Prepared for:
National Institute for Coastal and
Marine Management (RIKZ)

Selection of storm events and estimation of exceedance frequencies of significant wave height for five North Sea locations

AFGEHANDELD

April 1994

delft hydraulics
Selection of storm events and estimation of exceedance frequencies of significant wave height for five North Sea locations

C.F. de Valk

delft hydraulics
Executive’s summary

1 Introduction

The objective of the HYDRA project of the National Institute for Coastal and Marine Management (RIKZ) is to determine the joint statistics of waves and sea levels along the coast of the Netherlands which are needed for assessment of the risk of flooding during extreme storms. The approach adopted in HYDRA is to start with determining the appropriate statistics at the 20 m depth contour based on measurements and hindcast data. Numerical models are used to transform these statistics toward the shore and into the estuaries.

The present study is the first phase of the estimation of statistics near the 20 m depth contour. It follows a number of pilot studies which were carried out to determine the methodology. This reports describes the selection of storm events and the corresponding measurements of several wave, wind and sea level parameters at five offshore locations. Also, estimates of omnidirectional exceedance frequencies of the significant wave height $H_{m0}$ at these five locations are presented.

2 Selection of storm events

Available are 3-hourly measurements from January 1, 1979 to December 31, 1991 at the five North Sea locations SON (Schiemondkooog Noord), ELD (Eelde), YM6 (Ymuiden), EUR (Europlatform/Lichteiland Goeree) and K13 (see Figure 1) of the parameters:

- significant wave height $H_{m0}$,
- mean wave period $T_{m0}$,
- mean wave direction $\theta_0$,
- wind speed and direction,
- sea level and astronomical tide, and instantaneous set-up (the difference between observed sea level and astronomical tide).

Storm events are selected based on exceedances of thresholds of the parameters

- significant wave height $H_{m0}$,
- instantaneous set-up,
- wind speed.

For each parameter and location, the threshold is determined as the value which is exceeded during a fixed fraction $f$ of the time. However, $f$ is the same for all parameters and locations.

Next, storm events are selected jointly for all five locations as the intervals over which at each instant, at least one parameter at one location exceeds its threshold. In other words, a storm event will be considered ended only if all parameters at all locations are below their thresholds; the beginning of a storm event is defined analogously. This procedure is illustrated for the case of just two parameters in Figure 2.
The selection was carried out with the values 0.025 and 0.05 for the fraction \( f \) of data retained per parameter and location. The value 0.05 was chosen because it left a sufficient number of storm events in which \( H_{m_0} \) exceeds the storm selection threshold at the location EUR.

Several checks of the validity of the selected storm events were made. The lengths of the storm events were examined; a few very long storm events were found. Selection of storm events independently for each location reduces the storm lengths but has no effect on the highest storm maxima of \( H_{m_0} \), and also very little effect on the estimates of exceedance frequencies.

Clustering of storm events in which the significant wave height exceeds a high level must be absent in order to be able to apply the exceedance frequency estimates to risk assessment. Plots of the sample distributions of the intervals between such storm events indicate that this condition is satisfied.

It was found that restriction of the storm events to the storm season of October 1 to March 31 makes little difference in the values of the high storm maxima of \( H_{m_0} \).

3 Estimates of omnidirectional exceedance frequencies of significant wave height

Exceedance frequencies of significant wave height were estimated from the maxima during the storm events selected jointly for all locations as described above, using a fraction \( f \) of data retained of 0.05 and restricting storm events to the storm season October 1 to March 31.

Exceedance frequencies of \( H_{m_0} \) are defined as the mean number of storm events per year during which \( H_{m_0} \) exceeds a particular value. Estimates are derived from the estimate of the distribution function \( F^\omega \) of the storm maxima of \( H_{m_0} \) which exceed a threshold \( \omega \). The exceedance frequency of the value \( a \) of \( H_{m_0} \) is estimated as the product of \( 1 - F^\omega(a) \) and the observed mean number of storm maxima per year exceeding \( \omega \). The distribution \( F^\omega \) of the storm maxima of \( H_{m_0} \) exceeding the threshold \( \omega \) is assumed to be a conditional two-parameter Weibull distribution:

\[
F^\omega(a) = 1 - e^{-(a/\sigma)^\alpha} \cdot (a/\sigma)^\alpha
\]

In this expression, \( \sigma \) is a scale parameter and \( \alpha \) is a shape parameter. The parameter \( \alpha \) determines the curvature of the logarithm of \( 1 - F^\omega \). The case of \( \alpha = 1 \) corresponds to an exponential distribution. In this case, the probability of exceedance, plotted on a semilogarithmic scale, is a straight line with slope equal to \(-1/\sigma \). For a range of values of the threshold \( \omega \), the parameters \( \alpha \) and \( \sigma \) of the distribution function \( F^\omega \) are fitted by the maximum likelihood method. Examples of fitted exceedance frequencies for the five locations of wave measurements are shown in Figure 15. Estimates of the parameters \( \alpha \) and \( \sigma \) as well as of the value of \( H_{m_0} \) corresponding to an exceedance frequency of \( 10^{-4} \) per year are plotted in Figures 11-14 as a function of the number of samples exceeding \( \omega \).
The variation in the estimates of the $10^{-4}$ per year significant wave height with the number of samples is considerable. For most locations, the fluctuation is fairly stable within a band of one to three metres. The estimates for EUR show the smallest fluctuation. The estimates for ELD show the largest variation, associated with a trend extending over a wide range of thresholds.

The estimates of the $10^{-4}$ per year significant wave height as a function of threshold derived from the maximum likelihood estimates of the Weibull parameters were subjectively interpreted, resulting in the following guesses:

<table>
<thead>
<tr>
<th>Location</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>about 7.5 m</td>
</tr>
<tr>
<td>YM6</td>
<td>about 9.0 m</td>
</tr>
<tr>
<td>ELD</td>
<td>9.0-11.0 m</td>
</tr>
<tr>
<td>SON</td>
<td>about 10.5 m</td>
</tr>
<tr>
<td>K13</td>
<td>about 9.0 m</td>
</tr>
</tbody>
</table>

A modification of the estimation method was tested in which for various fixed values of $\alpha$, the scale parameter $\sigma$ was estimated for a range of thresholds. Results are shown in Figure 17. With $\alpha$ fixed, the variability of the estimate of the $10^{-4}$ per year $H_{m_0}$ is reduced considerably. A value of $\alpha$ can be chosen for which the trend vanishes in the estimates of $\sigma$ as a function of threshold. Also other criteria can be taken into account, such as the likelihood of the data for the given value of $\alpha$, conservatism, and the spatial coherence of $\alpha$. The procedure is similar to the two-stage approach suggested in (De Valk, 1993; Section 5.2) and is perfectly consistent with the estimator for exceedance frequencies of multivariate extremes in (De Valk and Zitman, 1992; Chapter 4).

The estimates of the $10^{-4}$ per year significant wave height as a function of threshold for various fixed values of $\alpha$ were subjectively interpreted, resulting in the following guesses:

<table>
<thead>
<tr>
<th>Location</th>
<th>$10^{-4}$ per year $H_{m_0}$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>about 8 m</td>
<td>2.5-3.0</td>
</tr>
<tr>
<td>YM6</td>
<td>9.5-10 m</td>
<td>2.0-2.5</td>
</tr>
<tr>
<td>ELD</td>
<td>9.5-10 m</td>
<td>2.0-3.0</td>
</tr>
<tr>
<td>SON</td>
<td>10.5-11 m</td>
<td>2.0</td>
</tr>
<tr>
<td>K13</td>
<td>about 10 m</td>
<td>2.0</td>
</tr>
</tbody>
</table>

(intervals do not represent error bands). These values are generally higher than those based on the maximum likelihood estimates of both $\alpha$ and $\sigma$.

