Global Estimates of Extreme Wind Speed and Wave Height

J. VINOTH AND I. R. YOUNG
Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Melbourne, Australia

(Manuscript received 17 February 2010, in final form 5 November 2010)

ABSTRACT

A long-term dataset of satellite altimeter measurements of significant wave height and wind speed, spanning 23 years, is analyzed to determine extreme values corresponding to a 100-yr return period. The analysis considers the suitability of both the initial distribution method (IDM) and peaks-over-threshold (POT) approaches and concludes that for wave height both IDM and POT methods can yield reliable results. For the first time, the global POT results for wave height show spatial consistency, a feature afforded by the larger dataset. The analyses also show that the POT approach is sensitive to spatial resolution. Since wind speed has greater spatial and temporal variability than wave height, the POT approach yields unreliable results for wind speed as a result of undersampling of peak events. The IDM approach does, however, generate extreme wind speed values in reasonable agreement with buoy estimates. The results show that the altimeter database can estimate 100-yr return period significant wave height to within 5% of buoy measurements and the 100-yr wind speed to within 10% of buoy measurements when using the IDM approach. Owing to the long dataset and global coverage, global estimates of extreme values can be developed on a 1° × 1° grid when using the IDM and a coarser 2° × 2° for the POT approach. The high-resolution 1° × 1° grid together with the long duration of the dataset means that finescale features not previously identified using altimeter data are clearly apparent in the IDM results. Goodness-of-fit tests show that the observed data conform to a Fisher–Tippett Type 1 (FT-1) distribution. Even in regions such as the Gulf of Mexico where extreme forcing is produced by small-scale hurricanes, the altimeter results are consistent with buoy data.

1. Introduction

The determination of extreme values of wind speed and wave height is a common requirement in many offshore applications, such as offshore structural design and operation. The usual approach is to estimate the value of the 50-yr or 100-yr return period wind speed or wave height based on a measured time series of limited duration. Here, for example, the 100-yr return period wave height is the wave height that would be exceeded on average once in 100 yr. As measured records are not available for 50 or 100 years, the normal approach is to fit a probability distribution to recorded data and extrapolate this to the required probability level (return period). Obviously, the accuracy of such an approach is limited by factors such as the length of the recorded time series and the ability of the chosen probability distribution to represent the extreme tail (low probability of occurrence) of the distribution. From a practical point of view, a major challenge is that long duration time series are seldom available at a particular location of interest. In such situations, it is common to use numerical model hindcast results in place of recorded data. Naturally, model data is only as reliable as the model physics and forcing wind fields. Hence, such an approach introduces an additional element of potential error. This is particularly the case when considering hindcasts of events many years in the past, where wind fields may be of low quality.

In recent years, the advent of satellite altimeter observations has provided the prospect of global coverage of measurements of wind speed and wave height. A number of studies have applied this data to the estimation of extreme wave and, to a lesser extent, wind conditions. Such studies have highlighted a number of unique challenges associated with the use of such data for extreme value analysis. Although altimeter data exists for approximately 25 years, previous studies have used only relatively short records, with a maximum duration of
approximately 11 yr. Use of longer duration time series would involve combining data from a variety of different platforms, and hence it would be necessary to ensure a consistent calibration of these multiparameter measurements over this extended period. The sampling pattern of the altimeter is also quite different to buoy data. In comparison to buoy data, altimeters sacrifice temporal coverage for high spatial density. That is, the satellite track revisits a particular location infrequently (typically of order 10 days) but provides excellent spatial coverage (typically observations every 5 to 7 km along the ground track). As a result, previous studies have averaged satellite data in a region surrounding the point of interest. Hence, point buoy measurements and spatial altimeter data measure slightly different quantities.

The present analysis extends these previous studies by using a far longer satellite altimeter dataset, consisting of measurements from seven separate altimeter missions dating back to 1985. This combined dataset has been consistently calibrated and validated by Zieger et al. (2009). The extended dataset provides a much more suitable basis for the assessment of the use of such data for extreme value analysis and to develop appropriate techniques for the estimation of extreme wind speed and wave height over the world’s oceans. In addition to developing such techniques, the present analysis also investigates the global distribution of extreme (100-yr return period) wind and wave conditions.

The arrangement of the paper is as follows. Section 2 provides a brief review of extreme value theory as applied to ocean wind and waves and an overview of previous applications of these approaches to altimeter data. Section 3 provides a summary of previous studies of extreme wind and wave climate using altimeter data. The altimeter dataset to be used in this analysis is described in section 4, followed by validation and testing against buoy data of the application of extreme value analysis to this data in section 5. Section 6 examines the global distribution of extreme wind speed and wave height. Finally, conclusions and discussion of the results are outlined in section 7.

2. Introduction to extreme value analysis

2a. Theoretical distributions

As outlined by Gumbel (1958), the aim is to determine the probability distribution from a sample of measured data. To form a valid distribution, the observations should be independent and identically distributed. The requirement for independence means that successive observations should not be correlated with one another. As typical wind generation systems will have durations of a number of hours, this implies that wind/wave observations may need to be separated by many hours to satisfy this condition. The requirement of identically distributed observations means that, should an area be exposed to quite different meteorological forcing events (e.g., trade winds and tropical cyclones), these systems should be considered separately.

Noting these basic requirements, there are three approaches that have been applied: the initial distribution method (IDM), the annual maximum method (AMM), and the peaks-over-threshold method (POT).

The IDM uses all recorded data and fits a cumulative distribution function (CDF) to this data. As there is no theoretical means to determine the most appropriate function, a number of CDFs are typically used—the one achieving the best fit to the data generally being accepted. Typical CDFs used with the IDM include (see Tucker 1991; Goda 1988):

- the Fisher–Tippett Type 1 (FT-1) or Gumbel distribution

\[ F(x) = \exp \left[ - \exp \left( \frac{x - A}{B} \right) \right], \quad (1) \]

where \( F(x) \) is the CDF of the variable \( x \), and \( A \) and \( B \) are parameters determined by the fitting process; the Weibull two-parameter (W2P) distribution

\[ F(x) = 1 - \exp \left[ -\left( \frac{x}{B} \right)^k \right], \quad (2) \]

where \( k \) is also a fitting parameter; and the Weibull three-parameter (W3P) distribution

\[ F(x) = 1 - \exp \left[ -\left( \frac{x - A}{B} \right)^k \right]. \quad (3) \]

The parameters in these distributions are generally determined by fitting the model CDF to the empirical cumulative distribution either by least squares, the method of moments, or the maximum likelihood method (see Holthuijsen 2007). Again, there is no theoretical way to choose between these fitting methods. Historically, (1) has been fitted using both the method of moments (in which case it is called the FT-1 distribution) and maximum likelihood (where it is called the Gumbel distribution) (Evans et al. 2000). In the remainder of the paper, the terminology FT-1 will be used when the distribution has been fitted with the method of moments and FT-1G when the distribution has been fitted using the maximum likelihood method. The Weibull distribution, together with the annual maximum and peaks-over-threshold methods considered below, are most commonly fitted using maximum likelihood (Evans et al. 2000).
The IDM suffers from two obvious deficiencies. First, as buoy observations are typically made at either hourly or three-hourly intervals, the data almost certainly violate the requirement that they are independent. Second, interest is in the tail of the distribution (extreme values), although the vast majority of the data used to fit the distribution comes from moderate conditions, that is, the body of the distribution. Thus, extrapolation to extreme events may result in significant error.

