Global Ocean Extreme Wave Heights from Spatial Ensemble Data

ALICIA TAKBASH AND IAN R. YOUNG

Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria, Australia

(Manuscript received 11 April 2019, in final form 11 July 2019)

ABSTRACT

A novel approach to estimation of extreme value ocean significant wave height is investigated, in which data from adjacent regions are pooled to form a spatial ensemble. The equivalent duration of this ensemble region is the sum of the durations of the data pooled to form the ensemble. To create such a spatial ensemble, data from regions to be pooled must be independent and identically distributed. ERA-Interim reanalysis data are used to investigate the requirement of independent and identically distributed data on a global basis. As a result, typical spatial ensembles are defined for a number of regions of the world and the 100-yr return period significant wave height is calculated for these regions. It is shown that the method can result in a reduction in the confidence interval for such extreme value estimates of between 30% and 60%. The method is demonstrated both with ERA-Interim data and altimeter data.

1. Introduction

The determination of extreme value estimates of environmental parameters (e.g., the quantity with a probability of occurrence of 0.01 in any year or 100-yr return period) is commonly undertaken using long-term time series of the measured quantity. In the context of ocean waves, such analyses traditionally use wave buoy/wave staff data from specific locations and model or satellite remote sensing data over the global domain. Irrespective of the data source, the aim is to fit an extreme value probability distribution function (PDF) to the measured data and then extrapolate this to the desired probability level (e.g., 0.01 for the 100-yr event). Extrapolation is almost always required as the desired return period is longer than the data record. The accuracy of the extreme value estimate is dependent on how well the chosen analytical PDF fits the low probability tail of the PDF of the recorded data and the extent of the extrapolation. To lessen the extrapolation and hence reduce the confidence limits (CL) on the extreme value estimate, the recorded time series should be as long as possible.

In the case of ocean waves, the time series of measured buoy data in some locations is as long as 40 years (Evans et al. 2003). Model data can be of any length for which the model is run, although for records longer than approximately 30 years the quality of the wind fields forcing the model declines (Dee et al. 2011). The satellite altimeter time series is now 33 years long (Young et al. 2017; Ribal and Young 2019). Studies that have attempted to apply traditional peaks-over-threshold (PoT) extreme value analysis (EVA) approaches to such altimeter significant wave height data (e.g., Vinoth and Young 2011) have generally not been successful. This is because the time series have been too short to produce estimates of 100-yr wind speed and wave height with acceptable confidence limits. Recently, however, Takbash et al. (2019) used the full 30-yr altimeter record to produce the first acceptable global estimates of 100-yr return period wind speed U_{100}^{100} and significant wave height H_s^{100} using this approach. Nevertheless, the Takbash et al. (2019) results still showed spatial variability of estimates as a result of relatively high statistical variability of these estimates (i.e., large confidence limits).

All such studies of altimeter wave height data divide the world into spatial (grid) regions and pool all altimeter data in each grid region. These grid regions are then considered as independent observation regions, and EVA is applied independently for each region. This raises the question of whether it is possible to use multiple spatial regions to obtain greater confidence in the

Denotes content that is immediately available upon publication as open access.

Corresponding author: Ian R. Young, ian.young@unimelb.edu.au

extreme value estimates. Breivik et al. (2014) and Meucci et al. (2018) have addressed a similar problem using forecast model data. These forecast models run multiple ensembles predicting the future sea state, each initiated with slightly different initial conditions. Under certain conditions, they show that data from these ensemble forecasts can be pooled, creating a synthesized data series, the equivalent length of which is longer than the length of the individual ensemble datasets.

In the present analysis, the data are considered as a spatial ensemble. Criteria are developed that identify regions that can be pooled to create effective datasets the equivalent length of which is longer than the 30-yr record of the original data. As a result, confidence limits on the resulting estimates can be significantly reduced, resulting in greater statistical confidence in the values of H_s^{100} .

The paper is organized as follows. In section 2, a brief review of relevant studies of the estimation of global values of extreme wave heights is provided. Section 3 outlines criteria that must be met to form spatial ensembles of data. Results for the extreme value estimates of significant wave height using such spatial ensembles are provided in section 4, followed by conclusions in section 5.

