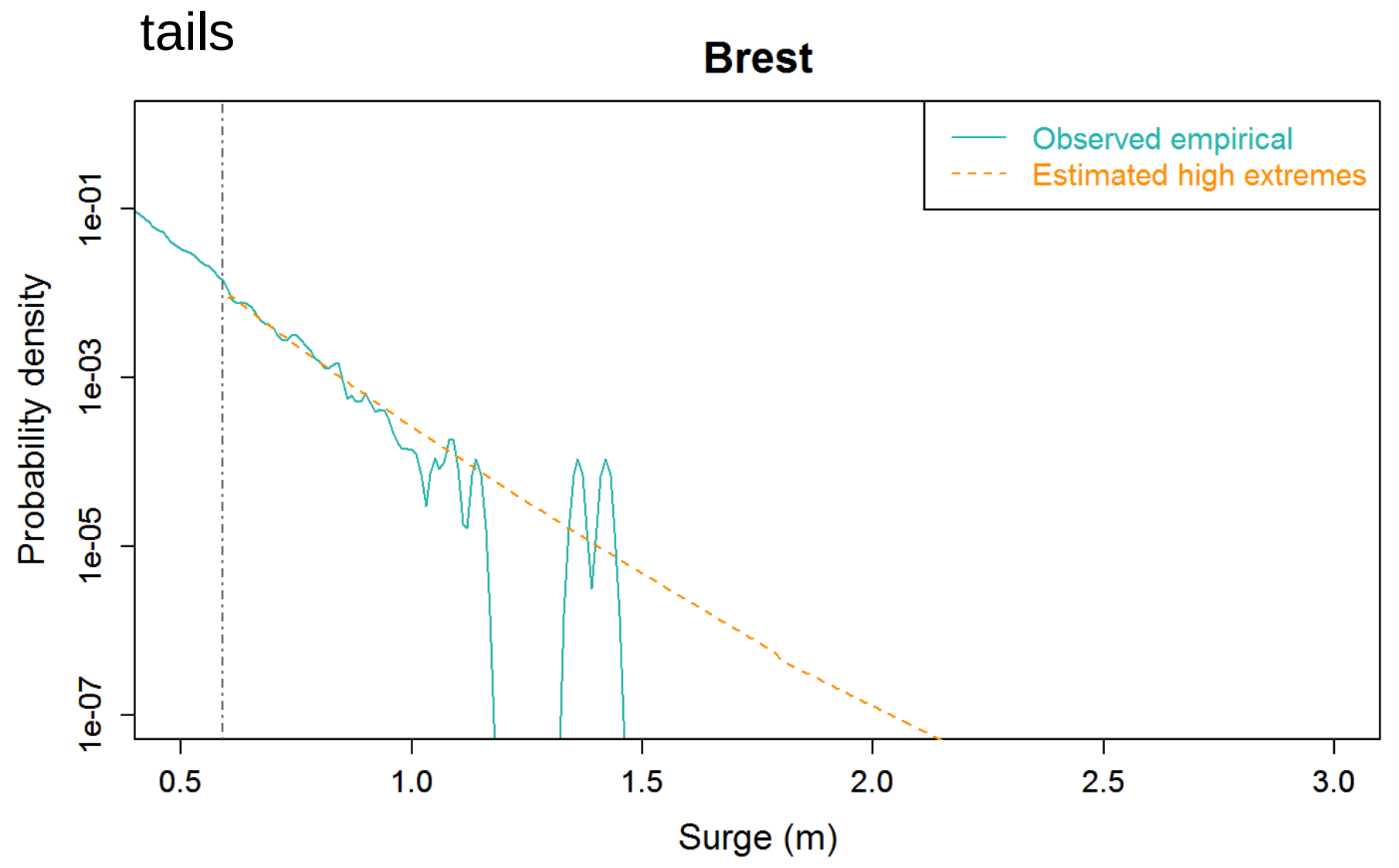

Figure 4 *Distribution of extreme hourly surges: model vs. observed*

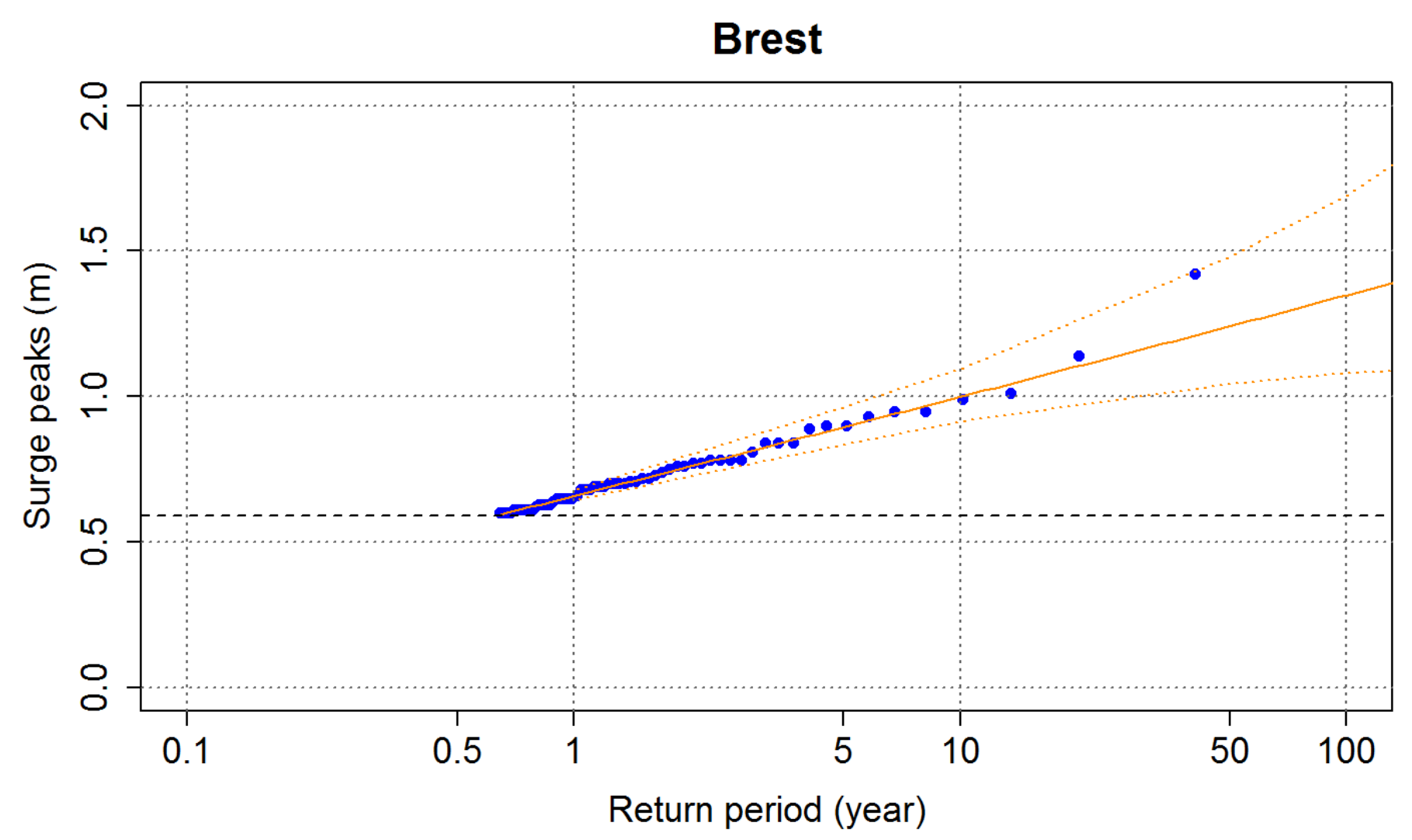

Figure 2 *Extrapolation of extreme surge peaks*

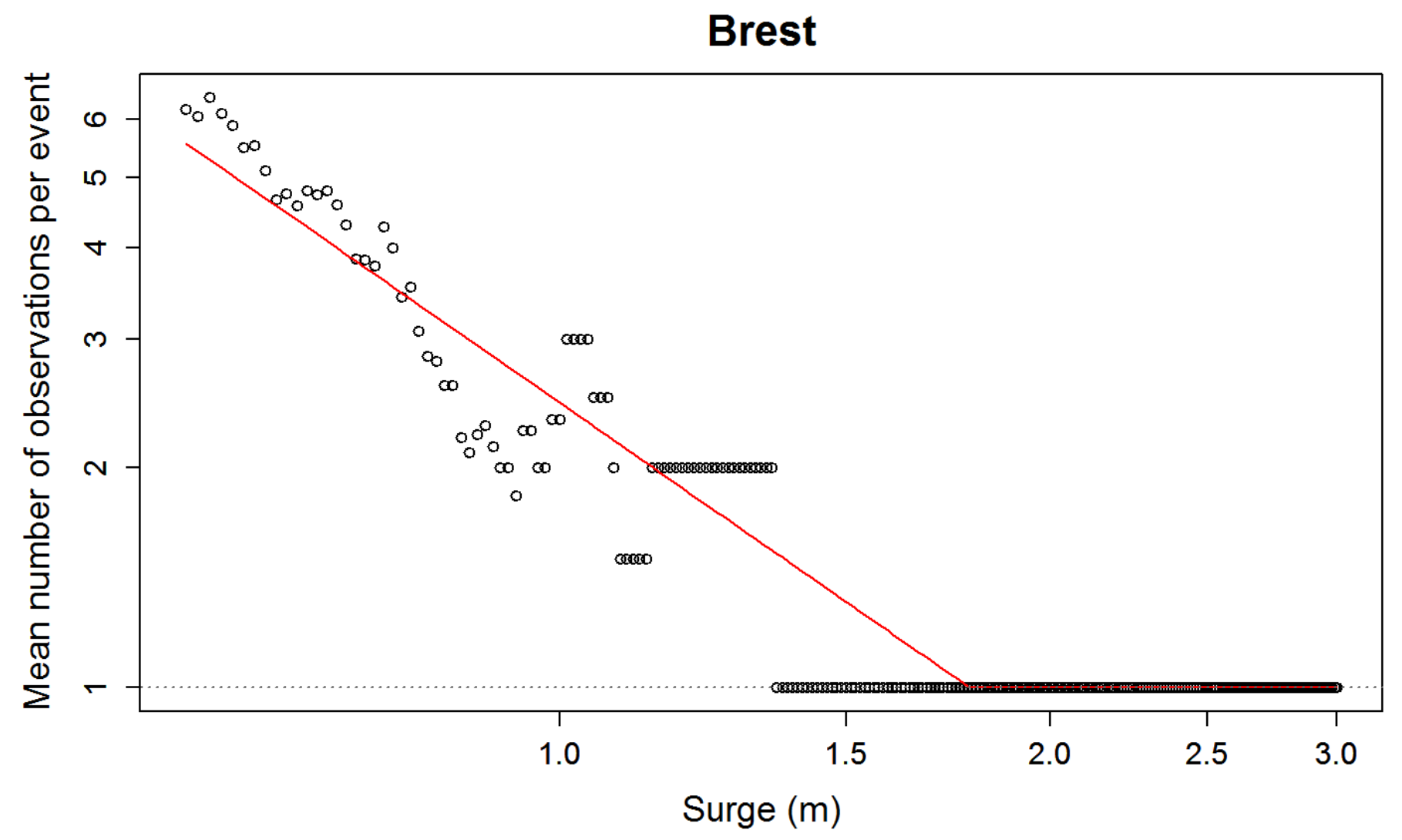

Figure 3 *Mean number of hourly surge values per event: model vs. observed*

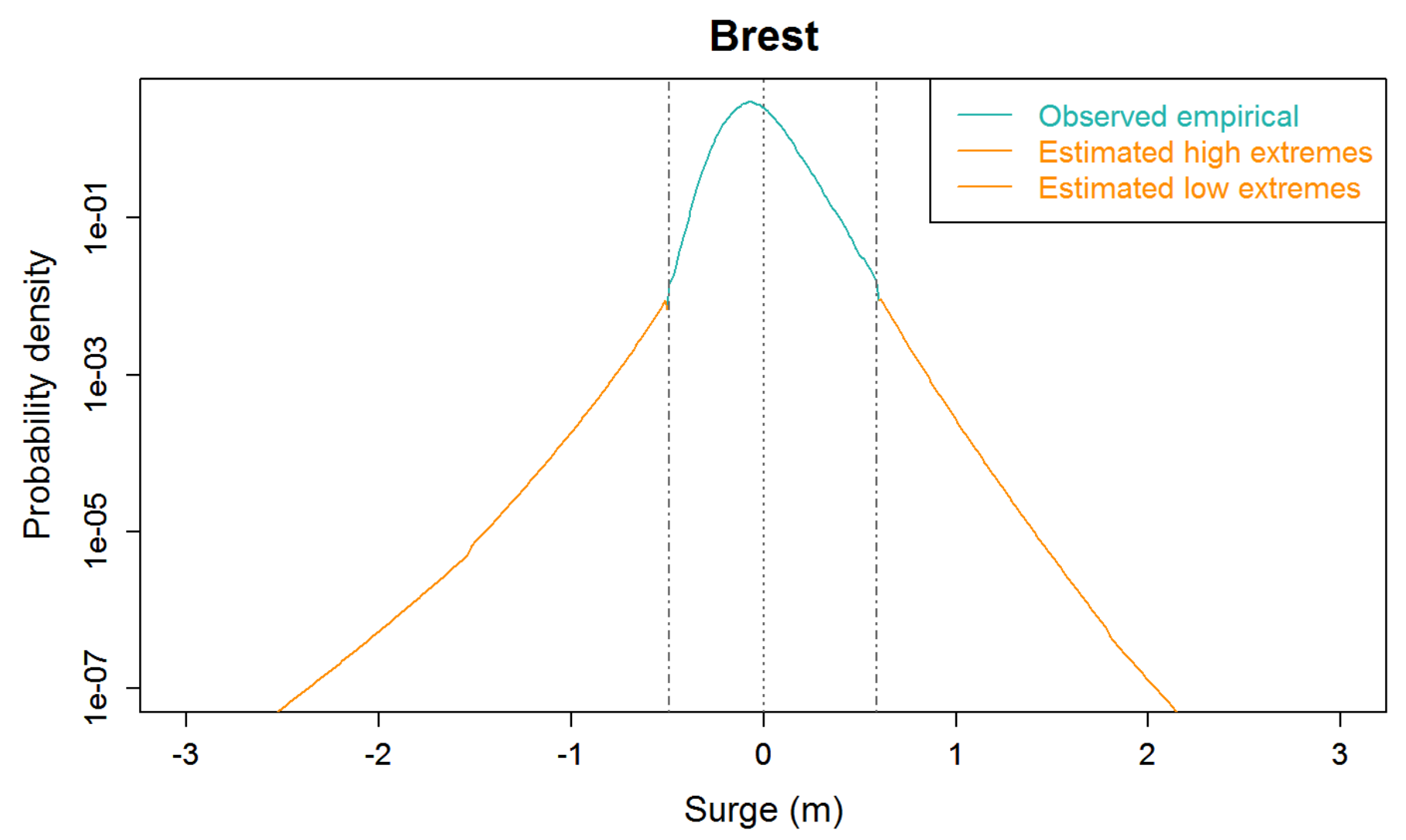

Figure 5 *Distribution of hourly surges: empirical bulk and estimated tails*

