Ocean Engineering 144 (2017) 277–295

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

A regional extreme value analysis of ocean waves in a changing climate

Erik Vanem

DNV GL Group Technology and Research, Høvik, Norway

ARTICLE INFO

Keywords:
Ocean waves
Marine structures
Extreme value analysis
Regional frequency analysis
Climate change
Ocean environment
Environmental loads
Spatial analysis

ABSTRACT

This paper applies regional frequency analysis on extreme significant wave heights in the North Atlantic ocean for historical and projected ocean wave climates. Regional frequency analysis is applied in order to perform spatial extreme value analysis over a large ocean area. One of the main advantages of regional frequency analysis is that one borrows strength for nearby locations by site pooling in order to get more accurate and robust estimates of extreme significant wave heights. Moreover, a set of homogeneous regions are identified as ocean regions with similar characteristics. In this way, the effect of climate change on the ocean wave climate can be studied, both with regards to changes in extreme quantiles at certain location and to overall spatial changes. The results from the regional frequency analysis are presented in this paper. The results are generally found to be reasonable, suggesting that regional frequency analysis can be useful in the analysis of extreme ocean waves. In particular, regional frequency analysis yields narrower uncertainty bounds and hence more robust estimation of extreme quantiles, corresponding to long return periods.

1. Introduction and background

In the design of ships and other marine structures, one needs to take extreme environmental events into account in order to ensure that the structural integrity can be maintained throughout the intended lifetime of the structure. Therefore, there is much interest in reliable descriptions of the extremes of important parameters such as the significant wave height. Typically, one is interested in estimating return values corresponding to long return periods, e.g. the 20- or even the 100-year return value. Due to the large uncertainty associated with the physical wave process, one must often use statistical methods to describe the observed extreme events. Often, one is not only interested in a point estimate of the extreme event, but also require a description of the uncertainty of that estimate. This may be achieved by establishing a probability distribution for the event in question. Frequency analysis refers to the analysis of how often a specified event, for example an extreme sea state, will occur.

Statistical frequency analysis of a particular dataset can be done in many different ways, and includes estimation of a probability distribution function. However, when the interest is in very extreme events corresponding to events far out in the tail of the distribution, the uncertainty becomes large due to limited data. Typically, the return period of interest is long compared to the available time series of observations. However, in some cases there might be data available for the same variable at other locations that might provide additional information about the variable at the site of interest. In such situations, more precise estimates can be obtained by utilizing information in data for similar locations. Hence, the main purpose of regional frequency analysis is to borrow strength from data collected at different sites to be able to provide more accurate conclusions compared to what can be done by analysing data from merely one particular site. This is often referred to as trading space for time. Another advantage of regional frequency analysis is that it provides means for analysis of extreme events at locations within the regions where there are no observations or data, based on data from similar locations. Hence, regional frequency analysis provides a basis for spatial extreme value analysis. This is useful in many marine engineering applications. For example, as input in the design process for fixed installations in areas where there are little or no data available or in risk analysis and route planning of maritime transport.

In this paper, the results of regional frequency analysis of significant wave height data are presented. Regional frequency analysis is presented for different sets of historical wave data, obtained by numerical wave models, as well as for future projections of significant wave height based on running the same numerical wave models with model output from climate models, assuming particular future emission scenarios. By doing this, the effect of climate change on the wave climate can be assessed in different ways. First, the analysis can indicate how large return values of significant wave height might change in a future climate. Furthermore, by comparing the homogeneous regions and the spatial return values in historical and future climate, one may assess possible changes in the spatial distribution of extreme wave events in a future scenario.

E-mail address: Erik.Vanem@dnvgl.com.

http://dx.doi.org/10.1016/j.oceaneng.2017.08.027
Received 12 December 2016; Received in revised form 21 June 2017; Accepted 17 August 2017

Regional frequency analysis is well established in various geosciences, but quite few applications of RFA on marine events are reported in the literature. A recent application of regional frequency analysis for the estimation of extreme waves is presented in (Campos and Guedes Soares, 2016); see also (Van Gelder et al., 2001; Weiss et al., 2014a; Ma et al., 2006; Goda et al., 2000) for earlier studies. Regional frequency analysis was also mentioned as a promising candidate for modelling of extreme wave events in (Vanem, 2011). Recent marine applications of regional frequency also includes modelling of extreme storm surges (Bernadara et al., 2011).

Numerous projections and analysis of potential changes to the extreme ocean wave climate due to climate change have been presented in recent years, and this is an important issue in many ocean and coastal engineering applications (Caires et al., 2006; Wang et al., 2004, 2015; Wang and Swail, 2006; Vanem et al., 2012). However, there are very few cases where regional frequency analysis is applied to assess the effect of climate change on ocean waves. An application of Regional frequency analysis to projected wave data around the Korean peninsula is presented in (Lim et al., 2013), demonstrating the ability to identify changes in the 50-year return wave heights in a future climate.