2. Global estimates of extreme wave heights

As the length of the global altimeter time series has grown, an increasing number of studies have investigated the use of such data for global estimates of H_s^{100} . Given the limited length of the time series, initial attempts focused on the use of initial distribution methods (IDM) (Tucker 1991; Cooper and Forristall 1997; Teng 1998) in which a predefined PDF form is fitted to all recorded data. Altimeter studies that have adopted this approach include Alves and Young (2003), Chen et al. (2004), Challenor et al. (2005), Wimmer et al. (2006), and Vinoth and Young (2011). As the IDM approach uses the full PDF it is relatively stable when only short time series are available, but its modeling of the low probability tail of the distribution is poor and hence it can produce unreliable extreme value estimates (Takbash et al. 2019). Vinoth and Young (2011) first attempted to apply a more rigorously sound PoT approach in which the extreme tail of the PDF is modeled. However, with 23 years of data, they found results were unusable for wind speed but showed some promise when applied to significant wave height. Takbash et al. (2019) applied the PoT approach to a 30-yr altimeter record and obtained global distributions of U_{10}^{100} and H_s^{100} that were consistent with buoy data and varied spatially in a reasonably smooth manner.

Similarly, EVA of wind speed and significant wave height can be based on model data obtained from hindcasts or reanalyses (Aarnes et al. 2012, 2015; Caires and Sterl 2005). In this terminology, reanalysis is used to indicate that the model results include assimilation of measured data, whereas hindcasts do not include assimilation.

The European Centre for Medium-Range Weather Forecasts (ECMWF) has generated a series of increasingly sophisticated reanalyses. The first of these was the ECMWF 15-yr Re-Analysis (ERA-15; Gibson et al. 1997), covering the period 1979–93. ERA-40 (Uppala et al. 2005) covered the period 1957–2002. Several global extreme value analyses of significant wave height have been based on the ERA-40 dataset [e.g., the Royal Meteorological Institute of the Netherlands (KNMI) Atlas; Caires and Sterl 2005]. However, as demonstrated by Sterl and Caires (2005), ERA-40 model results generally underestimate wind speed and wave height extremes. The most commonly used reanalysis for EVA has been ERA-Interim (Dee et al. 2011), which covers the period from 1979 until 2018. Note that ERA-Interim (hereinafter ERAI) is scheduled to be phased out in August 2019 and replaced by the higher-resolution ERA-5. ERAI has been used to evaluate ocean extremes by Aarnes et al. (2012, 2015). However, ERAI still underestimates wind speed and wave height extremes, and according to Stopa and Cheung (2014) particular attention must be paid to the analysis of the upper percentiles of the data, which may not be well represented by the model.

a. The relevance of the confidence interval

The uncertainty in the estimation of extreme values is commonly represented in terms of confidence limits, where $CL^{0.025}$ and $CL^{0.975}$ represent the 95% lower and upper confidence limits, respectively, of the values of the 100-yr return period significant wave height H_s^{100} . The difference between these values, $CI^{0.95} = CL^{0.975} - CL^{0.025}$, is the 95% confidence interval; statistically, there is a 95% probability that the true value of the 100-yr return period significant wave height lies within this interval around the estimate H_s^{100} (Muir and El-Shaarawi 1986; Mathiesen et al. 1994). A primary concern in the estimation of extreme values is reducing the CI and hence increasing the confidence in the extreme value estimate (Mathiesen et al. 1994; Breivik et al. 2013, 2014; Breivik and Aarnes 2017; Meucci et al. 2018). As a result, a number of recent studies have examined approaches to reduce the CI (Bulgakov et al. 2018; Takbash et al. 2019; Meucci et al. 2018). The magnitude of the CI can be reduced by either increasing the number of observations of extreme events, thus improving

confidence that the tail of the PDF is well defined or by increasing the duration of the record, thus reducing the magnitude of the extrapolation to the desired probability level. The desire to increase the length of the data record is common to extreme value applications for in situ, model, and satellite remote sensing data and has been discussed in numerous studies (Caires and Sterl 2005; Mazas and Hamm 2011; Young et al. 2017; Takbash et al. 2019). In the context of satellite remote sensing data, increasing the amount of data points is achieved by combining data values observed over a wider region. This is typically achieved by pooling data from a region around a given point (Young 1994; Cooper and Forristall 1997; Alves and Young 2003; Vinoth and Young 2011). In pooling data in this manner, there is an implicit assumption that the additional data are not statistically independent and hence all data are assumed to apply to a point at the center of the region. Obviously, as the region over which data are pooled increases, there is a reduction in the spatial resolution (Chen et al. 2004; Vinoth and Young 2011).

