WAVE EXTREMES IN THE NORTHEAST ATLANTIC

O.J. Aarnes[1,2], Ø. Breivik[1] and M. Reistad[1]

[1]Meteorological Institute of Bergen, Norway

[2]University of Bergen, Norway

Objective

The main objective of this study is to compute 100-year return values of significant wave height (H_s) using a new hindcast developed by the Norwegian Meteorological Institute. This regional hindcast covers the Northeast-Atlantic and spans the period 1958-2009.

The return value estimates are based upon three different approaches within classical extreme value theory, i.e. the Generalized Extreme Value(GEV) distribution, the r-largest order statistics and the Generalized Pareto(GP) distribution. Here, we investigate the qualitative difference between the estimates and their corresponding confidence intervals.

Norwegian Reanalysis - NORA10

NORA10 (NOrwegian ReAnalysis 10 km) is the latest contribution to the series of wave hindcasts developed by the Norwegian Meteorological Institute, see Reistad et al. (2007, 2010). This regional hindcast utilizes boundary conditions obtained with the ERA-40 reanalysis(Uppala and et al., 2005)(1958 - August 2002) and is extended using EC-analysis (September 2002 -2009), producing three hourly wave fields at 10 km grid spacing for the period 1958-2009. For wave simulations a modified version of the WAM cycle 4 (Komen et al., 1996) is run on the same grid as 10km HIRLAM model, nested inside a larger 50 km WAM-model, see Figure 1. No wave energy is advected into the model domain of the WAM50, only wind dependent wave spectra are set at the boundaries.

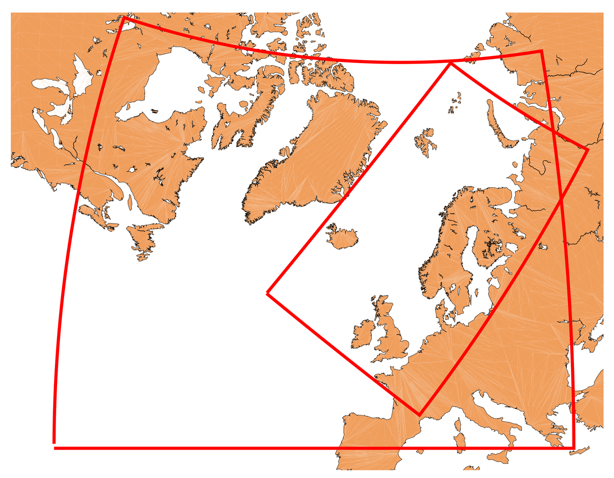

Figure 1: Outer domain: WAM50. Inner domain: WAM10/Hirlam10 - hindcast area.

Preliminary conclusions and further work

We obtain fairly similar 100-year return value estimates of H_s using three different statistical approaches, however, we see a few local exceptions we want to investigate further. Primarily, we see an increased precision in the confidence intervals when utilizing methods of increased data subsets, like the POT approach and the r-largest order statistics.

Future emphasis will be made on the POT approach and ways to optimize the threshold selection. Finally, a qualitative comparison of the three methods will be made using different goodness-of-fit tests.

References

S. Coles. *An introduction to Statistical Modelling of Extreme Values.* Springer-Verlag, 2001.

G. J. Komen, M. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen. *Dynamics and Modelling of Ocean Waves*. Cambridge University Press, 1996.

M. Reistad, O. Breivik, and H. Haakenstad. A high-resolution hindcast study for the north sea, the norwegian sea and the barents sea. In *Proceedings of the 10th International Workshop on Wave Hindcast and Forcasting and Coastal Hazard Symposium.*, 2007.

M. Reistad, O. Breivik, H. Haakenstad, O. J. Aarnes, and B. R. Furevik. A high-resolution hindcast of wind and waves for the north sea, the norwegian sea and the barents sea. *Journal of Geophysical Research, Oceans.(revision)*, 2010.