Despite these shortcomings, the IDM is commonly used (see for example Goda 1992, 1988; Ochi 1992; Tucker 1991).

One means to overcome the issue of independence of the data is to use only the maximum value in a 12-month period, the annual maximum method (AMM). Extreme-value theory indicates that the cumulative distribution of maxima will follow a generalized extreme-value (GEV) distribution (Castillo 1988)

\[ F(x) = \exp \left\{ \left[ 1 + k \left( \frac{x - A}{B} \right) \right]^{-1/k} \right\}. \] (4)

In addition to addressing the issue of independence, the AMM also provides a theoretical rationale for the choice of the CDF (4) and uses only extreme values associated with the tail of the distribution to determine the fitting parameters. However, as only one value per year is used, the available data is typically very small. As such, it is generally not practical to use the AMM approach with ocean wind and wave data.

A compromise between these two approaches is the peaks-over-threshold (POT) method (Goda 1992; Van Gelder and Vrijling 1999; Ferreira and Soares 1998). Under this approach, a threshold value is defined and a peak value is associated with each rise and fall of the measured data above the threshold. The rationale for this approach is that, if the threshold is selected correctly, the peak value associated with each storm (an extreme event) will be recorded. From extreme-value theory (Castillo 1988; Coles 2001), the distribution of the maxima in such a sequence of values above a threshold follows the generalized Pareto distribution (GPD)

\[ F(x) = 1 - \left[ 1 + k \left( \frac{x - A}{B} \right) \right]^{-1/k}. \] (5)

Although not having the same theoretical basis as (5), the Weibull distribution (3) has also been used to fit POT data (Soares and Henriques 1996).

Although the theoretical basis for the POT is compelling, the major issue is in selecting the value of the threshold. As pointed out by Coles (2001), there is no theoretical basis for the selection of such a value. For prediction at a single location, examination of storm data may enable a threshold value to be objectively determined. Alves and Young (2003) attempted this on a global scale using a database of storm conditions. A more straightforward approach has been proposed by Anderson et al. (2001), Caires and Sterl (2005), and Challenor et al. (2005), who propose simply to consider all data above a percentile value. Both the 90th and 93rd percentile have been suggested as reasonable limits.

b. Extrapolation to extreme values

Once the appropriate CDF is adopted, the aim is to use this distribution to determine the extreme wind speed or wave height. Therefore, it is necessary to determine the probability level that is associated with the 100-yr event (exceeded once in 100 yr), \( P(x < x^{100}) \), where \( x \) can be either \( H_s \) or \( U_{10} \). Following Goda (1988) and Mathiesen et al. (1994), for the POT approach \( P \) can be determined as

\[ P(x < x^{100}) = 1 - N_\gamma/(100N_{\text{POT}}), \] (6)

where \( N_{\text{POT}} \) is the number of data points in the POT analysis, and \( N_\gamma \) is the number of years covered by the analysis.

For the IDM analysis, \( P \) is commonly determined as

\[ P(x < x^{100}) = 1 - D/T_{100}, \] (7)

where \( D \) is a decorrelation time scale in hours for observations of \( x \), and \( T_{100} \) is the number of hours in 100 years. Consistent with the decorrelation time scale used in previous studies (Tucker 1991; Cooper and Forristall 1997; Teng 1998), a value of \( D = 3 \) h was used. Values of \( D \) ranging from 1 to 24 h were tested. The resulting values of 100-yr return period significant wave height \( H_s^{100} \) and 100-yr return period wind speed \( U_{10}^{100} \) did not vary significantly with the chosen value of \( D \).

In the subsequent analysis, both IDM and POT will be used to determine these extreme values. To clarify the method used to estimate the extreme value, the terminology \( H_s^{100} \) (IDM) or \( H_s^{100} \) (POT) (and similarly for wind speed) will be adopted. If no parentheses are included, then reference is to the generic extreme value, independent of the method of calculation.

c. Goodness of fit

In the application of any of the proposed extreme value approaches, the key issue is how well the recorded data fits the selected extreme value CDF. The “goodness of the fit” to the data can be determined from the statistics of the residual between the empirical cumulative distribution function (EDF), \( F_n(x) \), (i.e., the distribution
obtained from the recorded data) and the theoretical extreme value CDF model, \( F(x) \) (see Stephens 1986). There are two general techniques that have been proposed to measure this discrepancy.

Supremum statistics consider the largest positive and largest negative differences between the EDF and the CDF, where \( D^+ \) is the largest positive difference, and \( D^- \) is the largest negative difference:

\[
D^+ = \sup_x [F_n(x) - F(x)], \quad (8)
\]

\[
D^- = \sup_x [F(x) - F_n(x)], \quad (9)
\]

and

\[
D = \sup_x |F_n(x) - F(x)| = \max(D^+, D^-). \quad (10)
\]

The statistic \( D \) is the basis of the Kolmogorov–Smirnov test.

Quadratic statistics measure the discrepancy as

\[
Q = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \psi(x) \ dF(x). \quad (11)
\]

When \( \psi = 1 \), (11) defines the Cramér–von Mises statistic, \( W^2 \), and when \( \psi(x) = [F(x)][1 - F(x)]^{-1} \), (11) defines the Anderson and Darling (1952) statistic, \( A^2 \).

The values of \( D \), \( W^2 \), and \( A^2 \) can then be tested at the desired percentage point level to determine whether the null hypothesis (that the EDF is well approximated by the chosen CDF) is true or false.

Goda and Kobune (1990) have proposed an alternative to the above approaches by simply looking at the value of the correlation coefficient between the data \( x \) and its reduced variate: for example, the FT-1 distribution (1) can be expressed in linear form as \( y = (x - A)/B \), where \( y = -\ln \{-\ln[F(x)]\} \) is the reduced variate. In this case, the Goda and Kobune (1990) approach would evaluate the correlation coefficient between \( y \) and \( x \). A value of the correlation coefficient near one indicates a good fit.