Present-day weather prediction systems include a stochastic element to account for the intrinsic uncertainty in initial conditions by running an ensemble of forecasts, each initiated with slightly perturbed initial conditions, rather than a single deterministic forecast (Lewis 2005). Breivik et al. (2013, 2014) and Meucci et al. (2018) have taken advantage of the fact that at long lead time (9–10 days) these forecasts diverge to the point where they have low correlation. In such circumstances, each forecast in the ensemble potentially becomes an independent realization of a potential sea state. They show that provided the ensemble members are independent and identically distributed, they can be pooled to create a dataset with an equivalent duration much longer than the duration of the forecast time series. Using this approach, Meucci et al. (2018) created a dataset from ensemble forecasts equivalent to 750 years from a 6-yr archive taken between 2010 and 2016. As the equivalent duration of the dataset is longer than the desired return period (100 years), the extreme value estimates can be obtained without the need for extrapolation. This approach produced H_s^{100} estimates consistent with buoy data and with plausible spatial distributions. Importantly, because of the length of the synthesized dataset the $CI^{0.95}$ was reduced by more than 70% compared to a traditional “peaks over threshold” extreme value analysis (Meucci et al. 2018).

b. Spatial ensemble applied to extreme value analysis

Following the general concept developed by Breivik et al. (2013, 2014) and Meucci et al. (2018), the present study explores whether a spatial ensemble of data can be

used to reduce potential errors and the magnitude of confidence limits for estimates of H_s^{100} . Both the ERAI reanalysis model dataset and the altimeter dataset of Young et al. (2017) and Ribal and Young (2019) are used and the EVA methods described by Takbash et al. (2019) (i.e., PoT analysis). Rather than pooling independent model forecasts (Breivik et al. 2013, 2014; Meucci et al. 2018), we explore whether data from a number of independent spatial domains can be combined (pooled) to create a synthetic dataset with an equivalent duration that is the sum of the durations of the separate areas pooled.

To be able to pool data from spatial regions they must be independent and identically distributed (Goda 1988; Coles 2001; Breivik et al. 2013; Breivik et al. 2014). In the present context, these requirements become the following:

- 1) The regions must be far enough apart that the data from each of the regions are independently distributed (i.e., uncorrelated/poorly correlated). This essentially means that the extreme values are largely generated by different storms.
- 2) The wave climate in the regions to be pooled must be similar and representative of the larger aggregated region (i.e., identically distributed).

The present approach of pooling spatial ensembles has similarities to the Bayesian hierarchical models (Wikle et al. 1998) used to represent the spatial and temporal variations of H_s through conditional probabilities. These approaches have been used to examine trends in wave height by Vanem et al. (2012a,b).

3. Spatial ensemble data selection

As outlined above, regions can potentially be pooled for extreme value analysis if 1) wave heights between regions have low correlation and 2) the regions have comparable wave climate. This section will investigate these criteria globally.

a. Spatial coherence of waves

To assess the spatial coherence of wave height on a global basis, an approach similar to that adopted by Greenslade and Young (2005), for the analysis of anomaly correlation length scales, is used. In this approach, the aim is to determine the correlation coefficients between specified locations. The low spatial and temporal resolution of the altimeter data (Young et al. 2017), together with the irregular sampling, makes such data difficult to use for such an analysis. Therefore, as an alternative, ERAI reanalysis data (Dee et al. 2011) are investigated for this purpose. The ERAI wave height

FIG. 1. Correlation ellipses calculated at specified locations [monthly means subtracted from the time series for application in (1)].

data are available at 6-hourly intervals on a regular 0.75° spatial grid over the period 1984–2014 (Stopa and Cheung 2014). ERAI wave height data have been used in several studies to investigate climatology and/or variability of wave height as well as wave height extremes (Shanas and Kumar 2014; Shanas and Kumar 2015; Aarnes et al. 2015; Kumar et al. 2016; Young and Donelan 2018). As these studies show reasonable agreement between ERAI and altimeter data both in terms of the magnitude and spatial distribution of wave height, it is adopted here to determine spatial coherence (and climate).

