EFFECTS OF DISTRIBUTIONS AND FITTING TECHNIQUES ON EXTREME VALUE ANALYSIS OF MODELLED WAVE HEIGHTS

[1]W.D. Hogg and [2]V.R. Swail

[1] Reach Consulting, Bath ON
[2]Climate Research Branch, Downsview, ON

1. Introduction

Over the last two decades, there has been much discussion about the best distribution and most appropriate fitting technique to use in the analysis of extreme wave height probabilities. Investigation of the practical impact of such decisions has been hampered by the dearth of high quality, long-duration records of wave height measurements. In general, measurements in a single location have been too short to permit assessment of whether variability in predicted probabilities can be attributed to the distribution or fitting technique selected or are solely due to the small sample size. The availability of forty-one years worth of modelled wave heights in the AES-40 dataset provides one of the first opportunities to analyze wave fields that are close approximations to homogeneous, stationary, large samples, thus isolating the effects due to distribution and fitting technique selection.

The goal of this analysis was to examine the variability in wave height statistics introduced by the selection and use of various distributions and fitting techniques.

2. Data

The AES-40 data set used in this study is described in detail by Swail and Cox (2000). A brief summary description is provided in the following paragraphs.

The AES-40 hindcast wind and wave data base was produced for the Meteorological Service of Canada (formerly the Atmospheric Environment Service) by Oceanweather, Inc. Wind fields were generated on a fine mesh grid 0.625° latitude by 0.833° longitude over the entire North Atlantic Ocean from 20°E to 80°W, and from 0° to 76°N, every 6 hours for the period 1958-1998. Considerable attention was paid to the wind field generation, including overlaying detailed wind field analysis for more than 400 tropical storms, from the National Hurricane Center reconnaissance data, adjustment of all wind observations to a common reference height of 10 m, and intensive manual kinematic analysis of all storm wind fields. These high quality wind fields were then input to a proven 3rd generation spectral wave model (ODGP-3G). The resulting wave fields have been rigorously evaluated against both *in situ* wave measurements and satellite altimeter wave data, and have been shown to represent the wave conditions very well, including even the largest values which form the basis for extreme value analysis (Berek et al., 2000). Most AES-40 wind and wave validation statistics, and a wide range of climate analyses, including extremal analysis, are available on the project web site (http://www.oceanweather.com/aes40).

Forty-one years of modelled wave heights for four points in the North Atlantic were analyzed in this study. As shown in Figure 1, two of the points are in the western Atlantic (south of Sable Island, and Hibernia) and two are in the Northeastern Atlantic.

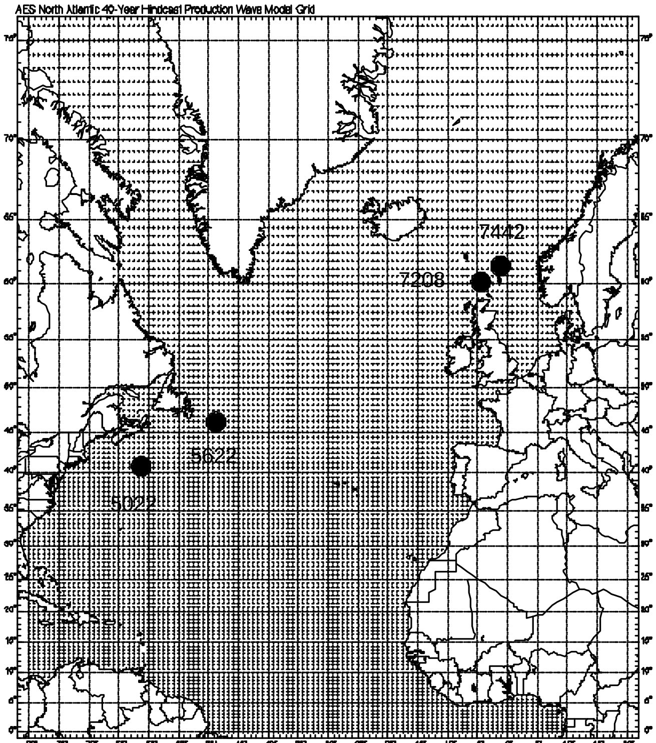

Figure 1. Grid Point locations.

3. Methodology

Frequency analyses were performed on the maximum wave height data using the method of L-moments (Hosking, 1990) to fit Gumbel, GEV and Pareto distributions to "wave year" maxima and, in the case of Pareto, to a POT sample where the threshold was chosen to generate approximately 41 points, i.e. the same as the number of years of record. For comparison purposes, samples were also fit to the Gumbel distribution using conventional moments. The sampling period "wave-year" has been defined as July-June to ensure that only one maximum per storm season is selected. The differences introduced by using a wave-year instead of a calendar year for selecting annual maxima were investigated.