This stepwise method for estimating the $10^{-4}$ per year value of $H_{m_0}$ seems viable both from a theoretical point of view and based on the results produced.
4 Conclusions and recommendations

1 The storm event selection proposed in section 2.1 appears satisfactory. Joint selection of storm events for all parameters and locations results in a few very long storm events, but the effect on the selected storm maxima of significant wave height is negligible.

2 Also the estimates of the $10^{-4}$ per year significant wave height are hardly affected by the choice between storm event selection per location and joint selection of storm events for all locations.

3 Clustering of storm events in which the significant wave height exceeds a high value should vanish in order to be able to apply the exceedance frequency estimates to risk assessment. Plots of the sample distributions of the intervals between such storm events indicate that this condition is satisfied.

4 Restriction of the storm events to the storm season of October 1 to March 31 makes little difference to the values of the high storm maxima.

5 Two-parameter Weibull fits to the storm maxima of significant wave height show a considerable dependence of the estimate of the $10^{-4}$ per year significant wave height on threshold. As a consequence, it is difficult to select a final estimate from these results.

6 Generally, the estimates of the $10^{-4}$ per year significant wave height increase along the coast from EUR to SON.

7 The variation in the estimates of $10^{-4}$ per year significant wave height with threshold is considerably reduced when the curvature parameter $\alpha$ is kept fixed; then only the estimate of the scale parameter $\sigma$ varies with the threshold. These estimates can be repeated for different values of $\alpha$ and a value of $\alpha$ can be chosen for which the trend in the estimates of $\sigma$ as a function of threshold vanishes. Also other criteria can be taken into account, such as the likelihood of the data for the given value of $\alpha$, conservatism, and the spatial coherence of $\alpha$. In this study, guesses based on these estimates appear to be higher than those based on maximum likelihood estimates of both parameters.
Contents

List of figures

1 Introduction .................................................. 1

2 Selection of storm events .................................. 2
   2.1 Method .................................................. 2
   2.2 Results .................................................. 5

3 Estimates of omnidirectional exceedance frequencies of significant wave height ........................................ 9
   3.1 Method .................................................. 9
   3.2 Two-parameter Weibull estimates ...................... 10
   3.3 Sensitivity of estimates to data-selection method .... 11
   3.4 Estimates with a fixed shape parameter .............. 11

References

Figures

Appendix A: Considerations related to independence in data selection
Appendix B: Three ways to define directional storm maxima and directional exceedance frequencies
List of figures

1. Locations of measurements
2. (a) Illustration of the storm selection procedure
   (b) Illustration of the definition of directional storm maxima in Tables 2 and 4.
3. Sample distribution of storm duration for the case $f=0.025$
4. Sample distribution of storm duration for the case $f=0.05$
5. Sample exceedance frequencies of storm maxima of $H_{m_0}$ at EUR, YM6, ELD and SON
6. Sample distribution of intervals between storm events during which $H_{m_0}$ exceeds a
   level, plotted for increasing levels: (a) EUR; (b) YM6; (c) ELD; (d) SON; (e) K13
7. Sample exceedance frequencies of storm maxima of $H_{m_0}$ during the storm season at
   EUR, YM6, ELD and SON
8. Sample exceedance frequencies of storm maxima of $H_{m_0}$ during the storm season at
   (a) EUR; (b) YM6; (c) ELD; (d) SON; (e) K13
9. Sample distribution of storm duration for ELD with storms selected separately for each
   location
10. Sample exceedance frequencies of storm maxima of $H_{m_0}$ at EUR, YM6, ELD and SON
    during the storm season, with storms selected separately for each location
11. Estimates of the $10^{-4}$ per year $H_{m_0}$ from two-parameter Weibull fits as a function of
    the number of storm maxima for (a) EUR and YM6; (b) ELD and SON; (c) K13
12. Estimates of the Weibull parameter $\alpha$ as a function of the number of storm maxima
    for (a) EUR and YM6; (b) ELD and SON; (c) K13
13. Estimates of the Weibull parameter $\sigma$ as a function of the number of storm maxima
    for (a) EUR and YM6; (b) ELD and SON; (c) K13
14. Threshold $\omega$ as a function of the number of storm maxima for (a) EUR and YM6;
    (b) ELD and SON; (c) K13
15. Two-parameter Weibull fits to the 50 highest storm maxima of $H_{m_0}$ at (a) EUR
    and YM6; (b) ELD and SON; (c) K13
16. Estimates of the $10^{-4}$ per year $H_{m_0}$, Weibull parameters $\alpha$, $\sigma$ and threshold $\omega$ as a
    function of the number of storm maxima for ELD, with storm events selected per
    location
17. Estimates of the $10^{-4}$ per year $H_{m_0}$ as a function of the number of storm maxima for
    fixed values 1.0, 1.5,..., 3.0 of $\alpha$: (a) EUR and YM6; (b) ELD and SON; (c) K13
18. Fits to the 50 highest storm maxima of $H_{m_0}$ with selected values of $\alpha$ at (a) EUR
    and YM6; (b) ELD and SON; (c) K13
1 Introduction

The objective of the HYDRA project of the National Institute for Coastal and Marine Management (RIKZ) is to determine the joint statistics of waves and sea levels along the coast of the Netherlands which are needed for assessment of the risk of flooding during extreme storms. The approach adopted in HYDRA is to start with determining the appropriate statistics at the 20 m depth contour based on measurements and hindcast data, taking into account existing results for sea levels (Dillingh et al., 1993) and wind (Rijkooord and Wierenga, 1983). Numerical models are used to transform these statistics toward the shore and into the estuaries.

The present study is the first phase of the estimation of statistics near the 20 m depth contour. It follows a number of pilot studies which were carried out to determine the methodology; see e.g. (Roskam, 1988; Roskam, 1993; De Valk and Zitman, 1992; Zitman, 1993; De Valk, 1993; De Valk and Otta, 1993).

The second chapter of this report describes the selection of storm events and the corresponding measurements of several wave, wind and sea level parameters at five offshore locations. One of the objectives was to select storm events in such a way that the data can be used later for the estimation of joint exceedance frequencies of several parameters as proposed in (De Valk and Zitman, 1992). Several aspects of the storm event selection were investigated, such as the numbers of directional storm maxima of significant wave height $H_{m0}$ available, the distributions of the storm length and of the interval between successive storm events, and seasonal dependence.

Estimates of exceedance frequencies of $H_{m0}$ from the (omnidirectional) storm maxima of the five locations are presented in chapter 3. Fits were made of the tail of the two-parameter Weibull distribution to the storm maxima of $H_{m0}$. Apart from maximum likelihood estimation of the Weibull parameters also an alternative approach was applied. The evaluation of the results focused on the estimates of the value of $H_{m0}$ with an exceedance frequency of $10^{-4}$ per year.