Specific points (on the $0.75^\circ \times 0.75^\circ$ ERAI grid) in the Pacific, Atlantic, Indian, and Southern Ocean were selected, and the decorrelation length scales of wave height in both zonal and meridional directions were calculated by using the Pearson correlation coefficient (Pearson 1895; Rodgers and Nicewander 1988):

$$r(i, j) = \frac{\sum_{l=1}^m [H_s(i, l) - \overline{H_s(i)}][H_s(j, l) - \overline{H_s(j)}]}{\sqrt{\sum_{l=1}^m [H_s(i, l) - \overline{H_s(i)}]^2} \sqrt{\sum_{l=1}^m [H_s(j, l) - \overline{H_s(j)}]^2}} \quad (1)$$

In (1), the sample correlation coefficient between points i and j is $r(i, j)$ and the summations are conducted over all $l = 1:m$ observations of $H_s(l)$ at each location. The overbar terms in (1) represent mean quantities that have been evaluated in a number of manners, as described below. The spatial coherence of the wave height field was evaluated using three different approaches. Initially, the monthly variation (monthly mean) was removed from the time series, leading to deseasonalized time series. Thus, we obtain time series of fluctuations

about the seasonal mean (storms) and any long-term trend (Young and Ribal 2019). Values of $r(i, j)$ were then determined between spatially separated points. That is, from the selected point, $r(i, j)$ was calculated for adjacent points at successively larger spatial separation. This spatial separation was successively increased in all radial directions from the selected point. This process continued until $r(i, j)$ fell below 0.5. Values of constant $r(i, j)$ between 1 and 0.5 were calculated and drawn for representative points in Fig. 1. The resulting correlation ellipses (CEs) demonstrate the gradual decorrelation of the wave height field with distance for various regions of the globe. As one might expect, the CEs in Fig. 1 decay more rapidly in the meridional direction than in the zonal direction, indicating larger spatial correlation scales in the zonal direction. This result is similar to the findings of Greenslade and Young (2005).

The CEs reflect patterns of both swell and wind speed directions: either the predominant direction of propagation of the swell, or the predominant surface wind direction for local wind sea. In high wind speed zones (storm zones), the predominant wind speed direction aligns with the longest axis of the CE. This is clear for the Southern Ocean, where the correlation length scale is longest along the direction of the strong westerlies. For much of the global oceans, however, it is swell that dominates (Semedo et al. 2011) and the CE longest axis is approximately aligned with the swell crests. This is clear in the Indian and South Pacific Oceans. Here, the great circle propagation paths for swell radiating out from storms in the Southern Ocean align from southwest to northeast. The CE long axis is approximately perpendicular to these great circles, indicating higher spatial correlations in these directions. As one moves from south to north along the line of points through the central

FIG. 2. Correlation ellipses calculated at specified locations [long-term means subtracted from the time series for application in (1)].

Pacific Ocean the CEs change orientation as the wave climate changes from being dominated by Southern Ocean swell to being dominated by North Pacific swell.

The largest CE of these samples occurs in the eastern Pacific, where the wave field is influenced both by Southern Ocean swell and also southeasterly trade winds. Both of these wave conditions tend to result in CEs with a long axis aligned from northwest to southeast. As both swell and local winds reinforce this orientation, the resulting CE is relatively large.

Areas where local winds dominate the shapes of the CEs include the central North Atlantic where the CE long axis is aligned from southwest to northeast, the South Atlantic (off South America) where the trade winds result in a CE long axis aligned from northwest to southeast, and the Pacific off the Asian coast where the northeast trades result in a CE long axis aligned from southwest to northeast.

Where the trade winds from both hemispheres converge at the equator (the doldrums) the shape of the CE becomes symmetric, with the longest axis parallel to the equator. A decrease in anisotropy (see Greenslade and Young 2005) for the sampled CEs can be seen in the midlatitudes (centered on 54°N, 200°E) of the Pacific Ocean. The more circular shape of the CE reflects the anticlockwise movement of wind speed in cyclonic systems in the Northern Hemisphere.