The regional frequency analysis presented in this paper is based on the L-moments, as outlined in (Hosking and Wallis, 1997) and is applied on annual maximum significant wave height data. The remainder of the paper is organized as follows: Section 2 gives an introduction to the methodology and presents the main steps of a regional frequency analysis. Section 3 presents the wave climate data used in the study and the actual analysis is presented in section 4. Section 5 provides a discussion on some of the results and issues of the modelling and section 6 concludes the paper.

2. Regional frequency analysis

The following gives a brief introduction to regional frequency analysis. For a more thorough introduction, reference is made to textbooks such as (Hosking and Wallis, 1997).

The probability distribution or the frequency distribution of a random variable is a fundamental quantity in statistical analysis, and it describes how probable or frequent certain values of the variable are. The (cumulative) probability distribution function of a random variable X evaluated at X is defined as the probability of the random variable being less than or equal to the value X,

$$F(x) = Pr[X \leq x] \tag{1}$$

For strictly increasing and continuous cumulative distribution function, the inverse of the cumulative distribution function is called the quantile function, and this quantile function expresses the value of the random variable corresponding to its non-exceedance probability, p,

$$q(p) = F^{-1}(p) \tag{2}$$

Thus, for any probability, p, the quantile function of p is the unique real number X such that $F(x) = p$. Quantiles are often expressed in terms of return periods. The return period T of an extreme high value for a random variable is defined in terms of the quantile Q_T as follows

$$T_{high} = \frac{1}{1 - F(Q_T)} \tag{3}$$

That is, an event with return period T is so extreme that its magnitude has probability $1/T$ of being exceeded by any single event. For extreme low values, the return value is defined as

$$T_{low} = \frac{1}{F(Q_T)} \tag{4}$$

The goal of a frequency analysis is to estimate, as accurately as possible, the quantile function of a random variable. Regional frequency analysis does this for certain locations by exploiting information from locations with similar characteristics in order to obtain more accurate estimates of the quantile functions. This is particularly useful if return periods of interest are long compared to the length of the data record. Typically, the quantile of a return period T can only be estimated accurately if the length of the data record, n is $n \geq T$. In practice, however, this is rarely the case and one is typically interested in quantiles of longer return periods than available record lengths. In essence, regional frequency analysis provides a means for including similar datasets in the analysis in order to effectively increase n.

Implicit assumptions in regional frequency analysis are that observations at a particular location are identically distributed, that the observations are serially independent and that the observations at different sites are independent. It may be questionable whether these assumptions are fully satisfied for ocean wave data, especially if recorded at nearby locations, but it is argued that regional frequency analysis still gives better estimates of extreme quantiles compared to individual frequency analyses at each location in isolation. Essentially, dependence in the data will reduce the effective number of samples in the pooled dataset.

In the analysis presented in this paper, quantile estimates are based on L-moments and an index-wave approach for pooling data from different locations. The analysis is applied to the annual maximum significant wave height, and it is believed that using only the annual maximum data significantly reduces both the serial and the spatial dependencies in the data. Hence, it is assumed that the data can be assumed to be iid (independent and identically distributed) at each location and that they are also spatially independent. Hence, regional frequency analysis can be applied to the data.

The regional frequency analysis presented in this paper is based on the index wave approach (this procedure is commonly referred to as the index flood procedure in hydrology, but renamed index-wave in this analysis on wave heights). The main assumption in the index wave approach is that certain locations constitute homogeneous regions and that the frequency distributions (and thus also the quantile functions) at locations within each homogeneous region are identical apart from a location-specific scaling factor, referred to as the index wave. Hence, denoting the index wave at location i as μ_i, the quantile function at location i can be written as

$$Q_i(F) = \mu_i q(F) \quad \forall i \in \text{the homogeneous region}, \tag{5}$$

where $q(F)$ is referred to as the regional growth curve. The regional growth curve is a dimensionless quantile function that is common to all locations within the homogeneous region, i.e. it is the quantile function of the regional frequency distribution. The index wave is simply the mean of the location-specific frequency distribution and is estimated from the data as the sample mean at location i. Rescaled data that are divided by the estimated index wave are then used to estimate the regional growth curve. Typically, a parametric form is assumed for the regional growth curve and estimation of this reduces to the estimation of the distributional parameters. In this study, this is done by way of the L-moments. Having estimated the index wave $\hat{\mu}_i$ and the regional growth curve $\hat{q}(F)$, the quantile function at location i can be found as

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F), \tag{6}$$

which can be used for inference about extreme return values.

2.1. Main steps of a regional frequency analysis

According to (Hosking and Wallis, 1997), there are four main steps involved in regional frequency analysis:

1. Screening of the data
2. Identification of homogeneous regions
3. Choice of a frequency distribution