S. M. Uppala and et al. The era-40 re-analysis. *Q. J. R. Meteorol. Soc.*, 131:2961–3012, 2005.

Method

The following analysis is based on the extreme value theory presented in Coles (2001). We are using two different but closely related families of distributions, determined by the selection of data subset, i.e.:

- The Generalized Extreme Value(GEV) distribution:

$$G(z) = exp\left\{-\left[1+\alpha\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1}$$

, where z represents the wave height and ξ , σ , and μ are the shape, scale and location parameters of the distribution function, respectively.

- The Generalized Pareto(GP) distribution:

$$H(z_u) = 1-\left(1+\frac{\xi z_u}{\tilde{\sigma}}\right)^{-1/\xi} \tag{2}$$

$$\tilde{\sigma} = \sigma + \xi(u-\mu) \tag{3}$$

, where $z_u = z - u > 0$ and u represents a relative large threshold.

Return value plots:

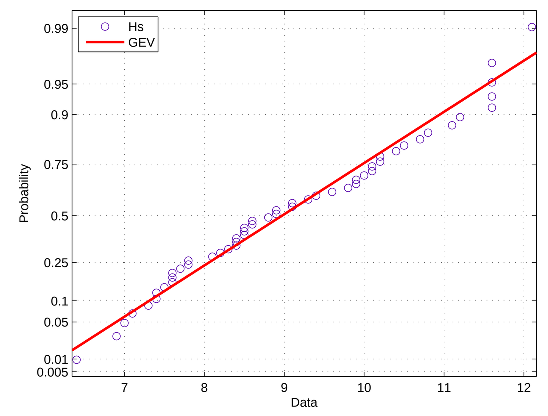

Figure 2: Annual maximum/GEV

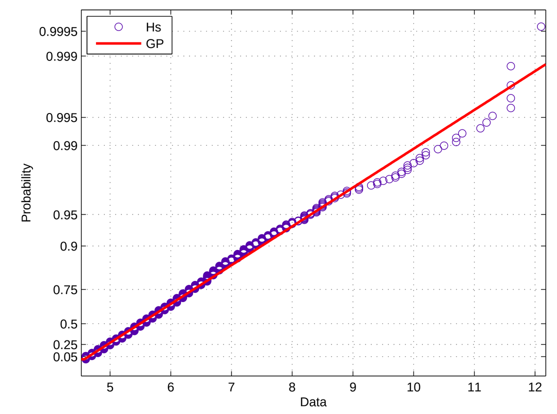

Figure 3: POT/GP

Distribution fitting

We utilize the *maximum likelihood* in our parameter estimation, i.e. maximizing the likelihood function L iteratively, or preferably minimizing the negative log-likelihood equation l

$$L(\theta) = \prod_{i=1}^{n} f_i(z_i;\theta) \Longrightarrow -l(\theta) = -logL(\theta) = \sum_{i=1}^{n} f_i(z_i;\theta) \tag{4}$$

, where f_i represents the probability density for each $z_i, i = 1....,n$, defined by the distribution parameters $\theta(\xi,\sigma,\mu)$.

Data subsets

- Annual maximum → GEV
  1 entry per year
- r-largest order statistics → GEV (modified likelihood function L)
  2-7 entries per year (based on the likelihood ratio test)
- Peaks-over-threshold(POT)→ GP
  ∼ 6-60 entries per year (based on the mean residual life plot)