In the second approach for calculation of the CEs, the long-term mean was subtracted from time series, rather than the monthly mean. As a result, the seasonal variation in the time series is retained and, hence the size of the CEs increases (Fig. 2). This is particularly the case in the midlatitudes of the Northern Hemisphere, where the seasonal variation is relatively large. At similar latitudes in the Southern Hemisphere, the seasonal variation is

much smaller (Young and Donelan 2018) and hence the CEs are similar in size to Fig. 1. The general shape and spatial variations of the CEs are, however, still similar to the case where the monthly means were removed (Fig. 1).

As the focus of the present work is on extreme wave heights, the third approach used only data greater than the 90th percentile. Again, the monthly mean of the values was subtracted before determining the correlation coefficients. The 90th percentile corresponds to the threshold which was subsequently used in the peak-over-threshold extreme value analysis. Therefore, this approach investigates the decorrelation scales of the storm events, rather than all the data. As the amount of data is significantly reduced by applying this threshold, the noise level increases in these calculations. However, the spatial distributions for these approaches remain very similar to Figs. 1 and 2, although the sizes of the CEs are reduced. The reduced correlation scale is as could be expected when considering only extreme conditions. That is, extreme conditions have shorter decorrelation scales than mean conditions. For the present application, large CEs represent a more demanding condition, as this limits the regions that can potentially be pooled for EVA. Therefore, the case shown in Fig. 1 is the more demanding test and is used in all future analysis.

To further illustrate the spatial variation of correlation, as represented by the CEs, data were considered at one selected CE in the North Atlantic. The location selected was centered on 30°N, 320.25°E. As shown in Fig. 1, at this point the CE has its long axis aligned from southwest to northeast. Scatterplots of the deseasonalized $H_s(i, j) - \bar{H}_s$ between this point and neighboring points on an approximate 12° circle around the point are

FIG. 3. Scatterplots of deseasonalized significant wave height $H_s(i, j) - \overline{H_s}$ between a location at 30°N, 320.25°E and locations surrounding that position at an approximate 12° radius; $\overline{H_s}$ was calculated as the monthly mean. Each panel shows the data scatter, a 1:1 linear relationship line, and the correlation coefficient $r(i, j)$. Data are obtained from ERA-Interim.

shown in Fig. 3. Each of the nine panels in this figure shows the scatterplot for $H_s(i, j) - \overline{H_s}$, a 1:1 linear relationship line for the data and the value of $r(i, j)$. Consistent with Fig. 1, the largest values of $r(i, j)$ lie along the southwest–northeast diagonal (panels 7, 5, 3) and the smallest values along the northwest–southeast diagonal (panels 1, 5, 9). The figure also clearly shows the reduction in the magnitude of extremes (variations from the seasonal mean) moving from north to south and the similarity of these extremes in the zonal

direction (same latitude). It is also clear that the probability distribution of H_s is skewed, with maximum values further above the mean (zero value in the figure) than minimum values are below the mean.

To illustrate the decorrelation of the storm events at this same location, Fig. 4 shows similar scatterplots but for data above the 90th percentile. Panel 1 shows $r(i, j)$ between the location 30°N, 320.25°E and 30°N, 332.25°E (i.e., the location 12° east of the point). This corresponds to panel 6 of Fig. 3 for the mean conditions. Comparison

FIG. 4. Scatterplots of deseasonalized storm significant wave height $H_s(i, j) - \bar{H}_s$ between locations $30^\circ\text{N}, 332.25^\circ\text{E}$ and $30^\circ\text{N}, 332.25^\circ\text{E}$ (same location as panel 6 of Fig. 3). Only data above the 90th percentile are included in the analysis to simulate storm conditions. The data at location $30^\circ\text{N}, 332.25^\circ\text{E}$ in panel 2 have been time-lagged by 24 h to account for the time of storm propagation. Each panel shows the data scatter, a 1:1 linear relationship line, and the correlation coefficient $r(i, j)$. Data are obtained from ERA-Interim.