This study was carried out by C.F. de Valk under the supervision of ir J.A. de Ronde and ir H. Keyser of RIKZ and in close cooperation with RIKZ.
2 Selection of storm events

2.1 Method

Available are 3-hourly measurements from January 1, 1979 to December 31, 1991 at the North Sea locations

SON (Schiermonnikoog Noord)
ELD (Eierland)
YM6 (Ymuiden)
EUR (Europlatform/Lichtelanda Goeree)
K13

(see Figure 1), of the parameters:

- significant wave height $H_{m0}$,
- mean wave period $T_{m0}$,
- mean wave direction $\theta_0$,
- wind speed and direction,
- sea level and astronomical tide, and instantaneous set-up (the difference between observed sea level and astronomical tide).

The data were collected, completed and validated by the National Institute for Coastal and Marine Management (RIKZ) of Rijkswaterstaat; see e.g. (Roskam, 1988). The effects of sampling with an interval of three hours were investigated in (Roskam, 1993).

Based on these data, storm events are selected, which are the distinct samples of data used to estimate frequencies of occurrence\(^1\) of extreme conditions such as high significant wave heights or sea levels. Without a precise definition, a frequency of occurrence of an extreme condition is a rather ambiguous concept. For example, due to a minor fluctuation, the sea level may exceed a certain value twice in a period of several hours. The question is whether these upcrossings should be considered as two distinct events, or as a single event. Because a small fluctuation of the sea level is insignificant to relevant parameters such as the probability of flooding in an arbitrary year or in a hundred year period, the latter option should be chosen or the effect of clustering of events should be removed in some other way. The statistical basis of selection of storm events is discussed in Appendix A.

In this study, storm events are selected based on exceedances of thresholds of the parameters

- significant wave height,
- instantaneous set-up,
- wind speed.

The thresholds are determined for each parameter and location separately as the value which is exceeded during a fixed fraction $f$ of the time that data of this parameter/location are available. However, $f$ is the same for all parameters and locations.

\(^1\) defined as the mean number of occurrences per year
Next, storm events are selected jointly for all five locations as the intervals over which at each instant, at least one parameter at one location exceeds its threshold. In other words, a storm event is considered ended only if all parameters at all locations are below their thresholds; the beginning of a storm event is defined analogously. The procedure is best illustrated for the case of just two parameters. Figure 2 shows time-series of these parameters together with their thresholds and illustrates how the threshold exceedances of the parameters are combined to storm events.

The approach sketched above is motivated by the following.

- The main reason for selecting storm events based on the values of the parameters significant wave height, instantaneous set-up and wind speed is that these parameters to a large extent determine the wave height in shallow water. Another important principle is that storm events should be related primarily to the unpredictable weather rather than to the (for our purposes) deterministic astronomical tide. A high sea level is composed of astronomical tide and a set-up generated by wind and air pressure variations, but only the height of the set-up is really uncertain; we can think of the contribution of the astronomical tide to the sea level maximum during a storm as a bounded disturbance. Therefore, the combinations of high astronomical tide and low set-up are of no practical interest. Also, storm events selected based on threshold exceedances of sea level only would tend to synchronize with the tidal maxima, and as a consequence, show strong clustering. For these reasons, instantaneous set-up provides a better basis for selection of storm events than sea level. Surge-tide interaction is rather small and will only have a minor effect on selection of storm events; therefore it is neglected. The wave period has not been considered because a long mean wave period may be related to swell from a distant storm field carrying little energy.

- Selection based on exceedances of thresholds is a well-established procedure known as the peak-over-threshold method in the case of a single parameter. The approach is readily extended to multivariate storm extremes, as opposed to selection of annual maxima.

- Storm events are defined as time-intervals over which at least one parameter of interest exceeds its threshold in order to make sure that the storm maximum of each parameter is included in the data, provided it exceeds the threshold. This ensures that estimates of joint distributions as in (De Valk and Zitman, 1992) are consistent with the univariate distributions of the individual parameters. This principle can be applied either
  - to each location separately, jointly for all parameters, or
  - jointly for all locations and parameters.

The latter option has been adopted. The advantage is that storm events are the same along the entire coast so they are easily compared and, if necessary, correlated. A possible complication is that different storm events may be joined because spatially distinct storm events overlap in time. For example, if the area of interest would be the entire globe, there would always be a storm event somewhere so they would all be joined. The correct approach would then be to identify storm events as distinct events in the space/time domain rather than only in the time domain. In the HYDRA project, we focus on a region which is smaller than the typical scale of a storm.
Therefore, we probably do not have to worry about this aspect. Tests have been carried out to confirm this (see section 2.2 below).

- Thresholds are parameter- and location-dependent to make sure that a sufficient number of data is available at each location and of each parameter. Physical effects (exposure related to coastline geometry, bathymetry, local storm climate) cause differences in the exceedance frequencies of the same parameter at different sites. In fact, we could also have chosen different thresholds for different wind directions or mean wave directions, but this would have made the procedure rather complicated.

- All high values of all parameters at all locations are still present in the selected data set. This makes it possible to derive various types of directional wave and wind statistics at a later stage of the project; for some possibilities, see appendix B.

The fraction $f$ which determines the thresholds for all parameter and locations should be chosen

- large enough that for each parameter a sufficient number of storm maxima exceeding the threshold remains, and
- small enough that storm events are clearly distinguished; they should not cover very long periods.

If correction is required, then all thresholds are adjusted simultaneously by changing the fraction $f$.

Note that we do not require the number of storm events to be limited to a small number, e.g. an average number of five per year. The only function of storm selection is to isolate events. When estimating exceedance frequencies for e.g. storm maxima of significant wave height, an additional selection can be applied of the storm maxima which exceed a high threshold; see Chapter 3.

The requirement that a sufficient number of data is selected is critical in particular for the estimation of directional exceedance frequencies of $H_{m0}$. These can be defined in several ways; see Appendix B. One approach considers exceedance frequencies of directional storm maxima of $H_{m0}$ for each of a fixed set of directional sectors. For a particular directional sector, a storm maximum is defined here as the maximum of those values of $H_{m0}$ which occur during the storm event while the (wave or wind) direction is in that sector; see Figure 2b for an illustration. Note that a single storm event may have directional maxima in several sectors if the wind or wave direction changes during that storm event. As a consequence, we cannot simply add directional exceedance frequencies for non-overlapping sectors to obtain the omnidirectional exceedance frequencies. Instead, exceedance frequencies should be determined not only for individual sectors but also for combinations of adjacent sectors; see Appendix B.

With the estimation of directional storm maxima in mind it should be verified that, for each directional sector of interest, a certain minimum number of storm events is selected during which the wave direction is in that sector while $H_{m0}$ exceeds its threshold, and similarly for wind direction. Wind- and wave directions are defined as opposite to the direction of the
wind vector or propagation direction of the waves. The directional sectors considered in the estimation of statistics along the 20 m depth contour are the 30-degree sectors

015-045 degrees  NNE
345-015 degrees  N
315-345 degrees  NNW
285-315 degrees  WNW
255-285 degrees  W
225-255 degrees  WSW

or combinations of these sectors. On an average, wind speeds decrease with wind direction turning from south-west to north, so lack of data will occur first in the north; see also (De Valk and Otta, 1993). Therefore, the number of directional storm maxima of $H_m$ in the sector N (345-015 degrees) exceeding the threshold was checked. Another reason for choosing the sector N is that some of the highest wave heights have occurred while the wind came from this direction (De Valk and Otta, 1993).