As interest is concentrated on how well the shape of the extreme value tail of the distribution is represented, rather than the body of the distribution, a variation of the above approaches is to “censor” the data by only applying the goodness-of-fit test to data above an upper percentile (Stephens 1986). For instance, the tests could be applied only to the upper 20% of data.

Tabulated values against which calculated values of \( D \), \( W^2 \), and \( A^2 \) can be tested have been presented for the FT-1, FT-1G, and Weibull (2 and 3 parameter) distributions by Stephens (1970, 1977, 1986) and Goda and Kobune (1990). Embrechts et al. (1997) and Kotz and Nadarajah (2000) have developed approaches to determine confidence limits for the GPD. Similar tests or confidence limits for the GEV do not appear to be available in the literature.

3. Altimeter observations of extreme wind and wave climate

A number of studies have demonstrated that altimeter data can be used to determine global wind and wave climatology, including monthly mean values and exceedance probabilities up to the 90th percentile (Challenor et al. 1990; Tournadre and Ezraty 1990; Young 1994, 1999; Young and Holland 1996). The first attempt to determine extreme wave heights from altimeter data was conducted by Carter (1993). He considered the three years of Geosat data for the North Atlantic. The data was “gridded” in 2° × 2° bins, and the IDM was used with a FT-1 distribution. Despite the very short record, the approach yielded values of 50-yr return period significant wave height, \( H_s^{50} \) (IDM) consistent with buoy and climate atlas results for the area.

Panchang et al. (1999) carried out a similar study, comparing \( H_s^{50} \) (IDM) from the altimeter with buoys along the Atlantic coast of North America. They concluded that short-duration altimeter records can be used to estimate extreme values but noted the “rule of thumb” suggestions of Hogben (1988) that the data duration should be at least a third of the return period (implying 16 yr for \( H_s^{50} \)).

Alves and Young (2003) used a 10-yr dataset from Geosat, ERS-1, and TOPEX/Poseidon. In contrast to previous studies, they applied both the IDM (FT-1 distribution) and POT (W3P distribution) to estimate global values of \( H_s^{100} \). For POT calculations, a spatially variable threshold was used, based on a global storm database. They noted that the POT results were spatially highly variable, indicating the sensitivity of the choice of threshold and the magnitude of the errors associated with extrapolating the short-duration time series to \( H_s^{100} \) (POT) values. As a result, they concluded that the IDM was superior for application to short-duration altimeter records.

Challenor et al. (2005) and Wimmer et al. (2006) considered approximately 11 years of data from ERS-1, ERS-2, and TOPEX/Poseidon. In a similar fashion to Alves and Young (2003), they considered both the IDM (FT-1 distribution) and the POT with a 2° × 2° grid but applied the GPD distribution. Also, they set the threshold for the POT as the 90th percentile value. They found that this approach gave good agreement when compared to a single North Atlantic buoy, NODC 44004. The spatial diagrams for the POT approach provided for the
North Atlantic did, however, show significant spatial variability, as also noted by Alves and Young (2003). Chen et al. (2004) calculated global values of both $U_{100}$ and $H_{s100}$ based on the IDM and a FT-1G distribution from eight years of TOPEX data. In contrast to other studies, they used a 1° × 1° grid. They concluded that the differences between buoy- and altimeter-derived extreme values were approximately 10% for wind speed and 5% for significant wave height.

4. Altimeter database

As noted above, to obtain reliable estimates of extreme values of wind speed and wave height, it is desirable to have as long a dataset as possible. Previous studies of the use of altimeter data for extreme value analysis have used datasets less than 11 yr long. To assess 50-yr or 100-yr return period statistics, it is desirable to have a dataset at least spanning two decades. In addition, it is essential that the calibration of the measurement instruments has been consistent over this period of time so as to minimize the effects of changes in measurement platform over the period. For this purpose, the altimeter database of Zieger et al. (2009) has been adopted. In the 23-yr period 1985–2008, a total of seven separate satellite altimeter missions were operational (see Fig. 1). This dataset provides almost continuous coverage over the 23-yr period, with a break from 1990 to 1991.

Zieger et al. (2009) calibrated each of the satellite altimeters for wind speed and wave height against deep water in situ buoys. These calibrated results were then cross-validated between satellites operating at the same time by considering collocated measurements at cross-over points between the satellites. In this manner, any discontinuities due to changes of hardware or software or instrument drift were eliminated. The resulting dataset is believed to provide a consistent, high quality, global view of oceanic wind speed and wave height throughout this extended period. As such, it represents a unique tool to investigate extreme values of wind speed and wave height on a global scale.

5. Buoy–altimeter extreme value comparisons

Development of the appropriate techniques for the extreme value analysis of altimeter data and the assessment of the performance of these techniques was undertaken by comparison with a number of deep water U.S. National Ocean Data Center data buoys (Evans et al. 2003). The buoys used in this analysis are the same as those adopted by Alves and Young (2003) and a subset of those used by Zieger et al. (2009) for the calibration of the database. These buoys were operational throughout the full 23-yr duration of the dataset and generally recorded wind speed and significant wave height at a 1-h interval (note that some buoys used a 3-h interval in the early years of the dataset). The buoy locations were chosen such that they were a minimum of 200 km offshore and in water depth of at least 300 m. These criteria were chosen to ensure that altimeter data would not be influenced by the transition from land to water and an averaging area of 2° × 2° centered on the buoy would not include significant regions of finite depth. Full details of the data processing are given by Zieger et al. (2009). As shown in Fig. 2, the buoy locations cover a variety of geographical regions including the North Pacific, North Atlantic, Gulf of Mexico, and the Pacific trade wind belt (Hawaii).

a. Averaging area

The significant difference between buoy and altimeter data is that buoys provide high temporal density (1-h measurements) but poor spatial density (single or small number of points), whereas the altimeter has poor temporal density (ground tracks are repeated on the order of once every 10 days) but excellent spatial coverage (data point every 5–7 km along the track). The approach that has been traditionally used is to define an area and combine all altimeter data for this area, the assumption being that the wave climate in the area is homogeneous.
As the area becomes larger, the validity of the assumption of homogeneity clearly becomes questionable. However, the larger amount of data will yield an extreme value estimate with smaller confidence limits (i.e., apparently more stable).

A variety of averaging areas have been used in previous studies. Carter (1993), Alves and Young (2003), Challenor et al. (2005), and Wimmer et al. (2006) all adopted a 2° × 2° region, whereas Chen et al. (2004) concluded that a smaller 1° × 1° region yielded acceptable results. Panchang et al. (1999) compared a number of different-sized regions and finally adopted an even smaller 50-km radius. In contrast, based on numerical experiments, Cooper and Forristall (1997) recommended a larger region of 200–300 km. Because of the significantly larger dataset available in the present study, the influence of averaging area can be directly investigated.