of the figures shows that $r(i, j)$ reduces from 0.49 for the mean conditions to 0.19 for the storm conditions (i.e., data above 90th percentile). These correlation coefficient calculations consider data at the same times at each of the pairs of points under consideration. As storms propagate in time, it is possible that higher correlation coefficients may result if $r(i, j)$ is determined with a time lag applied at location j . This is investigated in panel 2 of Fig. 4. In this panel, the second location j has been lagged by 24 h, relative to location i . As expected, the value of $r(i, j)$ increases to 0.42 when the data are time lagged in this manner. Importantly, however, comparison of Figs. 3 and 4 shows that the decorrelation scales of the storm waves are shorter than the mean conditions. (0.42 compared to 0.49). Testing at a range of locations showed that between points separated by 12° , as in Figs. 3 and 4, a time lag of 24 h produced the largest values of $r(i, j)$. A lag time of 24 h corresponds to a storm propagation speed of approximately 50 km h^{-1} , which seems reasonable. Also, as shown in Figs. 3 and 4, other locations showed that the storm waves are always more poorly correlated (smaller values of r) than the mean conditions.

There is no absolute level of $r(i, j)$ at which one can state that the regions are sufficiently decorrelated to pool. Following Meucci et al. (2018), we have adopted the criterion $r(i, j) < 0.5$. Under this condition, all neighboring locations in Fig. 3, with the exception of locations 3 and 7 (long axis of the CE), can be potentially pooled with the central location 5. That is, they are deemed to be

sufficiently decorrelated that they provide independent storm information. We have used $r(i, j)$ calculated from all the data as our criteria to determine whether the data are sufficiently decorrelated to pool, even though our interest is in storm conditions. This choice was made as this parameter is a more stable measure. As shown in Fig. 4, this would generally result in values of $r(i, j) < 0.4$ when storm waves are considered. Hence, using the mean conditions produces a conservative result.

b. Spatial variation of wave climate

The second criterion for ensemble data pooling requires that areas which have low correlation, still have a similar sea state (or wave climate). This was investigated by determining the relative percentage differences of both the mean monthly significant wave height and mean monthly 99th percentile significant wave height between the two locations:

$$\overline{\text{RPD}}(i, j) = \frac{1}{12} \sum_{k=1}^{12} \frac{|\bar{H}_s(i, k) - \bar{H}_s(j, k)|}{\bar{H}_s(i, k)}, \quad (2)$$

$$\text{RPD}^{99}(i, j) = \frac{1}{12} \sum_{k=1}^{12} \frac{|H_s^{99}(i, k) - H_s^{99}(j, k)|}{H_s^{99}(i, k)}. \quad (3)$$

In (2) and (3), $\overline{\text{RPD}}(i, j)$ and $\text{RPD}^{99}(i, j)$ are the mean and 99th percentile relative percentage differences between locations i and j respectively; $\bar{H}_s(i, k)$ is the mean monthly significant wave height for location i and month k and $H_s^{99}(i, k)$ is the 99th percentile significant wave

FIG. 5. QQ plots of significant wave height $H_s(i, j) - \bar{H}_s$ between a location at 30°N, 320.25°E and locations surrounding that location at an approximate 12° radius; $\bar{H}_s$ is calculated as the monthly mean. Each panel shows the QQ plot, a least squares linear fit to the QQ data, and the relative percentage differences $\overline{RPD}(i, j)$ and $RPD^{99}(i, j)$. Data are obtained from ERA-Interim.

height for location i and month k . As the requirement is that there is a similar wave climate at locations i and j in order for them to be spatially pooled, the requirement has been set that both $\overline{RPD}(i, j)$ and $RPD^{99}(i, j)$ are less than 0.1. That is, both the mean and 99th percentile conditions differ by less than 10%. This ensures similarity of both the mean conditions and the extremes, which are obviously important for EVA.

Figure 5 shows quantile–quantile (QQ) plots between the same locations shown in Fig. 3 (i.e., North Atlantic).