To examine the effect of sample size, 10-year samples throughout each record were analyzed separately and then additional frequency analyses were performed on samples which expanded one year at a time from an initial ten years to the full forty-one years of record. Finally, statistics for wave measurements at the Magnus site, which is in close proximity to one of the model grid points, were compared with the model statistics.

4. Results

Figure 2 shows time series for the points along with sample calculations of 100-year return period estimates. The diamonds are the "wave-year" annual maximum wave heights. Various estimates of the 100-year return period wave heights are shown on the graphs, including estimates based upon various windows of data of 10-years duration.

In Figure 2, the dashed line depicts the period of record estimate of the 100-year return period wave height obtained assuming a GEV distribution fit using L-moments. As will be shown later, this was marginally the "best fit" to the 41-year datasets for all four locations. The square points are 100-year estimates assuming Gumbel, fit to a moving window 10-years wide using standard method of moments fitting. The open circles and solid circles are also 100-year estimates for overlapping 10-year windows five years apart for Pareto and GEV distributions respectively using L-moments fitting. The Pareto distribution was fit to a POT sample selected using a threshold that generated approximately the same number of points as there were years of record in the sample at each location. Some of the things to note:

1. The difference due to distribution assumed or fitting technique is an order of magnitude less than the difference due to 10-year sample. The same statement is true to a lesser extent when 20-year samples are used (not shown).
2. GEV fits the sample better (discussed later) but is not any more stable at predicting the 100-yr event from different samples. Pareto with POT is no better than annual maxima analyses.
3. Trend in the 100-year values is not similar to trend in annual extremes. This is due to the dependence of the return period estimate on sample standard deviation (and skew for GEV and Pareto) as well as on the mean extreme.

Frequency Analysis of Extreme Wave Heights

Frequency analyses of various samples assuming Normal, Gumbel, General Extreme Value (GEV) and Pareto distributions and using L-moment fitting were carried out. Samples came from calendar year annual maxima, wave-year (July-June) annual maxima and peak-over-threshold samples selected by choosing a threshold that generated approximately the same number of data points as there were years of record. Events were assumed independent when separated in time by more than 4 days. Samples were also created by selecting daily maxima from the 3-minute model output.

Results for all four locations are shown in Table 1. We first examine the impact of choosing the annual maxima from the calendar year as opposed to the “wave year”. This concept is often used in hydrology to avoid selecting flow values twice from the same flood that happens to be spread over the December-January period. Because the storm season in the Northern Hemisphere is during the winter and since persistence of storm tracks and other climate factors can introduce memory to the wave height data, it seems reasonable to ensure that wave heights from the same storm season are not selected in the annual maximum series by separating sampling period in the summer. Focusing on the 100-year estimates from the entire 41-year record, we find that differences due to “wave year” can be as much as 0.5 m (Gumbel-AE5022) but are generally only about 0.1 m. Differences for smaller samples were greater.

We next want to consider the effect on the return period estimate of the distribution assumed. For Table 1, all samples were fit using L-moments, so differences due to fitting technique have been eliminated. The symmetric Normal distribution was included as a sort of lower bound or “worst case” fit to the known-to-be-skewed extreme wave height data. Any distribution generating long-return-period estimates similar to the Normal distribution was assumed to be a poor fit. All distributions were fit to both annual maxima and peak-over-threshold (POT) data even though the Pareto distribution was developed for POT.

Note the different return periods for the daily maxima series.

Table 1(a) AE5022

AE5022	Cal	Year			Wave	Year			POT	9.5m	43 pts	
R.P.	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par
10	12.0	12.1	12.1	12.3	11.8	11.9	11.9	12.1	11.7	11.6	11.7	11.8
20	12.6	12.9	13.1	12.8	12.3	12.7	12.7	12.7	12.0	12.5	12.3	12.7
100	13.7	14.6	15.1	13.5	13.3	14.5	14.6	13.5	12.7	15.8	13.7	15.4
1000	14.9	16.6	18.1	13.9	14.3	16.9	17.2	14.0	13.5	25.2	15.5	20.9
10000	15.8	18.3	21.1	13.9	15.2	19.3	19.8	14.2	14.1	46.4	17.4	29.4
	Daily	Max							POT	9.0m	61 pts	
R.P.	Norm	GEV	Gum	Par				R.P.	Norm	GEV	Gum	Par
2	7.2	12.7	10.0	9.0				10	11.2	11.2	11.2	11.3
5	7.6	15.0	11.1	9.5				20	11.5	11.9	11.8	12.1
10	7.9	16.9	12.0	9.8				100	12.1	14.4	13.0	13.9
20	8.2	19.0	12.8	10.1				1000	12.8	20.5	14.7	17.0
100	8.8	24.5	14.7	10.6				10000	13.4	31.9	16.4	20.4