A number of different-sized regions were considered, including: 1° × 1°, 2° × 2°, and a 50-km radius. In each case, these regions were centered on each of the buoys detailed in Fig. 2. Each pass of the satellite across the averaging region will generate a number of observations of wind speed and wave height, these observations clearly not being independent. Wimmer et al. (2006) refer to this process as clustering of the data. In this analysis a number of different approaches to declustering were investigated, including using all the data, replacing the pass with the mean, the median, or the maximum. All of these approaches yielded very similar results. Ultimately, the median value was adopted as this provided a stable measure, is statistically valid (i.e., independence of observations), and is consistent with previous studies (e.g., Wimmer et al. 2006).

The relative error of the altimeter estimate of the extreme wave height can be defined as \( \Delta r = (H_{\text{alt}}^{100} - H_{\text{buoy}}^{100}) / H_{\text{buoy}}^{100} \), where the subscripts refer to the altimeter and buoy values. Values of \( \Delta r \) for each of the buoys and each size of averaging area are shown in Table 1 for the IDM and FT-1 distribution and Table 2 for the POT and W3P distribution. In addition, the values summed over all \( n \) buoys of \( r_1 = (1/n)\sum |\Delta r| \) and \( r_2 = (1/n)\sum \Delta r \) are shown. The IDM results for \( r_1 \) and \( r_2 \) are very similar for all averaging areas and the values of \( \Delta r \) across all individual buoys are generally less than ±6%. The values of \( r_1 \) are all less than 3.5%, and values of \( r_2 \) are all less than ±1.0%.

The results of Table 1 clearly show that the IDM approach is insensitive to the averaging area. Adopting the larger 2° × 2° averaging area increases the amount of data available to construct the altimeter CDF, but this does not result in better agreement between buoy and altimeter.

In addition, the 50-km radius data was resampled such that only cases where the buoy and altimeter measurements were sampled within 30 min of each other were considered. Although this resulted in comparable values of sampled wave height, it significantly reduced the actual number of observations in the fitting process for the extreme value distribution. As a result, \( r_1 \) actually increased slightly to 3.07%. In situations where there are strong spatial gradients of wave height/wind speed increasing the size of the averaging area may compromise accuracy. Even for the Gulf of Mexico buoys (42001 and 42002), which are subject to the small-scale forcing of hurricanes, the largest area tested (2° × 2°) does not increase the relative error.

It should be noted that the excellent agreement between buoy- and altimeter-derived extreme values obtained using the IDM approach does not validate this extreme value method for altimeter data. Both the buoy and altimeter approaches using the IDM are estimates of the extreme value. It is possible that neither are good estimates of the actual extreme value (the same argument

<table>
<thead>
<tr>
<th>Geographic region</th>
<th>NODC buoy</th>
<th>1° × 1° region</th>
<th>2° × 2° region</th>
<th>50-km radius</th>
<th>Resampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. East Coast</td>
<td>41002</td>
<td>3.35%</td>
<td>1.67%</td>
<td>5.81%</td>
<td>−4.15%</td>
</tr>
<tr>
<td></td>
<td>44004</td>
<td>3.12%</td>
<td>1.12%</td>
<td>2.46%</td>
<td>−3.98%</td>
</tr>
<tr>
<td>Gulf of Mexico</td>
<td>42001</td>
<td>−3.86%</td>
<td>2.91%</td>
<td>−6.12%</td>
<td>0.49%</td>
</tr>
<tr>
<td></td>
<td>42002</td>
<td>−0.40%</td>
<td>−0.96%</td>
<td>−0.23%</td>
<td>−1.34%</td>
</tr>
<tr>
<td>U.S. West Coast</td>
<td>46001</td>
<td>−2.47%</td>
<td>−2.02%</td>
<td>−0.69%</td>
<td>0.63%</td>
</tr>
<tr>
<td></td>
<td>46002</td>
<td>1.81%</td>
<td>2.57%</td>
<td>0.37%</td>
<td>4.24%</td>
</tr>
<tr>
<td></td>
<td>46003</td>
<td>−0.45%</td>
<td>0.35%</td>
<td>−1.06%</td>
<td>−4.98%</td>
</tr>
<tr>
<td></td>
<td>46005</td>
<td>1.08%</td>
<td>1.35%</td>
<td>1.40%</td>
<td>−2.55%</td>
</tr>
<tr>
<td></td>
<td>46006</td>
<td>−1.78%</td>
<td>−1.69%</td>
<td>−1.42%</td>
<td>−1.78%</td>
</tr>
<tr>
<td>Hawaii</td>
<td>51001</td>
<td>−5.56%</td>
<td>−4.80%</td>
<td>−5.41%</td>
<td>−1.52%</td>
</tr>
</tbody>
</table>

\[ r_1 = (1/n)\sum |\Delta r| \]
\[ r_2 = (1/n)\sum \Delta r \]

TABLE 1. Values of the relative error, \( \Delta r \), between buoy and altimeter for different averaging areas for the altimeter. The results are shown for \( H_{\text{alt}}^{100} \) (IDM) using the IDM approach with a FT-1 distribution. Also shown are the mean absolute error \( r_1 \) and the mean error \( r_2 \). Each column shows a different averaging area with the buoy locations shown for each row.
Table 2. As in Table 1 but for $H_{100}^s$ (POT) using the POT approach with a W3P distribution. Also shown are the mean absolute error $r_1$ and the mean error $r_2$.

<table>
<thead>
<tr>
<th>Geographic region</th>
<th>NODC buoy</th>
<th>1° × 1° region</th>
<th>2° × 2° region</th>
<th>50-km radius</th>
<th>Resampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. East Coast</td>
<td>41002</td>
<td>−18.37%</td>
<td>−13.91%</td>
<td>−20.21%</td>
<td>−11.86%</td>
</tr>
<tr>
<td></td>
<td>44004</td>
<td>−17.62%</td>
<td>−13.80%</td>
<td>−39.16%</td>
<td>−2.61%</td>
</tr>
<tr>
<td>Gulf of Mexico</td>
<td>42001</td>
<td>−27.46%</td>
<td>3.56%</td>
<td>−41.61%</td>
<td>3.25%</td>
</tr>
<tr>
<td></td>
<td>42002</td>
<td>−22.46%</td>
<td>−23.30%</td>
<td>−27.73%</td>
<td>−7.52%</td>
</tr>
<tr>
<td>U.S. West Coast</td>
<td>46001</td>
<td>−16.01%</td>
<td>−14.12%</td>
<td>−14.12%</td>
<td>9.99%</td>
</tr>
<tr>
<td></td>
<td>46002</td>
<td>−25.11%</td>
<td>−9.64%</td>
<td>−13.56%</td>
<td>5.59%</td>
</tr>
<tr>
<td></td>
<td>46003</td>
<td>−17.38%</td>
<td>−15.58%</td>
<td>−27.66%</td>
<td>13.32%</td>
</tr>
<tr>
<td></td>
<td>46005</td>
<td>−7.49%</td>
<td>0.72%</td>
<td>−7.23%</td>
<td>7.46%</td>
</tr>
<tr>
<td></td>
<td>46006</td>
<td>−24.88%</td>
<td>−14.63%</td>
<td>−30.07%</td>
<td>4.87%</td>
</tr>
<tr>
<td>Hawaii</td>
<td>51001</td>
<td>−25.16%</td>
<td>−16.54%</td>
<td>−24.24%</td>
<td>10.93%</td>
</tr>
</tbody>
</table>