DISTRIBUTION OF SEA LEVELS

SEQUENTIAL (HOURLY) SEA LEVELS

- **Convolution** of distributions of sequential values of astronomical tide and surge (***Fig. 6***)
- Computation of return periods or return levels for hourly sea levels (***Fig. 7***):
  $T(z) = \frac{1}{\nu[1-F_Z(z)]}$

DISTRIBUTION OF EXTREME SEA LEVEL EVENTS

- Coastal protection requires estimation of return levels for extreme sea levels events, or peaks

→ Equivalent of **extremal index** for sea levels based on mean number of sequential observations per event (***Fig. 8***)

- Relation between conditional distribution of sea level peaks G_{P^Z} and distribution of sequential sea levels above a threshold u^Z:
  $G_{P^Z}(z) = 1 + \frac{n}{N^Z}\frac{1}{d_Z(z)}\left[1 - F_Z(z)\right]$
- Threshold u^Z d efining events taken as **MHWS** (Mean High Water Spring)
- Computation of return periods or return levels (***Fig. 9***):
  $T(z) = \frac{1}{\lambda^Z[1-G_{P^Z}(z)]}$
- When tide is highly predominant, difference is very little between extreme hourly sea levels and extreme sea level peaks (one hourly value per event)

CONFIDENCE INTERVALS

- Applying convolution on "equivalent CI parametric distributions" allows fast approximation of CI for sea level events