Table 1 (b) AE5622

AE5622	Cal	Year			Wave	Year			POT	10.3m		
R.P.	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par
10	12.8	12.9	12.9	13.0	12.8	12.9	13.0	12.9	12.6	12.7	12.7	12.9
20	13.4	13.5	13.8	13.3	13.4	13.4	13.9	13.2	13.0	13.4	13.3	13.5
100	14.4	14.6	15.9	13.7	14.5	14.2	15.9	13.4	13.7	15.2	14.6	14.5
1000	15.6	15.6	18.8	13.8	15.6	14.8	18.8	13.4	14.4	18.3	16.5	15.5
10000	16.6	16.2	21.8	13.8	16.6	15.2	21.7	13.4	15.1	22.2	18.3	16.0
	Daily	Max										
R.P.	Norm	GEV	Gum	Par								
2	8.0		11.0	8.3								
5	8.4		12.2	8.5								
10	8.7		13.1	8.6								
20	9.0		14.0	8.7								
100	9.6		16.2	8.9								

Table 1 (c) AE7208

AE7208	Cal	Year			Wave	Year			POT	11.2m		
R.P.	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par
10	13.8	13.8	13.9	13.9	13.7	13.8	13.8	13.9	13.6	13.7	13.7	13.8
20	14.3	14.4	14.8	14.3	14.2	14.4	14.7	14.2	14.0	14.4	14.3	14.4
100	15.4	15.5	16.9	14.6	15.3	15.6	16.8	14.6	14.7	16.1	15.7	15.4
1000	16.6	16.6	19.8	14.7	16.5	16.7	19.8	14.7	15.5	19.0	17.7	16.2
10000	17.6	17.2	22.7	14.7	17.5	17.4	22.7	14.7	16.2	22.4	19.7	16.7
	Daily	Max										
R.P.	Norm	GEV	Gum	Par								
2	8.8	14.7	12.3	10.4								
5	9.3	17.0	13.6	10.8								
10	9.7	18.9	14.6	11.1								
20	10.0	20.9	15.7	11.3								
100	10.7	26.2	18.0	11.7								

Table 1 (d) AE7442

AE7422	Cal	Year			Wave	Year			POT	11.2m		
R.P.	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par	Norm	GEV	Gum	Par
10	13.1	13.2	13.2	13.4	13.0	13.1	13.1	13.3	12.8	12.7	12.9	12.9
20	13.7	14.1	14.2	14.1	13.6	14.0	14.0	14.0	13.3	13.9	13.7	14.1
100	14.8	16.1	16.3	15.0	14.7	16.0	16.2	14.9	14.1	18.2	15.3	17.8
1000	16.0	18.7	19.4	15.5	15.9	18.6	19.3	15.4	15.1	31.2	17.6	25.9
10000	17.1	21.2	22.5	15.7	16.9	21.0	22.3	15.6	15.9	62.6	20.0	39.4

	Daily	Max		
R.P.	Norm	GEV	Gum	Par
2	8.5	14.0	11.7	9.9
5	8.9	16.2	13.0	10.3
10	9.2	18.0	14.0	10.6
20	9.6	20.0	14.9	10.8
100	10.2	24.9	17.2	11.2

Some points to note:

1. Differences between the calendar year and wave year annual maxima can result in differences of 0.5 m in the 100-yr value.
2. Pareto is always inappropriate (as low or lower than Normal) for analyzing annual maxima series.
3. Values estimated using Gumbel and GEV with annual maxima series and using Pareto with the appropriate POT series are generally similar up to the 100-yr return period.
4. POT techniques, including using Pareto, were very sensitive to the threshold selected as demonstrated in Table 1. Lowering the threshold selection by 0.5 m increased the number of points by 18 and decreased the Pareto 100-year estimate by 1.5 m.
5. Surprisingly, the 2-parameter Gumbel distribution even does a reasonable job of estimating the 100-yr value from daily maximum data but the other distributions do not.

Goodness of Fit

To aid in evaluating the L-moment fit of the various distributions and sampling techniques, the distributions and extreme wave height data were plotted on log/linear graphs. The observations were plotted using both the Gringorten plotting position formula ((m-.44)/(N+.12)) and the Hosking plotting position formula ((m-.35)/N). The different plotting position shifts the points a bit but does not make an appreciable difference to the apparent quality of the fit. Only the observations plotted according to the Hosking formula are included here. To simplify interpretation, only the curves for wave-year annual maximum GEV and Gumbel distributions and POT Pareto along with observations for return periods > 10-years are shown.