$r_1 = \frac{1}{n} \sum |\Delta r|$  
$r_2 = \frac{1}{n} \sum \Delta r$

holds for all approaches). Rather, the good agreement indicates that the altimeter data gives a very similar estimate of the CDF to the buoy and that this estimate is insensitive to the averaging area. The potential error in estimating the extreme value is introduced in extrapolating the limited-duration dataset to the extreme value. Here, the longer-duration dataset available in this study will reduce the extent of this extrapolation and hence the potential for error (compared to previous altimeter studies).

As a result of the comparisons in Table 1, subsequent analysis of the IDM approach has generally concentrated on the 1° × 1° averaging area, as the finer resolution has the potential to identify features in areas where there may be significant spatial gradients of wave height.

The results for the POT and W3P distribution comparisons between buoy and altimeter in Table 2 are quite different to the IDM results of Table 1. The values $\Delta r, r_1$, and $r_2$ are significantly larger for the POT analysis. Also, the averaging area has a clear impact on the results for the POT analysis. For the 1° × 1° analysis the values of $\Delta r$ at every location are negative, indicating that the altimeter approach underestimates compared to the buoy. This underprediction decreases when the averaging area is increased to 2° × 2° (i.e., $r_2$ reduces in magnitude from −20.20% to −11.73%). The underprediction disappears completely when the data is resampled, such that the buoy and altimeter data used are collocated in space and time ($r_2 = 3.34\%$). This result clearly indicates that, given comparable data, the POT approach will yield similar results for buoy and altimeter data. However, as the available altimeter data is limited, undersampling impacts the POT estimated extreme values. Increasing the size of the averaging area reduces this effect but does not completely eliminate it.

This result is not surprising. The IDM approach defines the body of the CDF and then attempts to extrapolate to extreme events. Such an approach is insensitive to undersampling, as the body of the CDF can be well defined with limited observations (the error comes in the extrapolation). In contrast, the POT approach attempts to define the extreme value tail of the CDF from observations of “peak” events. If undersampling means that insufficient “peaks” are measured, the result will directly impact on the resulting tail of the distribution (the observed distribution). As the POT approach appears to be sensitive to the averaging area, in subsequent analysis, both 1° × 1° and 2° × 2° averaging regions are considered.

b. Initial distribution method

The IDM was applied to 1° × 1° regions centered on each of the buoys for the FT-1 and FT-1G distributions. The resulting 100-yr return period values for both buoy and altimeter are shown in Table 3 for $H_{100}^s$ (IDM) and Table 4 for $H_{100}^s$ (IDM) (note that the 2° × 2° result is also shown, but not discussed here). Buoy and altimeter results are in excellent agreement for $H_{100}^s$ (IDM) for both distributions, as noted in Table 1. The distributions yield values of $r_1 = 2.11\%$ for the FT-1 distribution and $r_1 = 2.92\%$ for the FT-1G distribution. These values of $r_1$ are approximately half those reported by both Alves and Young (2003) and Chen et al. (2004) using IDM FT-1 and IDM FT-1G, respectively. The much larger datasets and longer duration of the time series significantly reduces the potential error in extrapolating the extreme value distributions. As a result, the buoy and altimeter results for the extreme values are in better agreement than these previous studies.

Figure 3a shows the values of $H_{100}^s$ (IDM) (FT-1) at each location, calculated for both the altimeter and buoy, together with 95% confidence limits calculated using the residual of the correlation coefficient (REC approach) proposed by Goda (2000). The excellent agreement at all buoy locations between buoy and altimeter $H_{100}^s$ (IDM)
both $H_s$ and $H_8$ averaging (Fig. 3a), and the buoy values often fall outside the confidence intervals at different buoy locations for different analysis methods and CDFs. Both $1^\circ \times 1^\circ$ and $2^\circ \times 2^\circ$ averaging areas are shown. The mean absolute errors between buoy and altimeter are also shown, as in Table 1.

Table 4 shows the comparable IDM results for wind speed, $U_{100}^2$ (IDM). The values of $r_1$ for each of the distributions are larger than the comparable values for $H_s^{100}$ (IDM)—for FT-1, $r_1 = 9.91\%$ and for FT-1G, $r_1 = 6.74\%$. These larger differences between buoy and altimeter reflect the greater spatial variability of wind speed compared to wave height and are consistent with the conclusion by Chen et al. (2004) that differences between buoy and altimeter extreme values are larger for wind speed than for wave height. Figure 3c shows a comparison at each buoy together with the respective 95% confidence limits. It is clear that the agreement is not as good as for wave height (Fig. 3a), and the buoy values often fall outside the confidence limits for the altimeter. The goodness-of-fit tests (see Table 4) show that the FT-1 distribution is slightly superior to the FT-1G distribution.

As the buoys used for the present comparisons are the same as adopted by Alves and Young (2003), a direct comparison can be made. The mean error between the two sets of altimeter estimates of $H_s^{100}$ (IDM) (as measured by $\Delta r$ between the two sets of altimeter estimates) was less than 5% at each buoy location, indicating comparable results.

c. Peaks-over-threshold

The POT approach was also applied at each buoy location to the altimeter and buoy data within both $1^\circ \times 1^\circ$
Table 4. As in Table 3 but for $H_{100}$ (m s$^{-1}$).