100-year return value estimates of significant wave height

Annual maximum/GEV

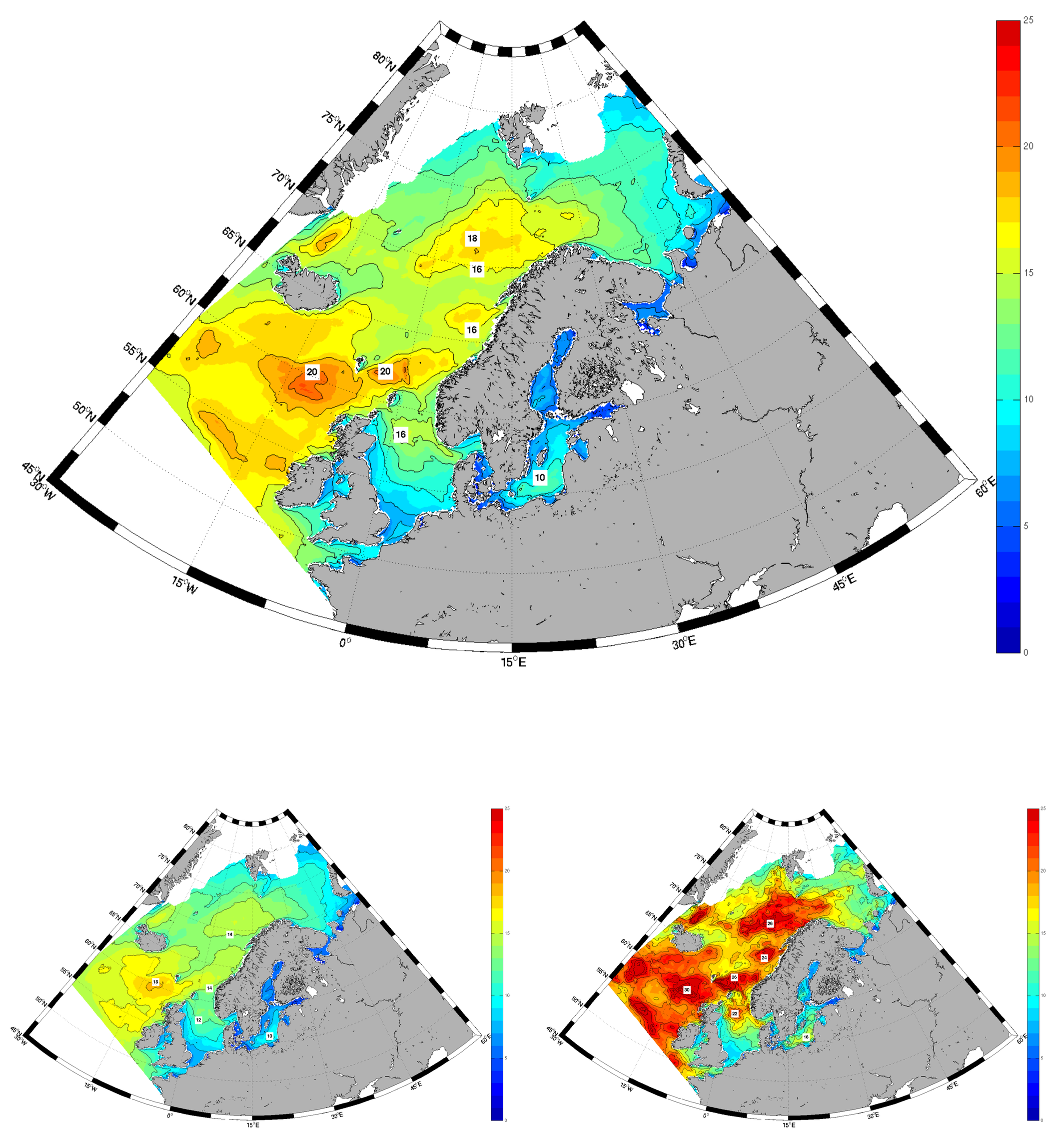

Figure 4: Top: best estimate.
Bottom: 5% and 95% confidence limits.