Things to note:

1. All probabilities are similar out to the 100-year value and generally speaking do a reasonable job of fitting the observations as plotted.
2. Gumbel is the worst fit for two locations and Pareto for the other two.
3. GEV fits the sample points best for all locations.
4. Referring back to the time series plots, this does not make GEV a superior predictor of low probability events because of the high variability from sample to sample.
5. The slope of the Gumbel distribution is similar for all locations.
6. A computational note: the 3-parameter distributions (GEV & Pareto) did not always converge to a reasonable answer unless a first guess plotting position was input to the L-moment software

supplied by IBM (Hosking, 1996). There was no error message, just disconcerting results. This was the only problem we had with what is otherwise an easy package to use.

Expanding Window Analysis

As previously discussed, the ability to closely fit small samples of environmental data is not necessarily a good measure for evaluating the suitability of a distribution to predict the "true" 100-year return period value. As shown in Figure 2, a series of small samples of wave maxima generate quite different estimates of the 100-year value and the closer the distribution fits the individual points in the sample, the greater the noise due to sampling in the resultant 100-year estimates. Since for small samples (< about 35 years), the error due to sampling is larger than all other sources of error including distribution assumed and fitting technique applied, it is imperative to select a distribution which minimizes the effect of sampling error when small samples are used. Given a well-behaved and large enough sample, the 100-year estimate is relatively insensitive to distribution assumed or fitting technique applied. The true practical measure of the worthiness of a distribution and fitting technique should be how quickly the correct return period estimate is reached as sample size goes from small to large.

Figure 4 is an attempt to measure how quickly the different distributions converge to the long-record 100-yr wave height. For each of the four AE model points the 100-yr height was computed using windows of consecutive years ranging in length from 10 to the full period 41 years. Gumbel and GEV L-moments, and the Gumbel moments with annual maxima and Pareto L-moments using POT derived from the same expanding windows as the annual maxima series, are shown. The absolute difference between the 100-yr value for each distribution at each AE point and its corresponding 100-yr value computed using all 41 years of data and the same distribution are plotted. The best distribution would be the one that converges on the long-record value most quickly and should have the lowest average differences (shown in the legend). The graphs for each point were produced but only the graph showing the average absolute differences for all 4 points is shown here. The Gumbel moments and Gumbel L-moments have the lowest differences, probably because the 2-parameter distribution is less sensitive to differences in individual small samples and is more stable as more data are added. Pareto is worst.

Analysis of Measured Wave Heights

Observations of RMS wave height at Magnus were also examined. The time-series for wave-year maxima are shown Figure 5. The apparent trend (almost .7 m/yr) is due to 3 low years at the start of the short sample followed by a step increase of 2-3 m in the average maximum. Extrapolation of the linear increase is certainly not justified but it would be interesting to investigate the step change in 1988. The wave-year maxima for the same period from the hindcast model results for the grid point closest to the Magnus location (AE7208) are also shown in Figure 5. The same step change and trend are not evident in the hindcast results.

Extreme value analysis was also carried out on the Magnus observations and some of the results are shown in Figure 6. Note that the Gumbel estimates seem very high and that all estimates are much higher than the model results for the grid point closest to the Magnus location, with 100-yr values around 20 m as compared to model predictions closer to 15 m. This is because both the mean and especially the standard deviation of the observed data are higher than the model results, not only for this 8-year period at the Magnus grid point but for hindcast results at all four grid points examined. Standard deviation essentially controls the slope of the return period prediction line and hence a small difference in it will generate large differences in predicted return periods. The standard deviation for modeled waves was typically around 1.5 m while for this short 8-year record of observed data the standard deviation was almost 2.2 m. The difference may be partially explained by the specific period of observation.

5. Conclusions.

Based upon this examination of the statistics of modeled wave heights we conclude:

1. Variability in the estimate of the 100-year return period value due to sample size is large compared to variability due to either distribution assumed or fitting technique employed.
2. If there are more than about 35 years of data, all distributions provide a good and stable estimate of the 100-year return period value.
3. GEV fit by L-moments provided the closest fit to sample points but that does not necessarily mean that it is the most appropriate distribution to use.
4. L-moment fitting may provide a closer fit to observations in small samples but this does not translate into faster convergence to the long-record probability estimates.
5. Pareto is inappropriate for annual maximum series and with POT datasets, exhibits all of the shortcomings of other distributions.
6. POT techniques are sensitive to the threshold selected.
7. Differences caused by inappropriately using a calendar year for selecting annual maxima instead of a year that breaks in the season of low winds can cause differences in the 100-year return period height of 0.5 m.
8. Even though the differences between measured and modeled extreme values in any one year is not large for the Magnus site, the differences in the predicted 100-year return period wave height is large, mainly because of the difference in standard deviation between the measured and modeled data.