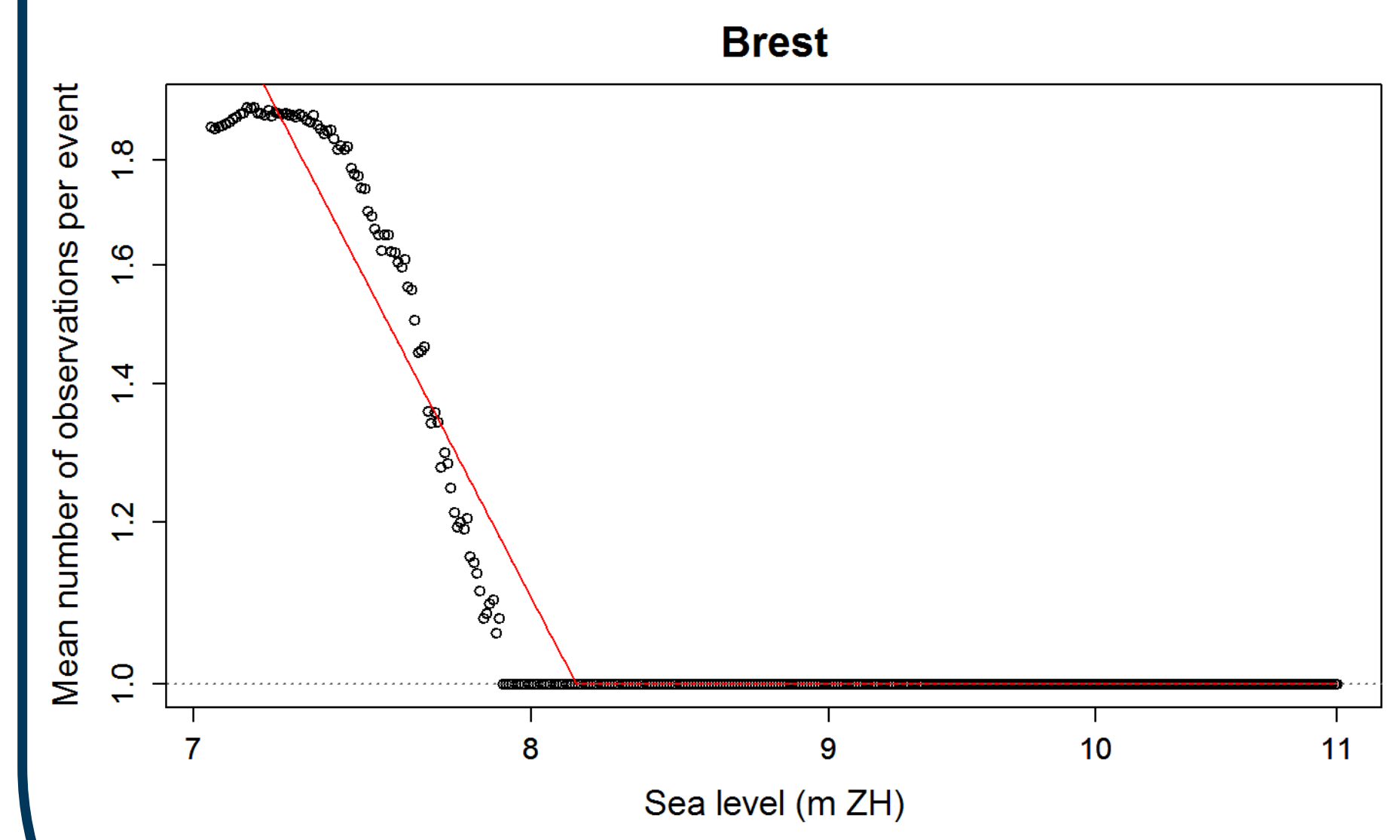

Figure 8 *Mean number of hourly sea level values per event: model vs. observed*

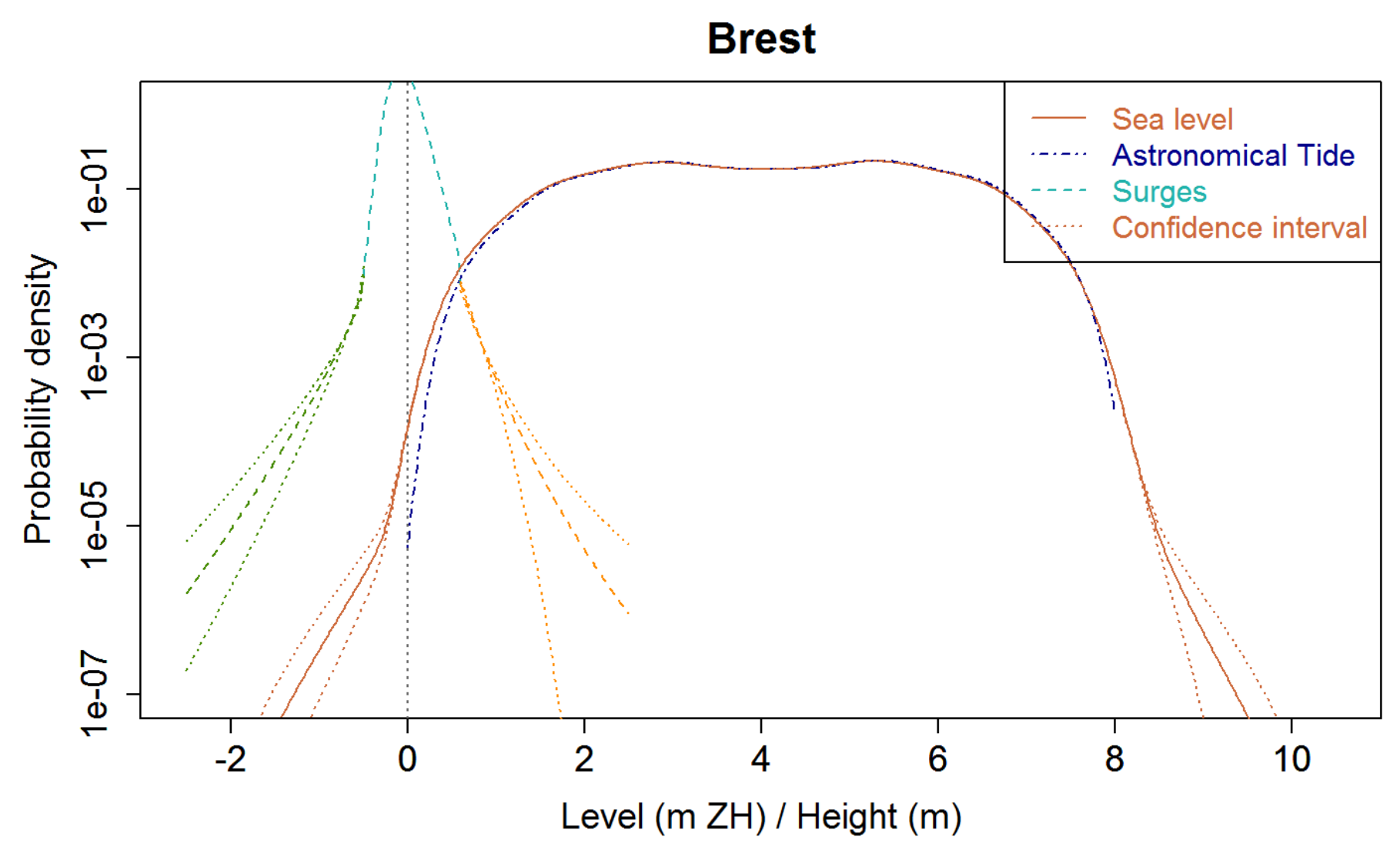

Figure 6 *Distributions of hourly values of tide, surge, sea levels*

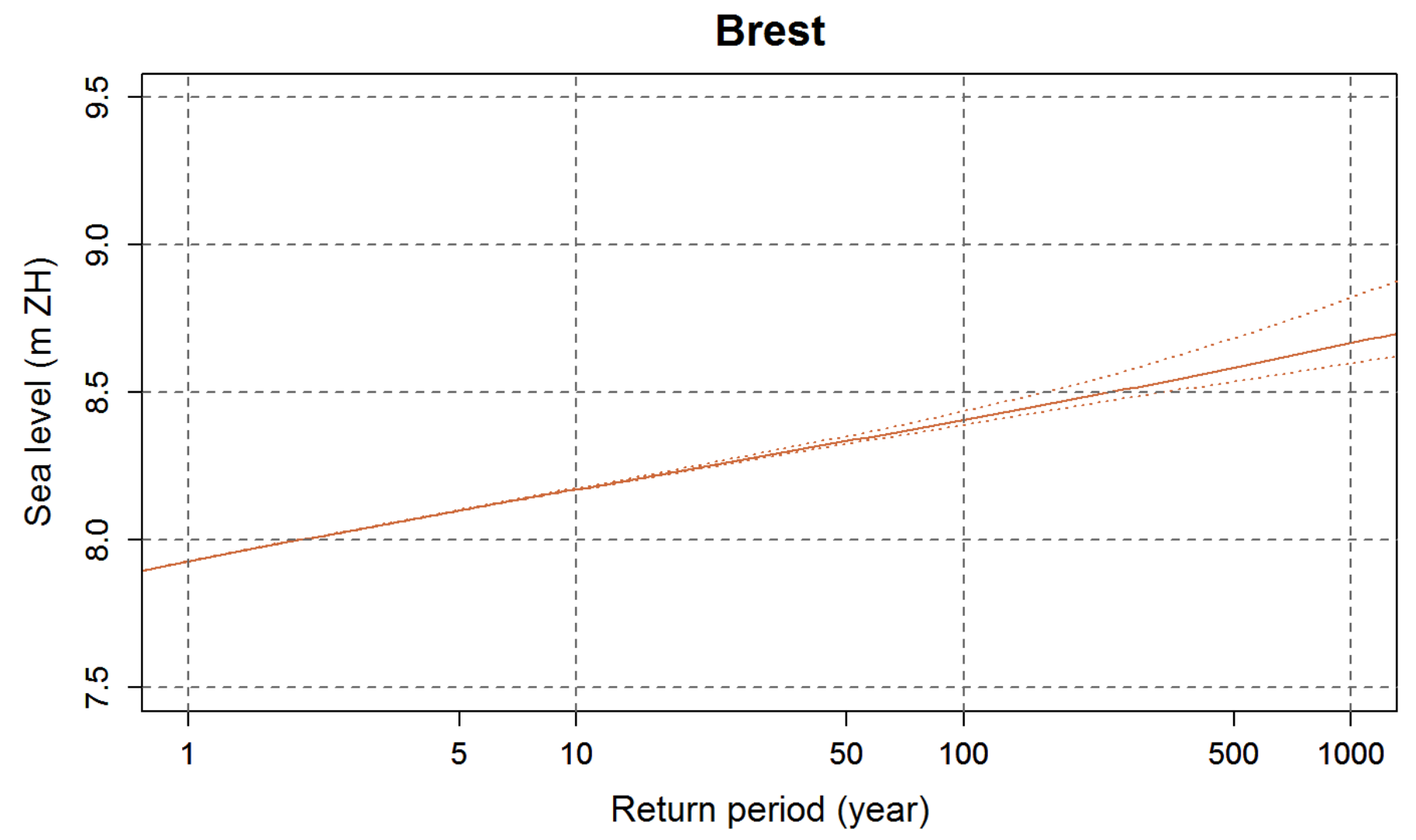

Figure 7 *Return periods for hourly sea levels*

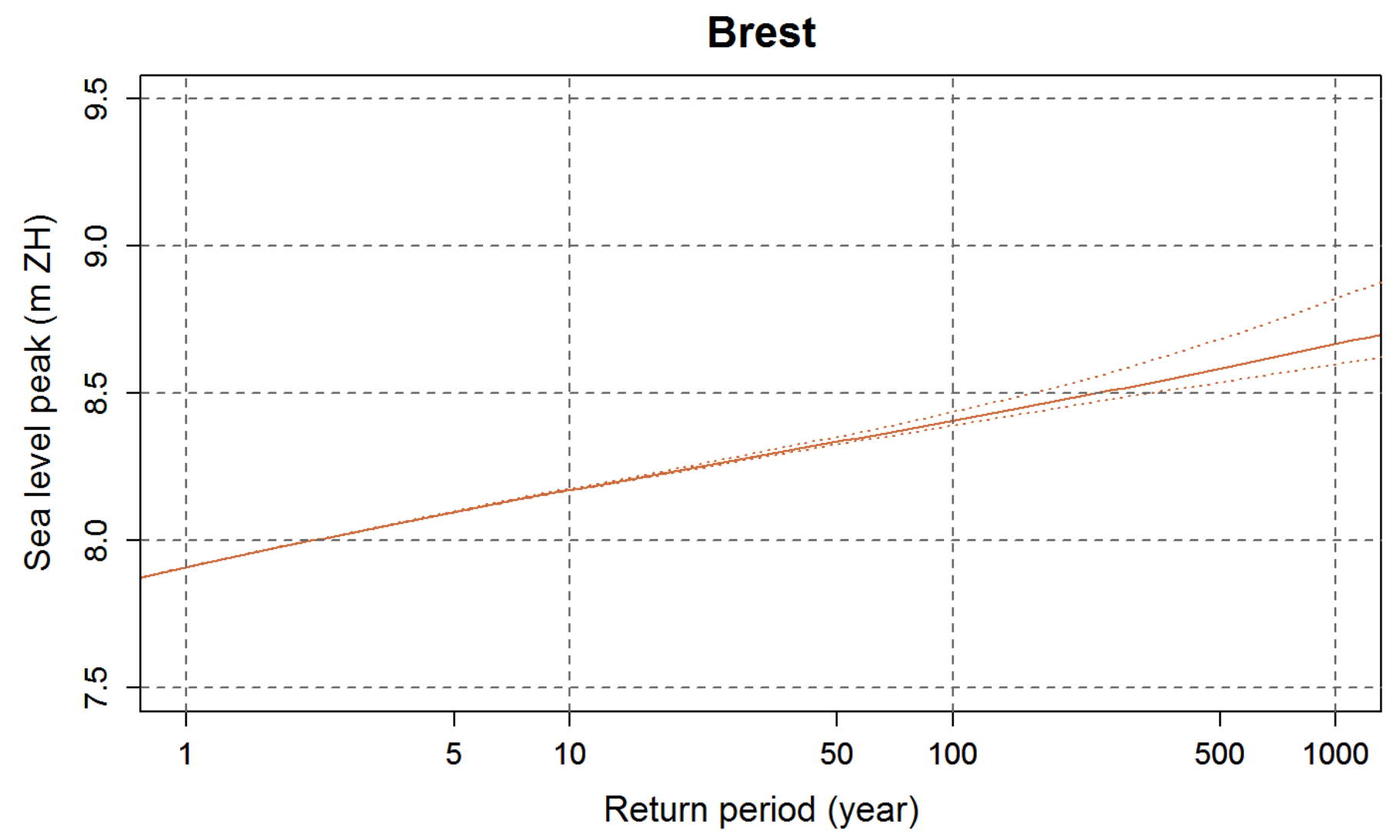

Figure 9 *Return periods for sea level events*

CONCLUSIONS AND PERSPECTIVES

CONCLUSIONS

