A review of statistical methods of analysis applied to the wave climate in the area of the Agulhas Bank in the Southern Indian Ocean

K. R. MacHutchon
Land and Shoreline Infrastructure Management Company, LasiMANCO (Pty) Ltd, South Africa

Abstract

A Wave Climate is normally defined in terms of (i) significant wave heights (Hs or $H_{m0}$), (ii) zero crossing, significant, mean or spectral peak wave periods (Tz, Ts, Tm or Tp respectively), (iii) wave directions, and (iv) joint frequencies of Hs or $H_{m0}$ and T. Wave climate data is statistically analysed to derive extreme values and design wave data for specified probabilities of exceedance or return periods.

This paper reviews the application of different methods of statistical analysis to the Southern Indian Ocean Area on the edge of the Agulhas Bank, using a data bank comprising wave data recordings, which were collected over a six-year period, from January 1998 to December 2003, on the Mossgas FA offshore gas drilling platform, on the Agulhas Bank in the Southern Indian Ocean.

Different samples from the data bank have been analysed in the paper and the results of these analyses have been extrapolated to derive extreme wave height estimates. The paper has taken cognisance of the previous work, which has been carried out in the same area by J. Rossouw and M. Rossouw [5, 8, 9].

1 Introduction

It is interesting to consider that the sea can have a very calming influence on one, even although its surface is never completely still itself. On the contrary, the sea surface is a continuously and chaotically moving medium which often vents its awesome energy, on a random basis, on sea shores and shoreline protection structures, on ships at sea, and on offshore structures.
Stochastic processes can be regarded as providing a closer approximation to reality, rather than a human invention to accommodate our deficiency in understanding the physical world. A real life problem cannot be completely described by an infinite number of parameters, thus it cannot be fully represented in a deterministic way. The stochastic representation of a system should not be considered, however, as a rival to determinism, but rather as a higher level of approximation to reality than the reality of the relevant deterministic laws governing the physical processes in the elements of the system [1].

Wave climates are generally defined in terms of wave parameters comprising; wave height (H); wavelength (L); and wave period (T) [2] as well as wave direction and they are governed by wind origins, speeds, fetches, durations and directions.

Wind-driven wave amplitudes, and range of excited wavelengths, depend on the turbulent energy flux from the atmosphere to the sea surface. This flux in turn is strongly dependant on the temperature stratification above the sea surface; for the same wind speed, a stronger flux yields more and/or bigger waves and roughness, while a weaker flux yields less and/or smaller waves and roughness [3].

2 Location of the recorded wave data analysed

The location where the data that has been analysed in this paper was recorded is some 85 km offshore of the southern South African Coast, on the Agulhas Bank and in the South Indian Ocean. The water depth in the area is 113 m. The area exists on a shelf, which is subject to very significant deep-sea swells from the west-southwest in the winter months, and it is fairly close to the strong, south-western flowing, Agulhas Current, and its retroflection which comprises return currents to the east of the Agulhas bank.

3 Description of the wave monitoring system used

The wave monitoring system, which recorded the wave data, is located on the FA Gas Drilling Platform and it forms part of a comprehensive Meteorological and Oceanographic data collection and processing installation.

The location of the FA platform has been shown in figure 1.

The system comprises two Marex Designed Processor and Display racks, one of which acquires wind speeds and wind directions as well as air temperatures and barometric pressures, while the other collects tide and wave data. The wave recorder is a downward looking radar type of instrument, which is positioned 20.0 above mean sea level, where it is attached to the platform support structure. The recorder has a minimum and maximum range of 7.0m and 50.0 m respectively. The wave data collects 20.0-minute time series records, comprising elevations of the sea surface, at hourly intervals. It processes the raw data, to derive relevant significant wave heights (Hs) and average zero crossing periods (Tz), before overwriting the original data and transmitting the maximum wave heights Hmax (according to PIANC standards), together with Hs and Tz to the shore, where it is stored in the CSIR’s offices in Stellenbosch, South Africa.
4 A systems approach to wave climates and the analysis of wave data

4.1 Wave climates

The term “wave climate” normally relates to a selected site or region and it describes what wave conditions can be expected there, in terms of wave statistics comprising wave heights, wave periods and wave directions. Wave climate data
generally has a time component in terms of either calendar years, or months of the year, or seasons of the year, and it is an essential input for navigation, as well as the design of offshore and coastal structures and coastal processes. The principal wave climate components have been shown in the figure 2.