At least for this dataset of modelled wave heights, if more than 35 years of data are used then all extreme value distributions and fitting techniques tested seem to do an equally adequate job and it doesn't matter which ones you use. For small samples, the 2-parameter distribution (Gumbel) fit by either standard moments or L-moments, makes smaller errors in comparison to the long-record estimate of 100-year return period than do the 3-parameter distributions. If it is reasonable that the goal of distribution fitting is to reach an estimate as close and as quickly as possible to the long-record estimate, then it follows that we should use the 2-parameter distribution.

6. References.

Berek, E.P., V.J. Cardone and V.R. Swail, 2000. Comparison of hindcast results and extreme value estimates for wave conditions in the Hibernia area – Grand Banks of Newfoundland. Proc. 6th International Workshop on wave Hindcasting and forecasting, Monterey, CA, 6-10, November, 2002, p. 250-260.

Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society, Series B, 52, 105-124.

Hosking, J. R. M. (1996). Fortran routines for use with the method of L-moments, Version 3. IBM Research Report RC20525.

Swail, V. R. and A. T. Cox, 2000. On the use of NCEP-NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *J. Atmospheric Oceanic Technology*, **17**, 532-545.

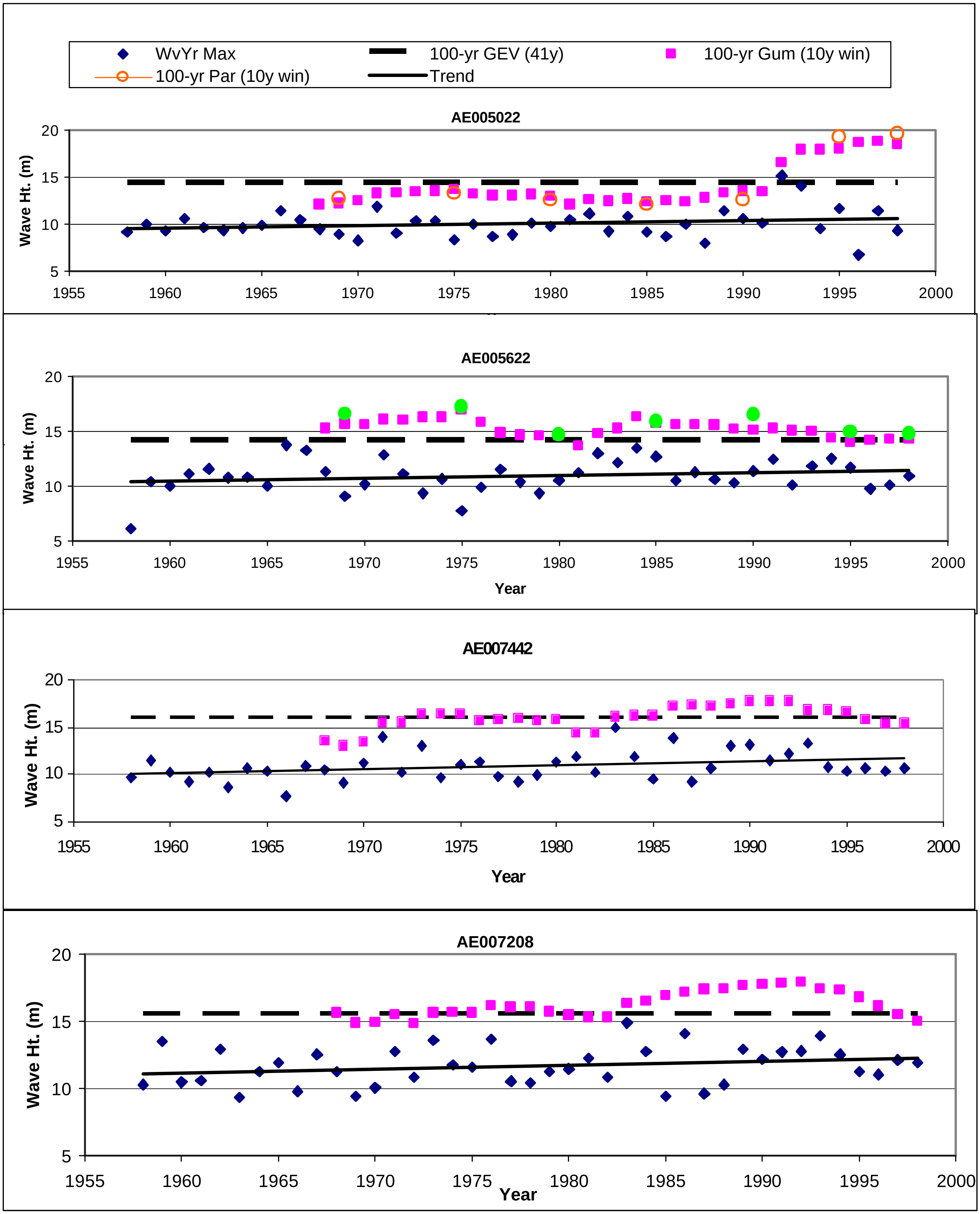

Figure 2. Wave Year Maxima for 4 locations. For AE5022 the open circles represent Pareto, whereas for AE5622 the solid circles represent the GEV.

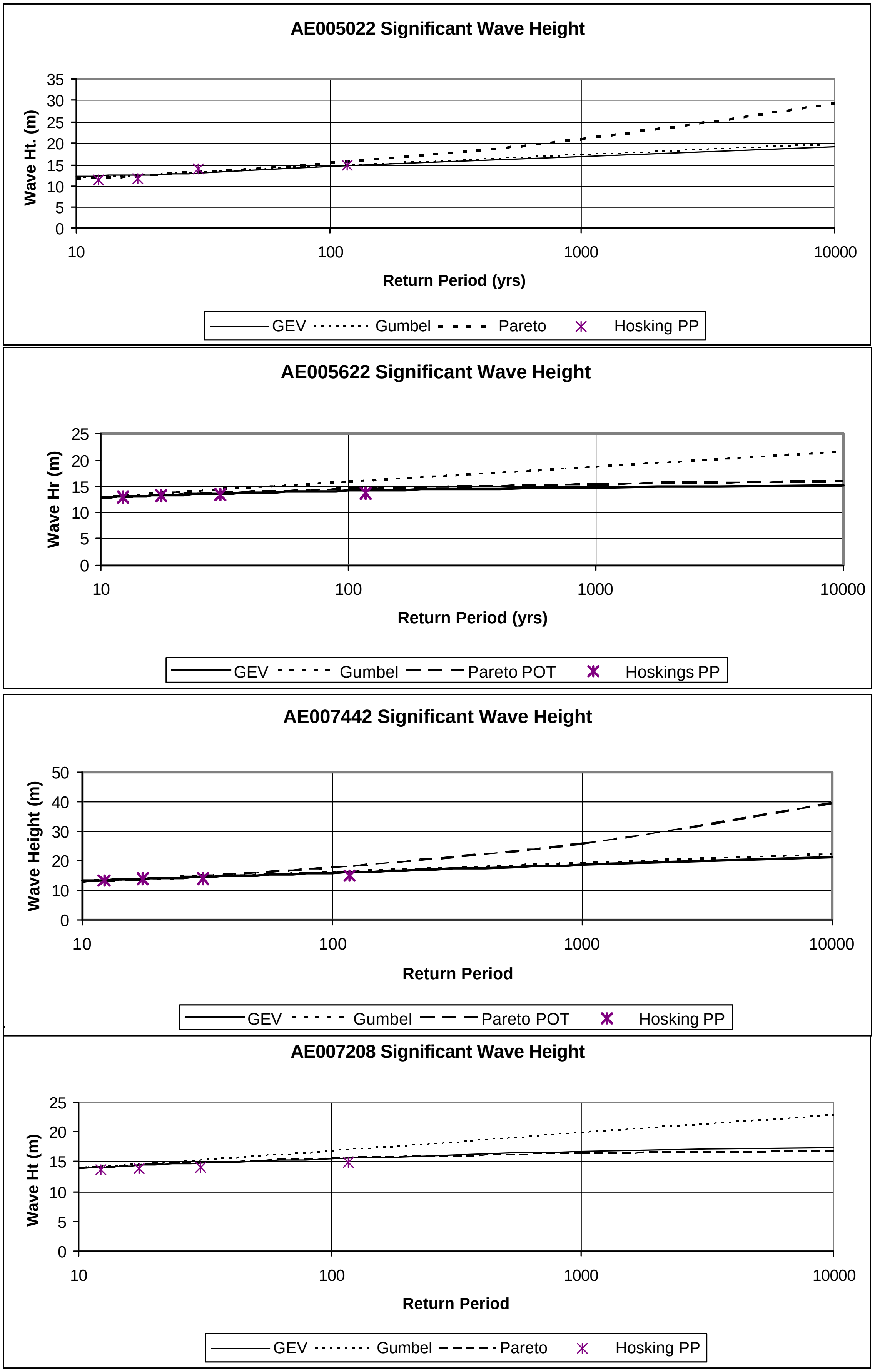

Figure 3. Extreme value analysis using L-moments for different distributions.