In each panel a linear fit to the QQ data is shown and the values of $\overline{RPD}(i, j)$ and $RPD^{99}(i, j)$. The variation in both the mean wave climate and the extremes are clear in moving from south to north (meridional direction). Locations where the best-fit line to the QQ data is significantly different from 45° signify differing wave climate. Further deviation between the locations are also clear at the upper percentiles in the figure (i.e., extreme sea states differ). Under the conditions that $\overline{RPD}(i, j) < 0.1$ and $RPD^{99}(i, j) < 0.1$, only the locations

in the zonal direction, locations 4 and 6, meet this criterion. The similarity of the wave climates between locations 4, 5, and 6 is also clear in the scatterplot, Fig. 3. As locations 4 and 6 meet all conditions set [$r(i, j) < 0.5$, $\overline{RPD}(i, j) < 0.1$, $RPD^{99}(i, j) < 0.1$] they could be pooled with location 1 to form a spatial ensemble for EVA. It is possible that further points in the zonal direction could also be pooled, making the spatial ensemble larger. Importantly, this would also further increase the effective duration of the ensemble pooled region. In this case, however, no points in the meridional direction can be pooled, given the changing wave climate in this direction.

4. Determination of extreme significant wave height from selected spatial ensembles

a. Selection of spatial ensemble regions

With the information on the global spatial variation of $r(i, j)$ and $RPD(i, j)$ provided above, the aim is now to define regions that satisfy the criteria set for both of these [i.e., $r(i, j) < 0.5$ and $RPD < 0.1$]. Note that for simplicity, we write $RPD(i, j)$ to signify both $\overline{RPD}(i, j)$ and $RPD^{99}(i, j)$.

The process used to define regions is shown diagrammatically in Fig. 6. Note that this process to define areas for spatial pooling is based on ERAI reanalysis data. Once the regions are defined, however, it can be applied to either ERAI or altimeter data to determine extreme value H_s .

- (i) An initial location 1 is defined. We proceed in the zonal direction until a location, 2, is found for which $r(1, 2) < 0.5$. We then check that $RPD(1, 2) < 0.1$. If this condition is satisfied, locations 1 and 2 can be pooled for the analysis.
- (ii) From location 2 we continue to move in the zonal direction until location 3 is identified, where $r(2, 3) < 0.5$. We check that the conditions $RPD(2, 3) < 0.1$ and $RPD(1, 3) < 0.1$ are satisfied. If both of these conditions are met, then locations 1, 2, and 3 can be pooled. This process continues in the zonal direction until the conditions are no longer met.
- (iii) With the extent of the region in the zonal direction defined, we then explore the extent in the meridional direction. Returning to location 1, we move in the meridional direction to location 4, where $r(1, 4) < 0.5$ and $RPD(1, 4) < 0.1$. However, we also need to check that the RPD criteria are met for the other combinations of locations, that is, $RPD(2, 4) < 0.1$ and $RPD(3, 4) < 0.1$. If all conditions are met, location 4 is added to the region.
- (iv) We then return to location 2, and again move in the meridional direction to identify 5, where $r(2, 5) <$

FIG. 6. The schema used to define regions for spatial ensemble pooling for various oceanic basins.

0.5 and $r(4, 5) < 0.5$. Again, all RPD values for all combinations of locations are checked.

- (v) The process then returns to location 3, and location 6 is identified in the same manner.
- (vi) Note that, for simplicity, the above description and Fig. 6 consider only locations north of the origin point 1. In reality this same process is mirrored south of the point as well, with all cross-checks for $r(i, j)$ and $RPD(i, j)$ in the whole region considered.

As RPD increases much more rapidly in the meridional direction than the zonal direction, this process tends to define regions with a much greater zonal extent than meridional extent.

Figure 7a shows a selection of spatial regions for various oceanic basins that can be pooled using these criteria. As expected, the regions tend to be elongated in the zonal direction, reflecting the similar wave climates that are found at the same latitudes. That is, these shapes tend to be determined by the RPD criteria rather than the r condition. In addition, the spatial extent of the regions is larger at high latitudes, reflecting the greater spatial extent of meteorological systems at these latitudes. The region with the largest spatial extent is the Southern Ocean, reflecting the relatively uniform wave climate in this area (Young 1999; Young and Donelan 2018; Semedo et al. 2011).

b. Spatial ensemble analysis of extremes

With representative areas defined by the above analysis, the aim is to pool the data for these regions and undertake extreme value analysis on the pooled data to determine 100-yr return period significant wave heights H_s^{100} . This process was applied to both the ERAI

FIG. 7. (a) Ensemble spatial regions for ERAI data with values of H_s^{100} (m) marked. The upper and lower confidence limits on values of H_s^{100} are shown by the superscripts and subscripts, respectively. (b) Ensemble spatial regions for altimeter data with values of H_s^{100} (m) marked. The upper and lower confidence limits on values of H_s^{100} are shown by the superscripts and subscripts, respectively. (c) Global values of H_s^{100} (m) obtained with a PoT analysis and a GPD distribution. Data are obtained from altimeter missions.