***r*-largest order statistics/GEV**

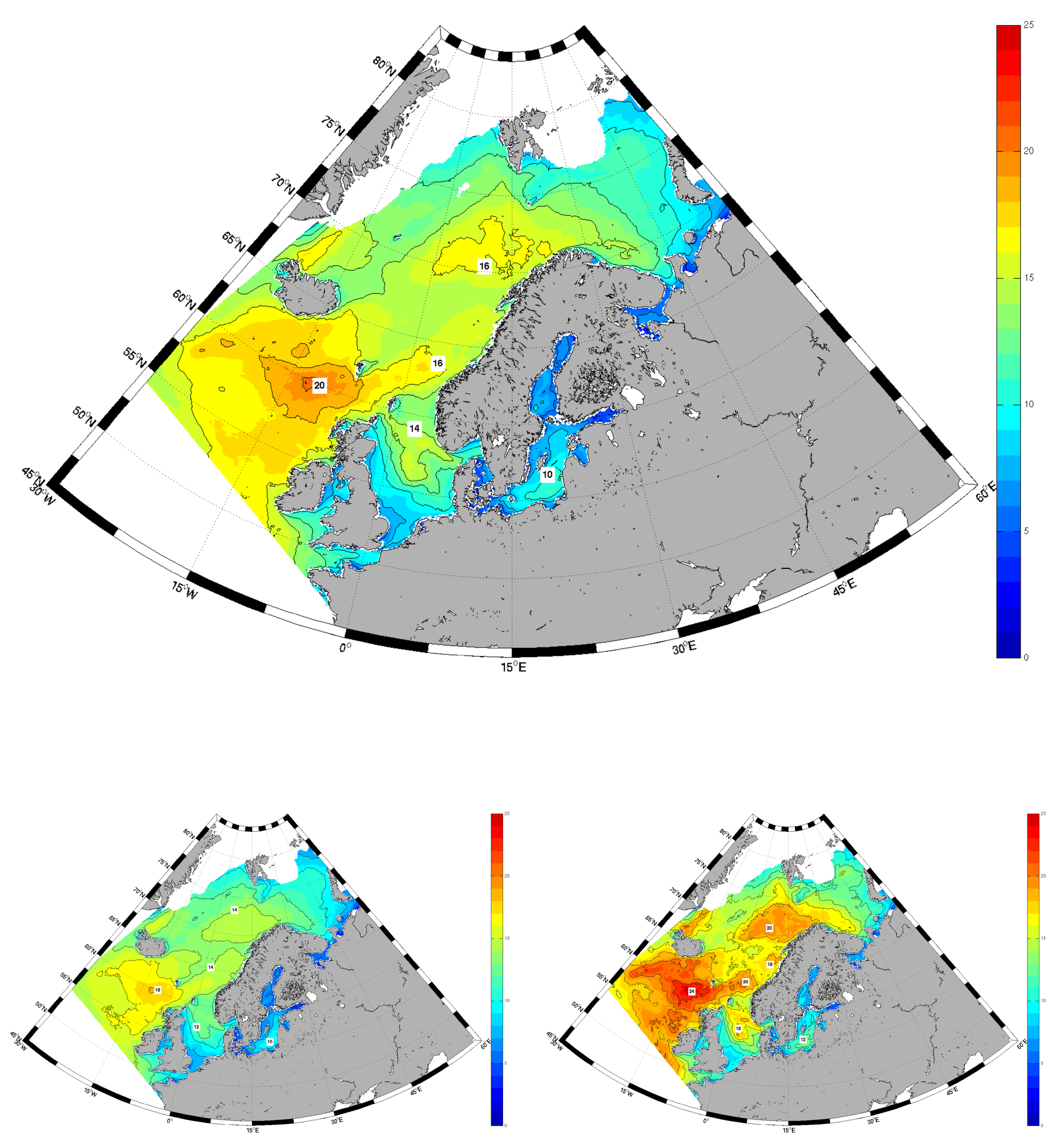

Figure 5: Top: best estimate.
Bottom: 5% and 95% confidence limits.