### 2.2 Results

Selections were made following the procedure described above with the fraction $f$ of data passing the data-selection threshold equal to 0.05 and 0.025. For the latter case, the thresholds per parameter and location are listed below.

<table>
<thead>
<tr>
<th>Location</th>
<th>$H_m$ [m]</th>
<th>$u_{10}$ [m/s]</th>
<th>set-up [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>3.40</td>
<td>14.5</td>
<td>0.55</td>
</tr>
<tr>
<td>YMB</td>
<td>3.55</td>
<td>16.0</td>
<td>0.66</td>
</tr>
<tr>
<td>ELD</td>
<td>3.75</td>
<td>13.5</td>
<td>0.67</td>
</tr>
<tr>
<td>SDN</td>
<td>3.58</td>
<td>14.5</td>
<td>0.70</td>
</tr>
<tr>
<td>K13</td>
<td>3.84</td>
<td>13.5</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Table 1 Data-selection thresholds with fraction of data retained $f=0.025$

Figure 3 shows the sample distribution of storm duration for this value of $f$. The numbers of directional storm maxima for the direction N (245-015 degrees) are given in Table 2 below for wind- and wave direction and for the three parameters significant wave height $H_m$, wind speed $u_{10}$ and instantaneous set-up at the location EUR.

<table>
<thead>
<tr>
<th>$f = 0.025$, wind direction in sector N</th>
<th>$H_m$</th>
<th>$u_{10}$</th>
<th>set-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>no of storm maxima:</td>
<td>29</td>
<td>11</td>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$f = 0.025$, wave direction in sector N</th>
<th>$H_m$</th>
<th>$u_{10}$</th>
<th>set-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>no of storm maxima:</td>
<td>30</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 2 Number of directional storm maxima at EUR selected using the thresholds in Table 1
Clearly, the numbers of storm maxima in the sector N are too small to allow fairly accurate estimates of exceedance frequencies from these data. Therefore, selection was also carried out with the value 0.05 for the fraction $f$. Thresholds corresponding to $f=0.05$ are:

<table>
<thead>
<tr>
<th></th>
<th>$H_{wo}$ [m]</th>
<th>$u_{10}$ [m/s]</th>
<th>set-up [m]</th>
<th>Total number of storm events: 1090</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>2.95</td>
<td>13.0</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>YMS</td>
<td>3.02</td>
<td>14.5</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>ELD</td>
<td>3.10</td>
<td>12.0</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>SON</td>
<td>2.91</td>
<td>13.5</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>K13</td>
<td>3.28</td>
<td>12.0</td>
<td>0.34</td>
<td></td>
</tr>
</tbody>
</table>

Table 3  Data-selection thresholds with fraction of data retained $f=0.05$

The thresholds are not very different from the thresholds with the smaller value of $f$. A number of 1090 storm events in 13 years may appear rather large, but it should be remembered that in general not all data are used to estimate exceedance frequencies; estimates are based on nondirectional or directional storm maxima of $H_{wo}$ which exceed a threshold which is higher than the storm selection threshold in Table 3 (see also Chapter 3). The sample distribution of storm duration is plotted in Figure 4, and the numbers of directional storm maxima for the direction N (245-015 degrees) are given in Table 4 below:

<table>
<thead>
<tr>
<th>$f = 0.05$, wind direction in sector N</th>
<th>$H_{wo}$</th>
<th>$u_{10}$</th>
<th>set-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>no of storm maxima:</td>
<td>52</td>
<td>30</td>
<td>52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$f = 0.05$, mean wave direction in sector N</th>
<th>$H_{wo}$</th>
<th>$u_{10}$</th>
<th>set-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>no of storm maxima:</td>
<td>56</td>
<td>25</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 4  Number of directional storm maxima at EUR selected using the thresholds in Table 3

The numbers of storm maxima for this direction are still not very large but at least sufficient to allow estimation of the distribution of threshold exceedances for a few thresholds exceeding the data-selection thresholds in Table 3. If these numbers are not sufficient, then directional sectors may be joined to obtain wider sectors with more data.

Sample exceedance frequencies from the omnidirectional storm maxima of $H_{wo}$ selected with $f=0.05$ were computed by simply dividing the rank of the sample (one for the highest, etc.) by the number of years covered by the data-set. Plots are shown in Figure 5 for the four measuring locations EUR, YMS6, ELD and SON near the 20 m depth contour. Symbols indicate the measured storm maxima; the drawn lines are the exact observed exceedance frequencies defined as the number of times that a value has been exceeded, divided by the length of the observation period. From Figure 5, it can be seen that the highest storm maxima of significant wave height generally increase along the coast from south-west to north-east.

The essential assumption underlying the application of exceedance frequencies of storm maxima to risk assessment is that the occurrences of storm maxima which exceed a high level behave in the limit as a Poisson process with seasonally-dependent intensity; see Appendix A.
This limiting Poisson character of storm maxima is the only assumption about *independence of storm events* that is required. If seasonal fluctuations of the wave climate are neglected, then the assumption implies that intervals between such successive storm maxima follow an exponential distribution. In order to check this, sample distributions of intervals between storm events having maxima of $H_{m_0}$, which exceed certain thresholds have been plotted in Figure 6. The interval between two storm events is expressed here by the number of hours between the beginnings of the storm events. Sample distributions have been plotted for thresholds increasing with a step of 0.20 m starting from the storm selection threshold in Table 3. The two plots for each location differ only in the extent of the vertical (10 log of probability of exceedance) scale. An exponential distribution would correspond to a straight line in the figures. If storms having high maxima of significant wave height have a limiting Poisson character, then the plotted sample distributions should tend to a straight line with increasing threshold. In Figure 6, curves which are higher correspond to higher thresholds.

In the figures with the extended vertical scale for EUR, YM6, ELD and SON, we see that for the lower thresholds (corresponding to the lower curves in the figure), intervals between storms events in the range of one to five days occur more often than should be expected from a Poisson process. This clustering may be related to typical meteorological time scales; removing closely spaced storm events would not change it. In fact, relatively few very short intervals are present. Note also that some deviation from an exponential distribution is likely to occur because of the seasonal dependence of frequency of occurrence of storm events. For the storm maxima exceeding higher levels, the sample distributions for most locations seem to tend to an exponential distribution. For ELD, for example, this occurs quite suddenly when the threshold exceeds 5 m.

It is common practice to estimate exceedance frequencies of wind speeds, sea levels etc. from data restricted to the storm season. In the Netherlands, the storm season is generally chosen from October 1 to March 31; (De Haan, 1990) motivates this choice for the case of high-tide sea levels. Selecting the data collected in this season from the set of storm data selected with $f=0.05$, the sample exceedance frequencies plotted in Figure 7 are obtained. A comparison with Figure 5 shows that restricting the data-set to the storm season hardly affects the selected storm maxima. Apparently, high values of $H_{m_0}$ rarely occur from April to September. The total number of storm events occurring in the chosen storm season is 734. The sample exceedance frequencies for the storm season are plotted separately for each location in Figure 8.

From Figure 4, we see that a few very long storm events are present. It should be realized that it makes little sense to judge the distribution of storm length solely on the basis of the length of the longest storm events in a sample. Figure 4 indicates that this distribution is well approximated by an exponential, so the length of the longest storm event will simply increase with the volume of data or, equivalently, with the number of years of data that are analyzed. However, examination of the data indicated that some storm events which would have been distinguished when selected separately for each location were joined into a single storm event by the procedure of section 2.1.