<table>
<thead>
<tr>
<th>POT (%)</th>
<th>GPD [100 (POT)]</th>
<th>W3P [100 (POT)]</th>
<th>IDM [100 (POT)]</th>
<th>FT-1G [100 (POT)]</th>
<th>Buoy IDM FT-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>17.31%</td>
<td>12.58%</td>
<td>10.08%</td>
<td>9.91%</td>
<td>5.04%</td>
</tr>
<tr>
<td>20</td>
<td>15.68%</td>
<td>11.36%</td>
<td>9.41%</td>
<td>9.28%</td>
<td>4.94%</td>
</tr>
<tr>
<td>30</td>
<td>14.05%</td>
<td>10.03%</td>
<td>8.47%</td>
<td>8.35%</td>
<td>4.61%</td>
</tr>
<tr>
<td>40</td>
<td>12.43%</td>
<td>9.00%</td>
<td>7.12%</td>
<td>7.00%</td>
<td>4.28%</td>
</tr>
<tr>
<td>50</td>
<td>10.80%</td>
<td>8.03%</td>
<td>6.24%</td>
<td>6.12%</td>
<td>3.95%</td>
</tr>
<tr>
<td>60</td>
<td>9.17%</td>
<td>7.05%</td>
<td>5.46%</td>
<td>5.34%</td>
<td>3.62%</td>
</tr>
<tr>
<td>70</td>
<td>7.54%</td>
<td>6.07%</td>
<td>4.88%</td>
<td>4.76%</td>
<td>3.29%</td>
</tr>
<tr>
<td>80</td>
<td>5.91%</td>
<td>5.09%</td>
<td>4.09%</td>
<td>3.97%</td>
<td>2.96%</td>
</tr>
<tr>
<td>90</td>
<td>4.28%</td>
<td>4.10%</td>
<td>3.19%</td>
<td>3.07%</td>
<td>2.63%</td>
</tr>
</tbody>
</table>

As noted above, there is no theoretical basis for selection of the threshold. The only practical approaches for application on a global scale are a fixed threshold (or range of fixed thresholds, which has little physical rationale) or the fixed percentile approach proposed by Anderson et al. (2001), Caires and Sterl (2005), and Challenor et al. (2005). Selection of such a percentile approach was explored in some detail. Figure 4 shows values of $H_{100}$ (POT) calculated at buoy 46002 for different values of the percentile threshold. Results are shown for both W3P and GPD distributions. It is clear from this figure that there is no clear choice for the selection of the threshold. As the threshold increases, the value of the resulting $H_{100}$ (POT) increases smoothly for the W3P approach. The GPD is less sensitive to the choice of the threshold. As pointed out by previous researchers (Alves and Young 2003; Challenor et al. 2005; Wimmer et al. 2006), selection of the threshold is a serious shortcoming of the POT approach. Consistent with Challenor et al. (2005), a value of 90% has been selected for further analysis, although it should be pointed out that the selection of an alternative level would impact the calculated values of extreme wind speed and wave height.

Table 3 provides values of $H_{100}^x$ (POT) using the POT with W3P and GPD distributions at each buoy location, and Table 4 provides the corresponding values of $U_{100}^x$ (POT). The mean error values for $H_{100}^x$ (POT) are: GPD, $r_1 = 17.31\%$, and W3P, $r_1 = 12.58\%$ ($2^\circ \times 2^\circ$ region); and for $U_{100}^x$ they are GPD, $r_1 = 40.62\%$, and W3P, $r_1 = 11.26\%$ ($2^\circ \times 2^\circ$ region). These values are significantly larger than values for the IDM, indicating significantly poorer agreement between buoy and altimeter for the POT. This poorer agreement between buoy and altimeter for the POT approach is also clear in Fig. 3b (for wave height) and Fig. 3d (for wind speed). This figure also shows the impact of the sampling region, with the $2^\circ \times 2^\circ$ region reducing the level of underprediction. The much wider confidence interval associated with the POT estimates can also be seen in these figures and Tables 3 and 4. Despite the larger confidence limits and the much less demanding criteria for goodness of fit, the POT W3P results satisfy a similar number of tests to the IDM results.

The POT and IDM results for $H_{100}^x$ are in reasonable agreement. Both GPD and W3P POT $H_{100}^x$ (POT) buoy values are approximately 3% lower than the corresponding buoy IDM FT-1 results when averaged across all buoys. In contrast, the wind speed results show much greater variability. The IDM FT-1G $U_{100}^x$ (IDM) buoy values
are approximately 11% higher than the corresponding IDM FT-1 buoy values. The differences between the POT and IDM estimates are even greater. The POT GPD and POT W3P buoy estimates of $U_{10}^{100}$ (POT) are 33% and 26% lower, respectively, than the IDM FT-1 buoy estimates when averaged across all buoys. Wind speed has much greater spatial and temporal variability than wave height, and this is clearly reflected in the broader range of values of $U_{10}^{100}$ resulting from these different estimation techniques.

The results of the buoy–altimeter comparisons of extreme value estimates can be summarized as follows.
1) FOR WAVE HEIGHT $H_s^{100}$
- The IDM and POT approaches produce similar results when applied to the same buoy data, irrespective of the CDF assumed (i.e., FT-1, FT-1G, GPD, or W3P).
- Altimeter IDM results are in good agreement with the buoy estimates and are not sensitive to the averaging area used (i.e., $1^\circ \times 1^\circ$ versus $2^\circ \times 2^\circ$).
- The POT altimeter results show greater variability than the IDM results when compared to corresponding buoy estimates. In addition, the POT results are sensitive to the averaging area with a $1^\circ \times 1^\circ$ averaging region significantly underestimating the extremes and a $2^\circ \times 2^\circ$ region moderately underestimating the values.

2) FOR WIND SPEED $U_{10}^{100}$
- There are significantly greater differences between buoy-derived values of extreme wind speed obtained by IDM and POT approaches than is the case for wave height.
- The differences between buoy and altimeter extreme value estimates are smaller for the IDM than the POT, although larger than for wave height.
- Whereas wave height POT GPD and W3P gave similar error results (comparing buoy to altimeter), for wind speed the POT GPD produces significantly larger differences than the corresponding POT W3P.

As a result of these comparisons, the global distributions to be considered below have adopted the following approaches:
- IDM FT-1 and $1^\circ \times 1^\circ$ averaging area and
- POT W3P and both $1^\circ \times 1^\circ$ and $2^\circ \times 2^\circ$ averaging areas.

**d. Altimeter calibration at extreme values**

Although the $H_s^{100}$ (IDM) comparisons above indicate quite good agreement between buoy and altimeter extreme values, it should be remembered that the calibration results of Zieger et al. (2009) involved few extreme wind/wave conditions. Therefore, some questions exist over the ability of the altimeter to measure extreme values of wind speed and wave height. As our focus is on the accurate determination of the shape of the extreme tail of the CDF, such data points are likely to significantly influence the accuracy of the results. To investigate the extreme value performance of the altimeter, a $2^\circ \times 2^\circ$ region was considered centered on each of the buoys. The full datasets for the respective buoy and altimeter pairs were then considered and percentile–percentile plots (also called Q-Q plots) were developed for each of the buoy–altimeter pairs for both $H_s$ and $U_{10}$. Figure 5 shows a typical result for Buoy 46002. The altimeter and buoy are in excellent agreement for all values, including the most extreme value plotted, the 99th percentile.