- Incorporation of OTM (POT approach for extreme data selection and Poisson-GPD model for statistical extrapolation) should improve **accuracy** and **reliability** of extreme sea level assessments
- Interest: clear distinction between **sequential observations** and **events** (usually defined as peaks)
- Possibility to apply this methodology to **skew surges** and **high tide sea levels**
- Method requiring **high-quality data over a long-period** (ideally a saros: 18.6 years) that are only available for a few locations

PERSPECTIVES

→ Application to **numerical modeling** output in order to have long-period continuous time series in any location needed

→ Necessity to account for **dependency** between surge heights and astronomical tide levels

References :

Bernardara, P., Mazas, F., Weiss, J., Andreewsky, M., Kergadallan, X., Benoît, M., Hamm, L., 2012. On the two-step threshold selection for over-threshold modelling. *Coast. Eng. Proc.*, **1**(33).

Mazas, F., Hamm., L., 2011. A multi-distribution approach of POT methods for determining extreme wave heights. *Coastal Eng.*, **58**, 385-394.

Pugh, D.T., Vassie, J. M., 1979. Extreme sea levels from tide and surge probability. *Proc. 16th Coast. Engng Conf.,* 1978, Hamburg. A. S. of Civil Engineers, Ed., **1**, New York, pp. 911-930.

Tawn, J. A., Vassie, J. M., 1989. Extreme sea levels: the joint probabilities method revisited and revised. *Proc. Instn Civ. Engrs*, Part 2, **87**, pp. 429-442.

Contact:

Franck Mazas - ARTELIA Maritime
6, rue de Lorraine
38130 Echirolles – FRANCE
franck.mazas@arteliagroup.com