In order to find out whether this would affect the estimates, the procedure of section 2.1 was applied to the data of the five locations separately, again with the value 0.05 for the fraction $f$ of data retained above the threshold. The distribution of storm duration has been plotted in Figure 9 for the location ELD. Clearly, the storm durations have been reduced considerably.
as a result of carrying out the selection for each location separately. However, the highest storm maxima of $H_{m_0}$ have remained practically the same, as can be seen in the plots in Figure 10; compare with Figure 7. In section 3.2, it is shown that also the fits of a conditional Weibull distribution to the data selected locally and jointly for all locations are practically identical. In conclusion, selecting storm events locally or jointly for all locations affects the storm length but has practically no effect on the values of the highest storm maxima of $H_{m_0}$. 
3 Estimates of omnidirectional exceedance frequencies of significant wave height

3.1 Method

Exceedance frequencies of significant wave height were estimated from the maxima during the storm events selected jointly for all locations, using a fraction of data retained \( f = 0.05 \) and restricting storm events to the storm season October 1 to March 31 (see section 2.1 and Figure 8). Only the storm maxima of \( H_{m_0} \) exceeding the data-selection threshold were retained.

Exceedance frequencies of \( H_{m_0} \) are defined as the mean number of storm events per year during which \( H_{m_0} \) exceeds a particular value. Estimates are derived from the estimate of the distribution function \( F^\omega \) of the storm maxima of \( H_{m_0} \) which exceed a threshold \( \omega \). The exceedance frequency of the value \( a \) of \( H_{m_0} \) is estimated as the product of \( 1 - F^\omega(a) \) and the observed mean number of storm maxima per year exceeding \( \omega \). The function of the threshold \( \omega \) is to exclude storm maxima of \( H_{m_0} \) with relatively low values which occur frequently in order to increase the influence of the highest storm maxima in the data set, which carry most information about the tail of the distribution. Estimates are made for a range of values of the threshold \( \omega \) above the data-selection threshold in Table 3 (section 2.2).

The distribution \( F^\omega \) of the storm maxima of \( H_{m_0} \) which exceed a threshold \( \omega \) is assumed to be a conditional two-parameter Weibull distribution:

\[
F^\omega(a) = 1 - e^{-(a/\alpha)^\beta} + \omega(a/\alpha)^\beta
\]

(1)

following the recommendations in (De Valk and Otta, 1993). In (1), \( \sigma \) is a scale parameter and \( \alpha \) a shape parameter. The parameter \( \alpha \) determines the curvature of the logarithm of \( 1 - F^\omega \). The case of \( \alpha = 1 \) corresponds to an exponential distribution. In this case, the probability of exceedance, plotted on a semilogarithmic scale, is a straight line with slope equal to \( -1/\sigma \). Note that the storm maxima of \( H_{m_0} \) are assumed to follow the distribution function (1) with the same \( \alpha \) and \( \sigma \) for all values of the threshold \( \omega \); of course, the estimates of \( \alpha \) and \( \sigma \) may vary with the threshold, but when (1) provides a fair approximation of the sampling distribution, the parameter estimates should be expected to fluctuate around fixed values over a certain range of thresholds.

For a range of values of the threshold \( \omega \), the parameters \( \alpha \) and \( \sigma \) of the distribution function \( F^\omega \) are fitted by the maximum likelihood method. The estimates of the parameters \( \alpha \) and \( \sigma \) as well as of the value of \( H_{m_0} \) corresponding to an exceedance frequency of \( 10^{-4} \) per year are plotted as a function of the number of samples exceeding \( \omega \). Also, a modification of this approach has been tested in which \( \alpha \) is kept fixed and the scale parameter \( \sigma \) is estimated for a range of thresholds, for various values of \( \alpha \). The ideas behind the latter approach are explained in more detail in section 3.4.
The procedure described above can also be applied to directional storm maxima. In this phase of the study, it has only been applied to omnidirectional storm maxima of $H_{m_0}$ at five locations as a first test of the approach. It should be noted that estimates from omnidirectional storm maxima are not necessarily the best estimates obtainable of the omnidirectional exceedance frequencies. Restricting the data-set to a subset corresponding to a particular sector of wind- or wave direction may lead to estimates of the $10^4$ per year $H_{m_0}$ which are higher than those from the omnidirectional data-set; see also (De Valk, 1993; Appendix A).

### 3.2 Two-parameter Weibull estimates

In Figures 11-14, the following results have been plotted as a function of the number of storm maxima of $H_{m_0}$ retained for each of the five locations EUR, YM6, ELD, SON and K13:

- (a) estimate of the value of $H_{m_0}$ corresponding to an exceedance frequency of $10^4$ per year;
- (b) estimate of the parameter $\alpha$;
- (c) estimate of the parameter $\sigma$;
- (d) threshold value $\omega$;

The variation in the estimates of the $10^4$ per year significant wave height with the number of samples is considerable. For most locations, the fluctuation is fairly stable within a band of one to three metres. The estimates for EUR show the smallest fluctuation. At ELD, the $10^4$ per year significant wave height estimates decrease steadily from 13 to 9 m with decreasing number of samples, but increase sharply below 30 samples. The data in Figure 8c indicate that this behaviour is caused by the abrupt change in the slope of the logarithm of the sample exceedance frequencies. At first, the downward curvature of $1 - F^\omega$ is increased in an attempt to follow the data until the threshold exceeds the value where the slope appears to change. It remains the question whether this phenomenon has a physical cause or is just a coincidence. In the former case, we might want to model it deterministically, rather than to try to fit a smooth distribution to these data. Directional estimates may provide more insight.

The estimates of the $10^4$ per year significant wave height as a function of threshold derived from the maximum likelihood estimates of the Weibull parameters were subjectively interpreted, resulting in the following guesses:

- $10^4$ per year $H_{m_0}$
  - EUR about 7.5 m
  - YM6 about 9.0 m
  - ELD 9.0-11.0 m
  - SON about 10.5 m
  - K13 about 9.0 m

Figure 15 shows the two-parameter Weibull fits to the 50 highest storm maxima of $H_{m_0}$ at the five locations.
3.3 Sensitivity of estimates to data-selection method

In Figure 16, estimates of the Weibull parameters and corresponding $10^{-4}$ per year significant wave height from the storm maxima selected per location have been plotted as a function of threshold for the location ELD. They are practically identical to the estimates from the maxima at ELD of storm events selected jointly for all five locations; see Figure 11c. This confirms the conclusion in section 2.2 that joint selection of storm events for all five locations hardly affects the highest values of the storm maxima of $H_{m_0}$.

3.4 Estimates with a fixed shape parameter

In an attempt to reduce the variability of the estimate of the $10^{-4}$ per year $H_{m_0}$, estimates with fixed values of the parameter $\alpha$ were made for the values 1.0, 1.1, 1.2, ..., 3.0 of $\alpha$. Estimates of the $10^{-4}$ per year $H_{m_0}$ as a function of threshold are shown in Figure 17. With $\alpha$ fixed, the variability is indeed considerably reduced. This shows that in maximum likelihood fitting of both $\alpha$ and $\sigma$, the estimates of $\alpha$ contribute most of the variation of the $10^{-4}$ per year $H_{m_0}$ as a function of threshold.