**6. Global distribution of extreme values**

Based on the results in section 5, both $1^\circ \times 1^\circ$ and $2^\circ \times 2^\circ$ grids were considered for the world and all data binned as appropriate. Each pass through a bin was represented by the median value and the resulting observations analyzed
using the various extreme value methods. For the IDM approach, the FT-1, FT-1G, and W2P were all applied, and for the POT approach the W3P and GPD were applied. Thresholds set at both 90% and 93% were used for each of these POT distributions. These approaches were used to calculate both \( U_{10}^{100} \) and \( H_{s}^{100} \) for every bin. In addition, the Kolmogorov–Smirnov, Cramér–von Mises, Anderson–Darling, and Goda goodness-of-fit tests were applied for each case tested. This full set of combinations is very large and it is not practical to present all results here. Figures 6–10 show a subset of this full set of cases.

Figure 6 shows color contour plots of \( H_{s}^{100} \) for each of IDM FT-1 (Fig. 6a), POT W3P (Fig. 6b), and POT GPD (Fig. 6c) for a 1° × 1° grid. The corresponding 2° × 2° results are shown in Fig. 7. The IDM FT-1G and IDM W2P distributions are not shown. The spatial distributions of these CDFs are both similar to the FT-1. The FT-1G distribution is also similar in magnitude to the FT-1 distribution, but the W2P is approximately 30%–40% lower than the FT-1. Figure 8 shows the corresponding set of plots for \( U_{10}^{100} \) with a 1° × 1° grid and Fig. 9 with a 2° × 2° grid. Figure 10 shows the global distribution of Cramér–von Mises goodness-of-fit results, censored at 20% for both IDM FT-1 \( H_{s}^{100} \) (IDM) and \( U_{10}^{100} \) (IDM) (1° × 1° grid).

The three panels in Fig. 6 clearly highlight the differences between the three approaches/distribution functions. The IDM FT-1 (Fig. 6a) is spatially consistent with the results of both Alves and Young (2003) and Chen et al. (2004). The extreme values at the high latitudes of both hemispheres and the relatively calm equatorial regions are the defining characteristics. As one would intuitively expect, the distributions of \( H_{s}^{100} \) (IDM) are spatially relatively smooth. The POT results (Fig. 6b, POT W3P; Fig. 6c, POT GPD) are, in contrast, more spatially variable, this small-scale variability being larger for the GPD results. Clearly, this variability is the result of undersampling in the much smaller dataset of peaks above the threshold. The larger datasets available in this study and the method of selecting the threshold (i.e., a fixed 90% value, rather than a variable level) do, however, mean that the results still reproduce the major spatial features (large wave heights at high latitude and calm equatorial regions). This is in contrast to the POT results of Alves and Young (2003) and Wimmer et al. (2006), which showed much greater spatial variability than reported here. It is believed that the larger dataset in the present analysis results in enhanced POT results. As shown by the buoy comparisons in section 5, the 1° × 1° POT results also underestimate the extreme values. This is clear in Fig. 6 with the POT results producing lower values than the IDM, particularly in the high latitudes.

In addition to the large-scale features outlined above, Fig. 6a shows many local features. The high waves, \( H_{s}^{100}(\text{IDM}) > 14 \text{ m} \), associated with the Somali jet (which is linked to the Asian monsoon) in the Arabian Sea is a clear local phenomenon. Also, a local maximum is clear within the Bay of Bengal, probably associated with tropical cyclone activity. Within the Mediterranean Sea, a local region of intensification is clear between France and North Africa, consistent with the strong mistral winds of this region. The effects of landmasses are also clear, with areas of reduced wave extremes east of South America and New Zealand where the strong westerly winds will have reduced fetch to generate high wave conditions. Even the isolated Kerguelen Islands in the southern Indian Ocean cast a significant “wave shadow” in the lee of the islands.

The 2° × 2° results for \( H_{s}^{100} \) (Fig. 7) are qualitatively similar to the 1° × 1° results although, as noted above, the POT results are slightly higher owing to the undersampling issue. As expected, the results are spatially smoother than the 1° × 1° results. As noted above, these results represent the first published global POT results that are spatially consistent. Whereas the quality of the IDM results will see little improvement by further extending the duration of the altimeter dataset, the POT results are clearly gaining in reliability with the longer datasets available.

Figure 10a shows the global distributions of the IDM FT-1 \( H_{s}^{100} \) (IDM) and the points (in black) where the Cramér–von Mises goodness-of-fit test (applied to the top 20% of the data) is accepted. With the exception of the Goda test, the other goodness-of-fit tests give very similar results. The Goda test is significantly less demanding, and the vast majority of the points across the globe satisfy this criteria. As seen in Fig. 10a, regions where the wave climate tends to form a single population satisfy the goodness-of-fit criteria. Areas such as the high latitudes and trade-wind belts tend to have similar meteorological conditions year round and hence result in CDFs that conform well to the (in this case) FT-1 distribution. In contrast, the tropics will have wave climates composed of tropical cyclones and larger-scale meteorological events, which are clearly separate populations. Similarly, the Arabian Sea is subject to intense local winds at the time of the Somali jet but otherwise experience locally light winds and Southern Ocean swell. These are quite distinct populations, and hence the goodness-of-fit test fails in this area. A more detailed analysis that separated these populations on a regional basis would be feasible but beyond the scope of this analysis.

Figure 8 shows the corresponding plots (to Fig. 6) of \( U_{10}^{100} \) (1° × 1°). As with \( H_{s}^{100} \), the IDM FT-1 result shows a relatively smooth distribution of wind speed with clear extreme wind belts at high latitudes of both hemispheres. The POT results (both W3P and GPD) are very noisy with a high degree of short-scale spatial variability. This
FIG. 6. Contour plots of the global distribution of $H_s^{100}$ (m) on a $1^\circ \times 1^\circ$ grid: (a) IDM method and FT1 distribution, (b) POT method and W3P distribution with 90% threshold, (c) POT method and GPD distribution with 90% threshold.
Fig. 7. As in Fig. 6 but on a $2^\circ \times 2^\circ$ grid.

(a) Global Values of $H_1$ computed using IDM method and FT1 Distribution

(b) Global Values of $H_2$ computed using POT method and W3P Distribution

(c) Global Values of $H_3$ computed using POT method and GPD Distribution
Fig. 8. Contour plots of the global distribution of $U_{10}^{100}$ (m s$^{-1}$) on a $1^\circ \times 1^\circ$ grid. (a) IDM method and FT-1 distribution, (b) POT method and W3P distribution with 90% threshold, (c) POT method and GPD distribution with 90% threshold.
FIG. 9. As in Fig. 8 but on a $2^\circ \times 2^\circ$ grid.