POT/GP

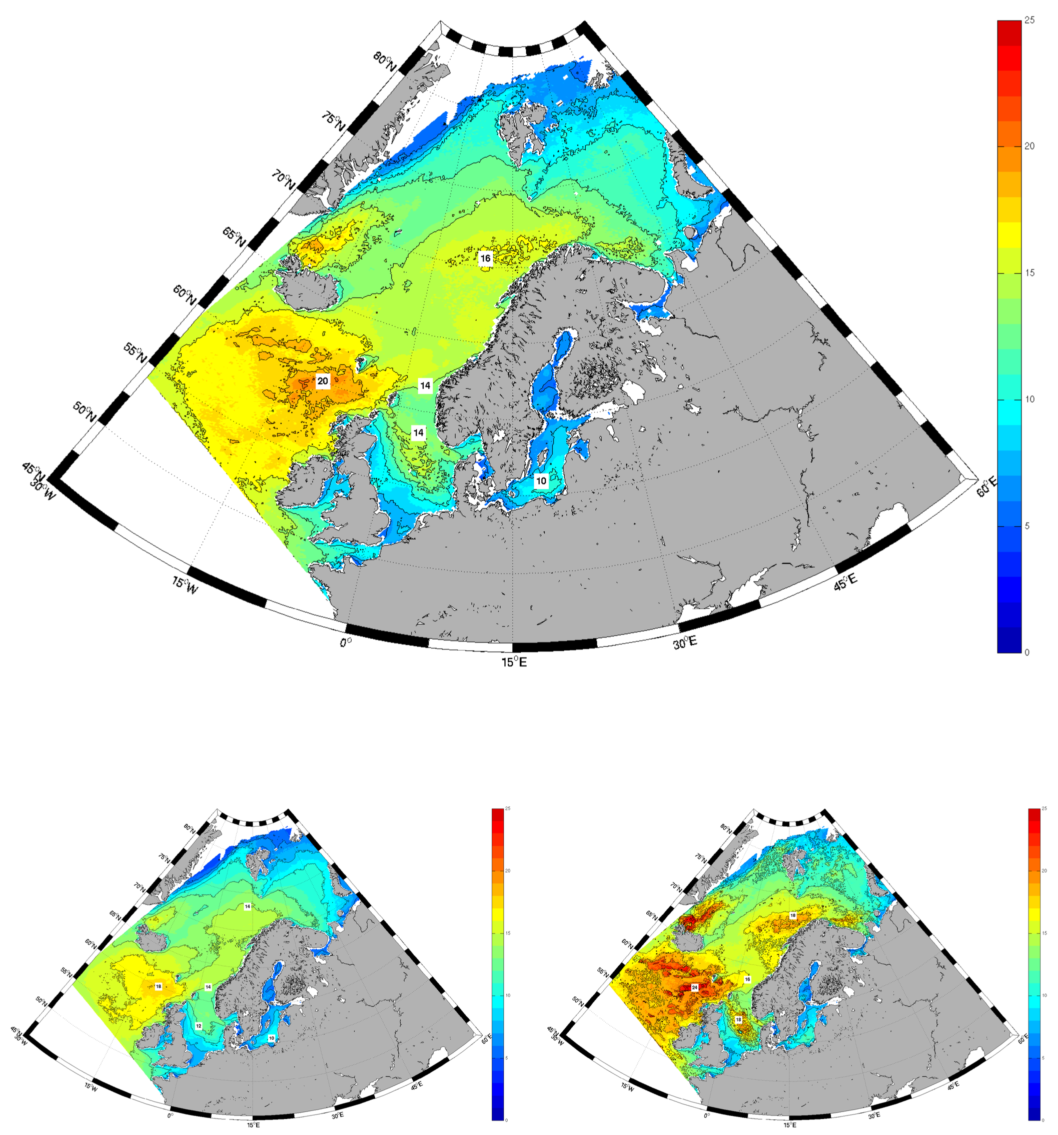

Figure 6: Top: best estimate.
Bottom: 5% and 95% confidence limits.