In (De Valk, 1993; Section 5.2), the following stepwise approach to tail fitting was proposed. First, any curvature present in the sampling distribution of storm maxima of $H_{m_0}$ is suppressed by applying a smooth transformation to $H_{m_0}$ which does not affect the limit distribution of threshold exceedances. An example of such a transformation is a power. Next, the limit distribution of threshold exceedances is determined from the transformed data. Most likely there will not be sufficient evidence to support the choice of any limit distribution other than the exponential so the exponential limit distribution will be assumed; a rather conservative assumption, which implies that there is no finite upper bound to $H_{m_0}$. The scale parameter can be fitted as a function of threshold, from which an appropriate estimate is derived. In the case that the transformation applied on $H_{m_0}$ is a power, the distribution for the storm maxima of $H_{m_0}$ obtained is of the form (1).

A weakness in that approach was that no recipe was given for determining the transformation to be applied to $H_{m_0}$, say, the value of the power $\alpha$. One way to do this has already been suggested above, namely, carry out the procedure for different values of $\alpha$ and determine an appropriate value for $\alpha$ by examining the behaviour of the estimates of $\sigma$ (or the $10^{-4}$ per year $H_{m_0}$) as a function of threshold. If a power is an appropriate transformation and $\alpha$ is properly chosen, the estimates of $\sigma$ are expected to approach a constant value with decreasing number of samples, apart from fluctuations which increase in magnitude. An advantage of this method is that other considerations can be taken into account as well. For example, a smaller value of $\alpha$ will lead to more conservative estimates, as can be seen in Figure 17. Also, the likelihood of the storm maxima with the given value of $\alpha$ can be considered; a suitable test statistic is easily constructed. Finally, spatial coherence of the values of $\alpha$ can be a point of consideration; see e.g. (Rijken, 1983).
The estimates of the $10^{-4}$ per year significant wave height as a function of threshold for various fixed values of $\alpha$ were subjectively interpreted, resulting in the following guesses:

$$10^{-4} \text{ per year } H_m \alpha$$

<table>
<thead>
<tr>
<th>EUR</th>
<th>about</th>
<th>8.0 m</th>
<th>2.5-3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>YM6</td>
<td></td>
<td>9.5-10.0 m</td>
<td>2.0-2.5</td>
</tr>
<tr>
<td>ELD</td>
<td></td>
<td>9.5-10.0 m</td>
<td>2.0-3.0</td>
</tr>
<tr>
<td>SON</td>
<td></td>
<td>10.5-11.0 m</td>
<td>2.0</td>
</tr>
<tr>
<td>K13</td>
<td>about</td>
<td>10.0 m</td>
<td>2.0</td>
</tr>
</tbody>
</table>

(intervals do not represent error bands). These values are generally higher than those based on the maximum likelihood estimates of both $\alpha$ and $\sigma$. It may be possible to replace the subjective choice of the most appropriate value of $\alpha$ by some test for the vanishing of the dependence of the scale parameter $\sigma$ on the threshold.

Figure 18 shows the fits to 50 storm maxima of $H_m$ for selected values of the parameter $\alpha$ close to the values indicated in the Table above.

In conclusion, the conventional approach of fitting all parameters of the tail of the distribution function simultaneously is straightforward in principle but application can be complicated by the variation of the estimates with threshold. The proposed stepwise method for estimating the $10^{-4}$ per year value of $H_m$ seems viable both from the theoretical point of view and based on the results produced. In particular, choosing a value from the estimates obtained over a range of thresholds becomes less of a gamble.
References


Locations of wave buoy measurements.
Thin straight lines indicate thresholds
Fat lines on the time-axis indicate the selected storms
ILLUSTRATION OF THE DEFINITION OF DIRECTIONAL STORM MAXIMA IN TABLES 2 AND 4

DELFt HYDRAULICS
SAMPLE DISTRIBUTION OF STORM DURATION
FOR THE CASE \( f = 0.025 \)

DELFt HYDRAULICS
H 1931 FIG. 3
Sample distribution of storm duration.

Fraction of exceedance

Storm duration [3-hour intervals]
Sample exceedance frequencies: EUR (x); YM6 (+); ELD (o); SON (*).

Sample exceedance frequencies of storm maxima of $H_m0$ at EUR, YM6, ELD and SON.

Delft Hydraulics H 1931 Fig. 5
EUR, Hm0 > 2.95, 3.15, 3.35,... m.

10log of probability of exceedance

EUR, Hm0 > 2.95, 3.15, 3.35,... m.

10log of probability of exceedance

SAMPLE DISTRIBUTION OF INTERVALS BETWEEN STORM EVENTS DURING WHICH Hm0 EXCEEDS A LEVEL PLOTTED FOR INCREASING LEVELS

DELFt HYDRAULICS
YM6, $H_m > 3.02$, $3.22$, $3.42$, ... m.

10log of probability of exceedance

lag [hours]

YM6, $H_m > 3.02$, $3.22$, $3.42$, ... m.

10log of probability of exceedance

lag [hours]

SAMPLE DISTRIBUTION OF INTERVALS BETWEEN STORM EVENTS DURING WHICH $H_m$ EXCEEDS A LEVEL PLOTTED FOR INCREASING LEVELS

DELT HYDRAULICS

YM6

H 1931 FIG. 6b
Sample distribution of intervals between storm events during which $H_m0$ exceeds a level plotted for increasing levels.

ELD, $H_m0 > 3.19, 3.39, 3.59, \ldots$ m.

10 log of probability of exceedance

Lag [hours]

DELFT HYDRAULICS
SAMPLE DISTRIBUTION OF INTERVALS BETWEEN STORM EVENTS DURING WHICH Hm0 EXCEEDS A LEVEL PLOTTED FOR INCREASING LEVELS

DELTFT HYDRAULICS

H 1931 FIG. 6d
K13, \( H_m > 3.28, 3.48, 3.68, \ldots \) m.

Sample distribution of intervals between storm events during which \( H_m \) exceeds a level plotted for increasing levels.

Delft Hydraulics

H 1931

Fig. 6e
Sample exceedance frequencies: EUR (x); YM6 (+); ELD (o); SON (*)

Sample exceedance frequencies of storm maxima of Hm0 during the storm season at EUR, YM6, ELD and SON.
EUR winter: sample exceedance frequencies.
YM6 winter: sample exceedance frequencies.
ELD winter: sample exceedance frequencies.
SON winter: sample exceedance frequencies.

Sample exceedance frequencies of storm maxima of $H_m0$ during the storm season at SON.
K13 winter: sample exceedance frequencies.

SAMPLE EXCEEDANCE FREQUENCIES OF STORM MAXIMA OF $H_{m0}$ DURING THE STORM SEASON AT K13

DELFt HYDRAULICS  H 1931  FIG. 8e
ELD: sample distribution of storm duration.

Sample distribution of storm duration for ELD with storms selected separately for each location.
Sample exceedance frequencies: EUR (x); YM6 (+); ELD (o); SON (*).
EUR winter: estimates of once in 10,000 yr Hm0.

YM6 winter: estimates of once in 10,000 yr Hm0.

Estimates of the $10^{-4}$ per year Hm0 from two-parameter Weibull fits as a function of the number of storm maxima.

Delft Hydraulics
ELD winter: estimates of once in 10,000 yr $H_m0$.