4.2 The analysis of wave data

The analysis of wave data normally comprises the following elements which have been shown in the network in figure 3:

- The analysis of raw time series data in either the Time Domain or the Frequency Domain to determine parameters which can be analysed statistically to provide wave climate information
- The estimation of extreme conditions for Statistical Probability Analysis to determine Engineering Design Criteria

![Wave data analysis network breakdown structure](image)

Figure 3: Wave data analysis network breakdown structure. Wave data analysis network breakdown structure.

5 Wave climate for the study area in the South Indian Ocean

5.1 Significant wave height and average period plots

Traditionally sea state is a scale for the average wave height somewhat similar to the Beaufort scale for wind, and it refers to wave height, wave period and character of waves on the surface of the ocean.
Plots of significant wave heights (Hs) and average zero crossing periods (Tz), for either years or months or seasons, can be useful when used with Pierson-Moskowitz Sea Spectrum data, to compare the area under investigation with Open Sea State conditions. Annual data at the recording station for the period January 1998 to December 2003, inclusive, has been shown in figure 4.

**Figure 4:** Significant Wave Heights (Hs) and Average Zero Crossing Period (Tz) plots at the recording station, with sea states based on the Pierson-Moskowitz Sea Spectrum [www.oceandata.com].

**Figure 5:** Wind direction diagram for the study area.
5.2 Wind direction diagram

The growth of waves in deep water is governed mainly by three factors comprising the wind velocity, the duration of time during which the wind blows and the length of the fetch over which the wind blows. Initially, when the duration is small, the wave heights and periods are small, but with increasing duration, they increase throughout the fetch length, as a function of either fetch and wind velocity, or duration and wind velocity or wind velocity only, depending on different combinations of the controlling factors [4].

In the Southern Indian Ocean, the fetch is long, in the order of 900 km, [5] the duration can be long and the velocity can be great, as the result of the regular typical cyclonic weather patterns associated with low-pressure conditions in the southern ocean. A wind direction diagram for the recording station in the study area of the Southern Indian Ocean is given in figure 5.

5.3 Joint frequency plots

Joint Frequency plots of wave heights and concurrent wave periods give useful information on wave climate [6]. A plot for the study area has been given in figure 6, where the contours represent the percentage of the hourly occurrences when the respective wave heights and periods occurred simultaneously from 0.5% to 15%.

![Joint percentage frequency contours of hourly Hs and Tz for the recording station.](image)

Figure 6: Joint percentage frequency contours of hourly Hs and Tz for the recording station.
It can be seen from the plot that there is a weak correlation between the wave heights and periods. This is due to the fact that the wave climate is governed by local winds and swells. In enclosed waters, where the wave climate is generally governed by local winds only, there is a much stronger correlation between wave heights and periods.

6 Statistical analysis of wave climate data

6.1 Fitting of distributions and data selection

Selecting data for extremal analysis, estimating parameters in the distribution function and choosing an extremal distribution function must be done carefully. Each of these choices can significantly influence the estimated extremal values, especially those in very rare events [7].

The effect of missing data, which may have been lost during severe storms, has been investigated by J Rossouw [5] who found that the Extreme I Distribution, using the methods of moments to estimate the parameters of the distribution, is fairly robust to missing data and unbiased inaccuracies in the recording of wave heights.

6.2 Independent randomness and stationary requirements for data

Theoretically each data value should be from a different event to ensure statistical independence between values. It is assumed that the statistics of extreme events are stationary over the period of record and in the future [7]. The issue of independent event sampling, as the basis for analysis of the data at the recording station, is discussed in more detail below.

Ideally, a sampling plan needs be developed whereby independent and identically distributed data samples can be selected from a total data set [5], before the analysis of the data can be commenced.

The standard statistical test for independence of data recorded as a time series is to do serial correlation between the data recorded at increasing shifts in the sampling interval. As the shift increases the serial correlation coefficient reduces. J Rossouw [5] found that wave heights recorded at six hourly intervals on the Agulhas bank are highly correlated and that the serial correlation coefficient only dropped to below 0.1 after a sixty-hour shift. Considering the above, as well as possibly selecting independent $H_{m0}$ values on an event (weekly) basis, it was finally found that predicted design waves in the Southern Indian Ocean area are insensitive to the method of sampling [5].