(a) Global Values of $U_{10}$, 100 computed using IDM method and FT1 Distribution

(b) Global Values of $U_{10}$, 100 computed using POT method and W3P Distribution

(c) Global Values of $U_{10}$, 100 computed using POT method and GPD Distribution
trend is stronger for the POT GPD results than the POT W3P. In both cases the results are clearly degraded by undersampling of extreme events. It is clear that variations in sampling density caused by satellite tracks can be seen. The $U_{10}^{100}$ (POT) results are poorer than the comparable $H_{10}^{100}$ (POT) values. This occurs because wind systems are generally spatially smaller than the waves they generate. Once generated, waves tend to propagate away from the generation source. Hence, it is necessary to have greater sampling density for wind speed than for wave height.

The $2^\circ \times 2^\circ$ results in Fig. 9 show some improvement in the spatial consistency of the POT results. Satellite track patterns are, however, still visible and, even with this larger averaging region, undersampling clearly degrades the results. Whereas the $2^\circ \times 2^\circ$ data for $H_{10}^{100}$ (POT) produced quite acceptable results, the corresponding $U_{10}^{100}$ (POT) results clearly indicate that a longer duration record, and possibly also greater spatial density of observations, is required to obtain reliable results for wind speed.

As for wave height, the many small-scale features are again evident with the IDM FT-1 approach. One notable difference between the wind speed and wave height results is the extent of “shadowing” down wind of land in the Southern Hemisphere. This was quite evident in the wave height results but is much less evident for wind speed. This indicates that the effect is caused by the reduction in wave fetch down wind of the land, rather than a reduction in the strength of extreme winds.

The Cramér–von Mises goodness-of-fit test acceptance results for $U_{10}^{100}$ (IDM) FT-1 are shown in Fig. 10b. The number of points that satisfy the criteria are much
reduced compared to wave height. This is consistent with the greater level of spatial and temporal variability of wind speed data. As a result, the data does not conform to the chosen CDF as well as for wave height. This is particularly the case in the equatorial and tropical regions of all oceanic basins. Of course, it must be remembered that this is a particularly demanding test and the Goda test was satisfied at most points. Again, resampling the data on a seasonal basis is likely to improve the goodness-of-fit.

As noted above, the spatial distributions of 100-yr return period values for both wind speed and wave height developed in this study with the IDM FT-1 are similar to those of Alves and Young (2003) and Chen et al. (2004). The magnitudes of $H_{100}$ (IDM) values also appear to be very similar to those of Alves and Young (2003). A quantitative comparison with the results of Chen et al. (2004) is not possible. However, a visual comparison of global extreme value charts for both wind speed and wave height indicate the Chen et al. (2004) estimates are 5%–10% lower than the present results. This presumably reflects the shorter dataset (less than half as long) and the different fitting method for the adopted FT-1 CDF used by Chen et al. (2004). One can, however, conclude that the present results are consistent with these previous studies but enhance the spatial resolution.

7. Discussion and conclusions

This analysis presents results from a much longer dataset of global altimeter observations than previously reported to determine extreme wind speed and wave height. The analysis shows that the IDM approach with an FT-1 distribution produces results with a mean error less than 5% for wave height and less than 10% for wind speed when compared to the same analysis conducted for buoy data (see $r_1$ values in Tables 3 and 4). The IDM FT-1 distribution also generally satisfies demanding goodness-of-fit tests on a global basis. The goodness-of-fit tests for the IDM FT-1 are also more often satisfied for wave height than wind speed. This indicates that the FT-1 distribution is a better approximation to the altimeter wave height data than the wind speed data. Such results reflect the nature of the present analysis in which no attempt has been made to separate wind and wave populations in geographical areas where multiple populations might exist. Because of the long duration of the present records, such an approach may be a possible extension of the present work.

Despite the better theoretical underpinning of the peaks-over-threshold analysis, this approach does not perform as well as the IDM. The POT analysis is much more sensitive to undersampling than the IDM analysis. As a result, the averaging area over which data is binned is an important consideration for the POT analysis. Even with the extended dataset, a bin size of $1\degree \times 1\degree$ consistently underestimates extreme values compared to buoy data. The larger $2\degree \times 2\degree$ POT analysis, which uses a significantly longer altimeter dataset, produces the first global plots of extreme significant wave height using this technique with an acceptable level of spatial variability. The POT wave height results do capture many of the global features evident in the IDM analysis. However, when applied to wind speed, the POT approach yields unacceptable results. These poor results are attributed to the greater spatial variability of wind speed compared to wave height. As a result, the dataset is still too small to capture sufficient “peaks” to adequately define the extreme tail of the CDF.

The present analysis does, however, show that, when applied to buoy data, the IDM and POT analyses yield comparable results. The analysis also shows that, given longer altimeter datasets, the POT analysis will produce acceptable extreme value altimeter estimates.

The theoretical limitations of the IDM center on the ability of the approach to accurately model the tail of the CDF and the validity of using data that may not be independent. As there is no absolute measure of the extreme values for comparative purposes, the ability of the IDM to model the tail cannot be absolutely determined. However, the approach gives results that are in excellent agreement with buoy data, which has greater temporal sampling density and hence more extreme data points in the analysis. Although comparison with buoy data does not provide an authoritative answer, it provides some level of confidence in the altimeter results. When applied to altimeter data, the issue of independence of the data is not considered a major issue. Altimeter passes through the $1\degree \times 1\degree$ bins occurred only once every few days, even when there were multiple satellites operational in the latter years of the dataset. As a result, the data will generally satisfy the requirement for independence.

The present analysis clearly shows that with approximately 23 years of data, the satellite altimeter can provide high quality estimates of extreme wind speed and wave height conditions on a global basis. The volume of the data also means that small-scale features can be resolved, opening up the possibility for regional studies of extreme wave climate. With the duration and sampling density of the present dataset, either the IDM or POT approaches can be used to obtain extreme-value wave height estimates comparable with buoys. As the duration of the dataset continues to grow, it is likely that the POT
approach, with its sounder theoretical underpinning, will replace the IDM approach for extreme wave height estimates. The greater spatial variability of wind speed, however, means that the altimeter data is still too sparse to produce POT extreme value results comparable to buoy estimates. For wind speed, only the IDM approach produces acceptable results, a situation that is likely to remain unless future satellite missions can provide greater spatial and temporal sampling density.

Acknowledgments. This work was financially supported by the Australian Research Council (ARC) through Linkage Grant LP0882422. The support of the ARC and our industry partner RPS MetOcean is gratefully acknowledged.

REFERENCES