SON winter: estimates of once in 10,000 yr $H_m0$.

Estimates of the $10^{-4}$ per year $H_m0$ from two-parameter Weibull fits as a function of the number of storm maxima.
K13 winter: estimates of once in 10,000 yr Hm0.
EUR winter: Weibull parameter estimates.

YM6 winter: Weibull parameter estimates.

ESTIMATES OF THE WEIBULL PARAMETER $\alpha$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA

DELT HYDRAULICS
ELD winter: Weibull parameter estimates.

SON winter: Weibull parameter estimates.

ESTIMATES OF THE WEIBULL PARAMETER $\alpha$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA

DELFt HYDRAULICS
K13 winter: Weibull parameter estimates.

ESTIMATES OF THE WEIBULL PARAMETER $\alpha$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA

DELFt HYDRAULICS

H 1931 FIG. 12c
EUR winter: Weibull parameter estimates.

YM6 winter: Weibull parameter estimates.

Estimates of the Weibull parameter $\sigma$ as a function of the number of storm maxima.

Delft Hydraulics
ELD winter: Weibull parameter estimates.

SON winter: Weibull parameter estimates.

ESTIMATES OF THE WEIBULL PARAMETER $\sigma$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA

DELFt HYDRAULICS
K13 winter: Weibull parameter estimates.

ESTIMATES OF THE WEIBULL PARAMETER $\sigma$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA

DELFt HYDRAULICS
EUR winter: threshold Hm0.

YM6 winter: threshold Hm0.

Threshold ω as a function of the number of storm maxima.
Threshold \( \omega \) as a function of the number of storm maxima

DELFT HYDRAULICS

H 1931 FIG. 14b
EUR; 2-parameter Weibull fit, 50 samples.

\[ \omega = 4.50 \text{ m} \]

YM6; 2-parameter Weibull fit, 50 samples.

\[ \omega = 4.64 \text{ m} \]

TWO-PARAMETER WEIBULL FITS TO THE 50 HIGHEST STORM MAXIMA OF \( Hm0 \)

DELFt HYDRAULICS

H 1931  FIG. 15a
TWO-PARAMETER WEIBULL FITS TO THE 50 HIGHEST
STORM MAXIMA OF \( H_{m0} \)

DELFT HYDRAULICS

H 1931  FIG. 15b
K13; 2-parameter Weibull fit, 50 samples.

\[ \omega = 5.05 \text{ m} \]

exceedance frequency [yr]

Hm0 [m]
ELD winter: estimates of once in 10,000 yr Hm0.

ELD winter: Weibull parameter estimates.
ELD winter: Weibull parameter estimates.

![Graph showing sigma vs. no. of storm maxima]

ELD winter: threshold Hm0.

![Graph showing Hm0 vs. no. of storm maxima]

Estimates of the $10^{-4}$ per year $Hm0$, Weibull parameters $\alpha$, $\sigma$ and threshold $\omega$ as a function of the number of storm maxima for ELD, with storm events selected per location.

DELFt HYDRAULICS
ESTIMATES OF THE $10^{-4}$ PER YEAR $Hm0$ AS A FUNCTION OF THE NUMBER OF STORM MAXIMA FOR FIXED VALUES 1.0, 1.5,..., 3.0 OF $\alpha$
ELD winter: \(0.0001 \text{ yr}^{-1} \text{ Hm0, } \alpha = 1.0-3.0\).

\[\begin{array}{c}
\text{no. of storm maxima} \\
\text{SON winter: } 0.0001 \text{ yr}^{-1} \text{ Hm0, } \alpha = 1.0-3.0.
\end{array}\]

ESTIMATES OF THE \(10^{-4}\) PER YEAR Hm0 AS A FUNCTION OF THE NUMBER OF STORM MAXIMA FOR FIXED VALUES 1.0, 1.5, ..., 3.0 OF \(\alpha\)

DELFt HYDRAULICS

H 1931 FIG. 17b
K13 winter: .0001 /yr Hm0, alpha = 1.0-3.0.

ESTIMATES OF THE 10^-4 PER YEAR Hm0 AS A FUNCTION OF THE NUMBER OF STORM MAXIMA FOR FIXED VALUES 1.0, 1.5, ..., 3.0 OF \( \alpha \)

DELFIT HYDRAULICS
EUR winter; alpha = 2.5, 3; 50 storm maxima.

\[ \omega = 4.50 \text{ m} \]

YM6 winter; alpha = 2.0, 2.5; 50 storm maxima.

\[ \omega = 4.64 \text{ m} \]

FITS TO THE 50 HIGHEST STORM MAXIMA OF Hm0 WITH SELECTED VALUES OF \( \alpha \)

DELFT HYDRAULICS

H 1931 FIG. 18a
ELD winter; alpha= 2.5, 3.0; 50 storm maxima.

$\omega = 4.75 \text{ m}$

SON winter; alpha= 2.0, 2.5; 50 storm maxima.

$\omega = 4.65 \text{ m}$

FITS TO THE 50 HIGHEST STORM MAXIMA OF Hm0 WITH SELECTED VALUES OF $\alpha$

DELFt HYDRAULICS
Appendix A

Considerations related to independence in data selection
To estimate meaningful frequencies of occurrence of extreme events or annual probabilities of occurrence of these events from a sequence of the variable(s) of interest, the effects of clustering of occurrences needs to be eliminated. Roughly, there are two approaches. We will first consider the case of a random sequence of a single variable $u$.

1 In the first approach, all data exceeding a particular level are selected to estimate the tail of the distribution function $G$ of the variable $u$, or equivalently, its survivor function $G_c = 1 - G$. The value $G_c(a)$ for some level $a$ can be regarded as the mean fraction of the time that $u$ spends above $a$. In general, the values of $u$ at different instants are not independent. In that case, $G_c$ cannot be used to determine the probability of $u$ exceeding some value in a certain time-interval. However, assuming that the random sequence is stationary and satisfies a so-called mixing property (Leadbetter et al., 1983), the mean length of an excursion above a high level converges to a limit when the level is increased, so this limit can be used to normalize $G_c$ to obtain the mean number of upcrossings of a high level per unit of time (frequency of occurrence of an upcrossing). An advantage of this approach is that it makes efficient use of the data, because all high values observed are used to estimate $G_c$. A disadvantage may be the reliance on a mixing property, which has been doubted for sequences of high-tide set-up by (De Haan, 1990). At least, the data should provide sufficient evidence that a limiting length of an excursion is indeed approached within the range of exceedance frequencies of interest.