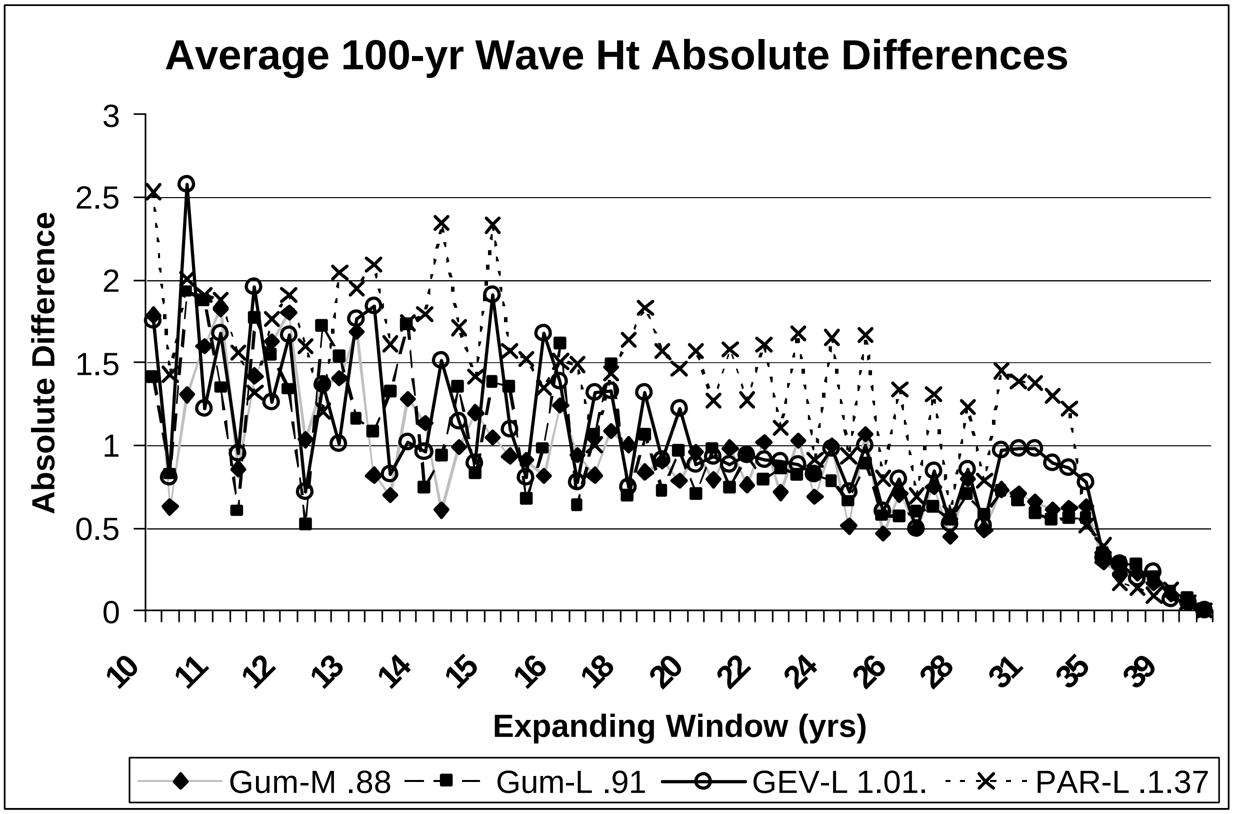

Figure 4. Absolute difference (m) between 100-year wave height computed from 41-year sample and that computed from the expanding window ranging from 10-40 years.