The Coastal Engineering Manual [7] details the generally accepted distributions, to be used for the fitting of probability distribution functions for long term statistical analysis, as the Fisher-Tippett I (or Gumbel), the Weibull, the Fisher-Tippett II, the Log Normal, the Log Pearson Type III, the Pearson Type III, the Binomial and the Poisson (discrete) Distributions.

The method currently preferred in South Africa for estimating extreme or design wave heights uses the method of moments to fit a two parameter
Extreme I (Fisher-Tippett I) probability distribution to a total sample from the winter months. This method has been found to provide stable estimates, especially for the southern Atlantic area [8].

The IAHR committee recommends the fit of a three parameter Weibull Distribution to wave data selected using the Peaks Over Threshold (POT) sampling method, but it has been found that the threshold level influences the estimates when using the IAHR method. This can be attributed to the fact that too much emphasis is placed on the upper tail of the distribution, when high thresholds are chosen [8].

The Extreme I method is considered to be applicable in areas having strong correlation between the mean and extreme wave heights. This is the case when storms of similar origin occur regularly, such as the storms occurring during the winter months in the study area. It is also important that these storms should be responsible for the extreme events, e.g. intense frontal systems moving on a regular basis over the area of interest. The Extreme I method is not considered to be applicable to areas where cyclones occur, or in semi protected areas such as bays or in areas which have mostly calm conditions but with occasional severe storms such as the Mediterranean [9].

![Figure 7: Probability density function plot.](image)

**6.3 Probability density functions and probability distribution plots**

Probability Density Functions of the Significant Wave Heights in the study area for the Full Hourly Annual sample, the Six Hour Interval sample and the Full Hourly Winter Period sample have been given in figure 7:

A Typical Extreme I Distribution plot for the Full Hourly Winter Period Sample is presented in figure 8. The position of the significant wave height, with a return period of 100 years has been marked on the plot and it can be calculated by either using the Extreme I distribution parameters as indicate below, or by
linear regression of the Distribution Plot for a given probability \( p = 1 - 1/T_r \), where
\( T_r \) is called the recurrence interval or return period [5].

The Extreme I distribution for the significant wave height \( H_{m0} \) is [5]:

\[
H_{m0} = \alpha - \beta \ln(-\ln(p)).
\]

Based on the method of moments, \( \alpha \) and \( \beta \) can be calculated from the average (\( \mu \)) and variance (\( \sigma^2 \)) of the sample as follows [5]:

Average = \( \mu = \alpha + 0.5772 (\beta) \)

Variance = \( \sigma^2 = \beta^2 \pi^2 / 6 \)

Figure 8: Probability distribution plot for full hourly Winter Period sample.

The different 1 in 100 year significant wave heights for the different samples, referred to above have been summarised in the following table:

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>( H_{m0(100)} ) by Parameter</th>
<th>( H_{m0(100)} ) by Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Winter Sample</td>
<td>13.42 m</td>
<td>12.24 m</td>
</tr>
<tr>
<td>Full Annual Sample</td>
<td>12.89 m</td>
<td>11.93 m</td>
</tr>
<tr>
<td>6 Hour Lag Sample</td>
<td>11.46 m</td>
<td>11.17 m</td>
</tr>
</tbody>
</table>

7 Conclusions

The conclusions of the review carried out in this paper are that:

- The review confirms that the Extreme I distribution is applicable to the study area, as recommended by J Rossouw [5]
- The Study Area is subject to well-known and well-established weather patterns, which governs the wave climate in the region.
The statistical methods used to analyse extremal wave data on the Agulhas Bank in the Southern Indian Ocean are:
  o Particular to the area, due to the nature of the wave climate and its causes
  o Robust, from the point of view of estimating distribution parameters and the loss of records during storms.

Acknowledgements

PetroSA and CSIR for the supply of wave and meteorological recorded data; E Bosman, Senior Lecturer, Dept of Civil Eng. University of Stellenbosch, and M Rossouw, Coastal Research Engineer, CSIR for technical input and support.

References

[5] Jan Rossouw, Design for the Southern African Coastline Dissertation Approved for PhD Degree at the University of Stellenbosch