2 The second approach focuses on extreme events, for example excursions above a certain level, as the basic data on which the estimates are based. For a stationary process satisfying certain conditions, the upcrossings of a high threshold are a Poisson process and maxima during the excursions above that level become independent in the limit; see (Leadbetter et al., 1983). Weather-related variables such as significant wave height can be modelled as seasonally-dependent processes and upcrossings of a high level behave in the limit as a Poisson process with seasonally-dependent intensity $\mu$. For such a process, the probability of $m$ occurrences in an arbitrary time-interval $J$ is given by

$$P[N_J=m] = \frac{M_J^m}{m!} \exp(-M_J) \quad M_J = \int_J \mu_\tau d\tau$$

with $\mu_\tau$ periodic with period one year, so $\mu_{t+1} = \mu_t$.

for all $t$. The probability of at least one occurrence in a period of $n$ years ($n$ integer) can be obtained from the annual frequency of occurrence $M_{(0,1)}$ as

$$P[N_{(0,n)} \geq 1] = 1 - P[N_{(0,n)} = 0] = 1 - \exp(-nM_{(0,1)})$$

When strictly applying this approach, the only practical requirement is to make sure that the threshold is high enough such that the number of upcrossings in an arbitrary time-interval follows the Poisson distribution closely enough. Then no additional measures are needed to leave out maxima which are too close to be considered
independent in some vague sense. In practice, there may be some doubt as to whether the threshold is high enough that the effect of minor fluctuations around this threshold can be neglected. A simple approach to get rid of such parasitic maxima is to select storm events based on a rather low threshold and then to compute the estimates from the storm maxima exceeding a higher threshold \( \omega \), for increasing values of \( \omega \). One may expect that clustering will vanish for these storm maxima, more rapidly with increasing level than for level upcrossings. The reason is that small fluctuations around a low level do not affect the number of high storm maxima whereas small fluctuations around a high level do not affect the selection of storm events. Independence of storm maxima exceeding a sufficiently high level should hold in the limit. However, strict independence of the storm maxima is not even required for maximum likelihood fitting of the parameters of a distribution. The maximum likelihood estimator converges provided that the sampling distribution converges, regardless of dependence. Generalization of the approach to multivariate extremes is straightforward.

The Poisson character of sufficiently high storm maxima can be checked from plots of the sample distribution of intervals between successive storm maxima. In the stationary case, they should correspond to an exponential distribution. This is not so in the more general case of a Poisson process with seasonal intensity. Still, clustering of storm events at certain time scales can be detected from such plots. An alternative method is described in (Dillingh et al., 1993). Independence of the sequence of storm maxima exceeding a level can also be checked for increasing levels.

A similar approach to data-selection can be applied when estimating annual probabilities of exceedance. In this case, not the storm maxima are used, which exceed a high threshold but the \( m \) highest storm maxima in each year of data. This approach is not easily generalized to multivariate extremes.
Appendix B

Three ways to define directional storm maxima and directional exceedance frequencies
1 Single direction per storm event

description: Each storm event has a single wind or wave direction, which is the direction coinciding with the storm maximum of the most important parameter. This parameter can be sea level, set-up, significant wave height, or wind speed. The significant wave height and the sea level affect the nearshore wave heights most directly; probably, sea level would be the best choice because it has the biggest impact on the nearshore wave height. It is logical to extend this approach to all other parameters; for each storm event, only the values coinciding with the sea level maximum are considered. This is the most elementary way to deal with multivariate extremes. Storm events with the storm direction in a particular sector are collected and exceedance frequencies are determined of for example the significant wave height values of these storm events.

advantages: — Directional exceedance frequencies can be added to obtain omnidirectional exceedance frequencies.
— The form of the results is suitable for assessment of the risk of failure of coastal defenses.

drawbacks: — Does not recognize the fact that wind is generally varying during a storm event, which is a basic characteristic of a storm wind field.
— Related to this, any observations of large wave heights in directions other than that at the storm maximum will disappear from the data set. This may lead to serious bias in exceedance frequencies for directions which generally do not coincide with the storm maximum.

2 Direction may vary during a storm event

description: For a particular directional sector, the directional storm maximum is defined as the maximum of the values during a storm event which coincide with wind or wave direction in a particular sector; see Figure 2b for an illustration. The frequency of exceedance of an arbitrary level by directional storm maxima is defined straightforwardly. In principle, this type of directional exceedance frequency should be estimated for arbitrary sectors, so in practice e.g. for 30-degree sectors and combinations of these.

advantages: — It recognizes the fact that wind generally varies during a storm event, which is a basic characteristic of a storm wind field. Therefore no bias caused by considering only the direction at the storm maximum (see 1 above).
— The form of the results is suitable for assessment of the risk of failure of coastal defenses. The frequency of occurrence of a combination of significant wave height and mean wave direction in an arbitrary set of values can be approximated in a conservative, though fairly accurate, manner.
drawbacks: — Omnidirectional estimates cannot be derived by adding the results for the individual sectors. However, this should not need to be a drawback, because besides the estimates for the different sectors, also estimates for all possible unions of adjacent sectors can be made. The consistency requirement in this case is that exceedance frequencies for a subsector do not exceed those for the sector which it is part of.

3 Projections of the wind vector on directional axes

description: Axes pointing in four orthogonal/opposite directions are chosen, e.g. SW, NW, NE and (if relevant) SE. Consider the joint distributions of pairs of wind components along these axes, e.g. of the SW and NW component, during an arbitrary storm, as defined in (De Valk and Zitman, 1992; Section 4.1.3). The components of the wind vector in opposite directions are independent, so at most four pairs of orthogonal components need to be considered, one pair for each quadrant in the plane. From the joint distribution of two orthogonal component, also the distribution of the projection of the wind vector in any direction in that quadrant can be approximated. Deriving omnidirectional exceedance frequencies of wind speed is more complicated; a conservative approximation can be derived based on the assumption of independence of the wind speed maxima in each of the four quadrants, but this may not be accurate enough. Extension to wave spectral parameters is possible e.g. by combining the significant wave height and the direction to a vector in some manner. Maybe more interesting from a physical point of view is to present omnidirectional statistics of wave energy (or significant wave height) and directional statistics of the energy flux vector; the combination should also give information about the wave period.

advantages: — It recognizes that directions may change during a storm event.
— The projection of a wind vector along an axis varies smoothly in time during a storm.

drawbacks: — Exceedance frequencies for derived parameters such as projections on arbitrary axes or wind speed can only be approximated; the achievable accuracy of the approximation is yet unknown.
— In general, results are not easily converted into the statistics required for assessment of the risk of failure of the coastal protection.
— It will require much effort.
For definition 2 above, there are two different approaches to estimate directional exceedance frequencies:

- Select the directional storm maxima for each sector of interest separately and estimate the exceedance frequencies for that sector;

- Consider the conditional probability of the wave or wind direction being in a given sector during a storm event while the significant wave height or wind speed exceeds a high level. Assume that this conditional probability tends to a limit upon increasing this level for every sector. All that needs to be done is
  - to determine the sectors for which the limiting conditional probability is nonzero;
  - for these sectors, to estimate the limiting conditional probabilities.

Note that the existence of nonzero limiting probabilities for certain sectors implies that the directional exceedance frequency curves in the limit on a logarithmic frequency scale differ only in a shift along the log-frequency axis. The idea behind the assumption of fixed nonzero limiting probabilities for certain sectors is that one or several storm types dominate the extreme wave climate at a particular location. The wind and wave direction need not to remain in one particular sector during such a typical storm; however, only for directions which occur during a typical storm while no additional limitations on wave height or wind speed are present, the conditional probability will remain nonzero. After identifying these critical sectors, the exceedance frequency curve for their union can be estimated (which is more accurate than separately for all these sectors), and the directional exceedance frequencies derived from the result using the limiting conditional probabilities. Of course, the omnidirectional exceedance frequencies must be higher, and in the limit, proportional to the exceedance frequencies for the critical sectors. For the remaining sectors for which the limiting conditional probability is zero, exceedance frequency curves need to be estimated separately. Advantages of this approach are that it will be much more reliable than simply making estimates separately for all directional sectors, and also, it may be possible to achieve better directional resolution in the critical sectors without loss of accuracy.