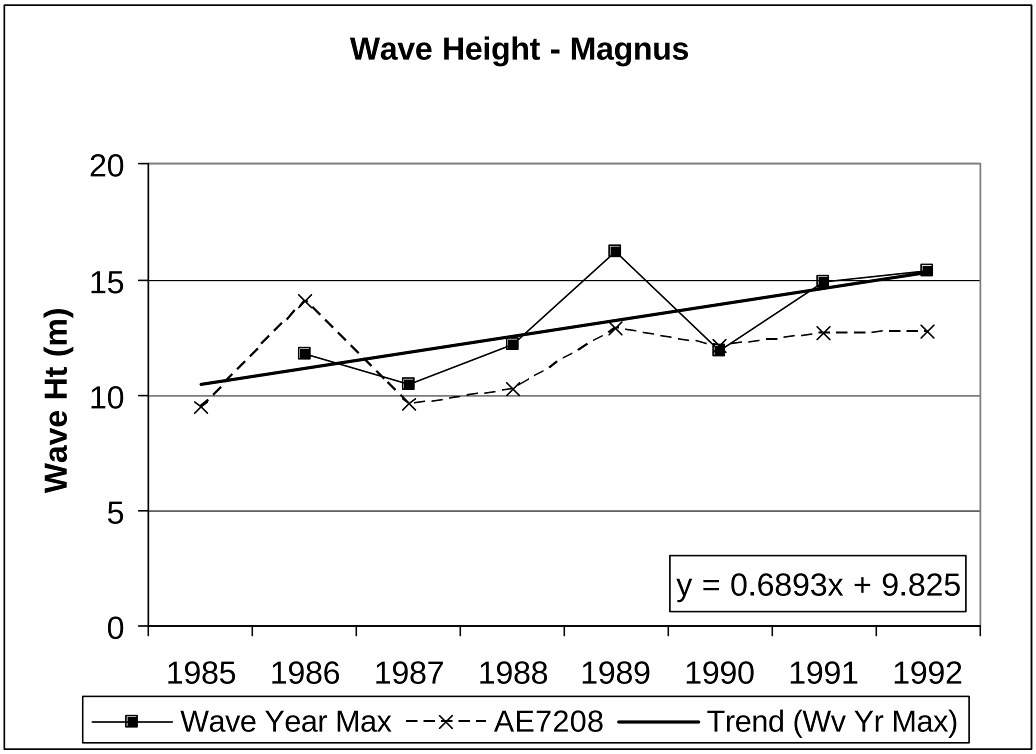

Figure 5. Time series for wave-year maxima at Magnus.

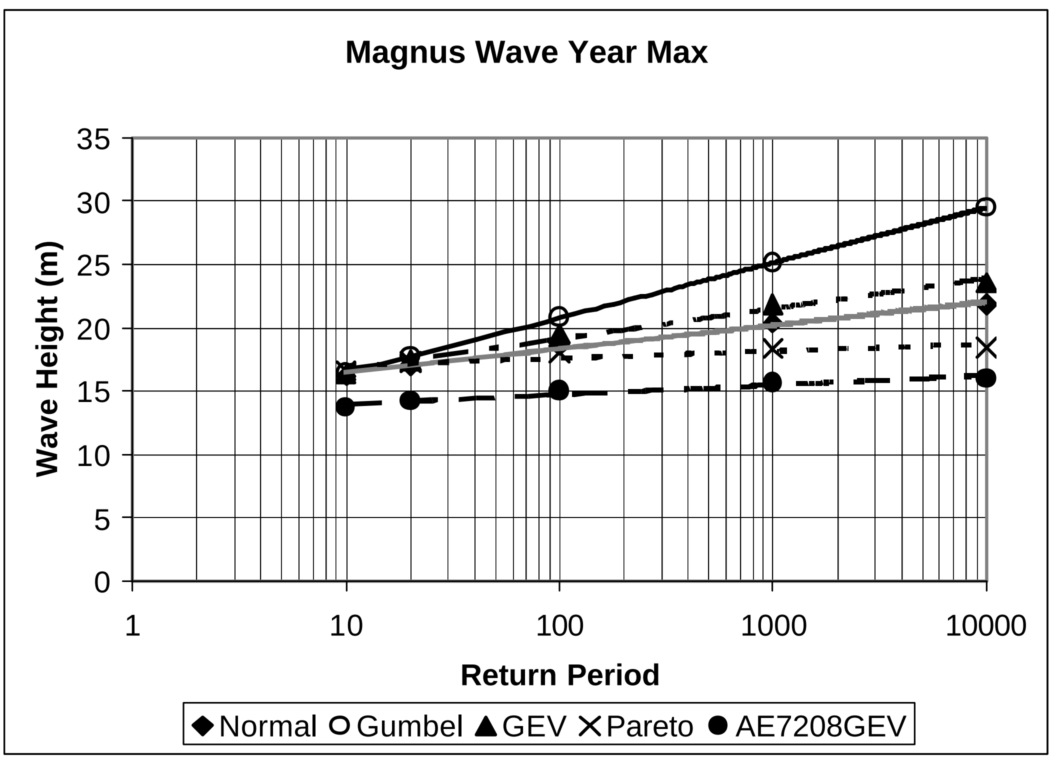

Figure 6. Extreme value analyses at Magnus based on wave-year maxima.