TABLE 1. Values of H_s^{100} (in parentheses) and 95% confidence intervals ($CI^{0.95}$, in italics) for four selected ensemble regions (see Fig. 7a). The values of H_s^{100} and $CI^{0.95}$ for each of the individual subareas pooled to create the ensemble region are also shown. Values are shown of ERAI data in the left columns and altimeter data in the right columns.

Ensemble region name	ERAI: (H_s^{100}) and $CI^{0.95}$ (m)				Altimeter: (H_s^{100}) and $CI^{0.95}$ (m)			
	PN1	PN2	PE	SP	PN1	PN2	PE	SP
Subarea								
Area 1	(13.8) 1.6	(17.8) 2.8	(4.7) 0.3	(13.1) 1.5	(13.8) 2.9	(17.9) 5.7	(5.8) 1.3	(13.3) 2.1
Area 2	(12.5) 1.4	(17.4) 2.4	(5.0) 0.3	(12.7) 1.1	(12.7) 3.0	(17.4) 5.2	(5.5) 0.8	(13.3) 2.5
Area 3		(17.3) 2.2		(12.3) 1.1		(18.3) 5.4		(15.1) 3.9
Area 4		(15.3) 1.9		(13.0) 1.3		(16.0) 4.4		(12.7) 1.7
Area 5				(12.6) 1.0				(12.9) 2.1
Area 6				(12.6) 1.1				(13.0) 2.2
Area 7				(11.7) 1.1				(12.3) 2.5
Ensemble region	(13.2) 1.0	(17.2) 1.2	(4.9) 0.2	(12.6) 0.5	(13.4) 2.1	(17.7) 2.7	(5.7) 0.7	(13.3) 0.9
Percent reduction in $CI^{0.95}$	33%	50%	30%	58%	30%	50%	30%	61%

reanalysis data and the altimeter data. In the case of the ERAI data, the total duration of the original dataset is 30 years (1984–2014). As each location to be pooled has a data duration of 30 years, the equivalent duration for the pooled regions will be integral multiples of 30 years. With the selection criteria used here, the pooled regions had an equivalent duration of between 60 and 210 years (mostly 90 years).

The process was also undertaken for the altimeter data. In this case, the original time series again spans 1984–2014, although the effective duration is 27 years, as no satellites were operation from 1987 to 1990, effectively removing approximately 3 years from the analysis. As a result of the spatial pooling, the resulting effective duration of the pooled areas was between 54 years (two adjacent subareas pooled) and 189 years (seven adjacent subareas pooled).

Following Takbash et al. (2019), both sets of pooled data were analyzed using a peaks-over-threshold analysis with a threshold set at the 90th percentile. A generalized Pareto distribution (GPD) was fitted to the data and extrapolated (or interpolated if the equivalent duration of the data was longer than 100 years) to the 100-yr return period probability level. Figure 7a shows values of H_s^{100} for each pooled region, calculated from ERAI data. The equivalent result from the altimeter data is shown in Fig. 7b. For reference purposes, global values of H_s^{100} are shown contoured from the original altimeter $2^\circ \times 2^\circ$ data, with each 2° region considered independently (i.e., no pooling). The values of H_s^{100} for all cases are comparable. Consistent with Takbash et al. (2019), the present results show the largest values of $H_s^{100} \approx 18$ m in the North Atlantic and North Pacific. The Southern Ocean shows extensive regions of extreme waves heights, with $H_s^{100} \approx 14$ m around the globe at latitudes greater than 40° S. Equatorial regions show much lower values with $H_s^{100} \approx 4$ m. As pointed out by Takbash et al. (2019) both the altimeter data and the ERAI model reanalysis will

underestimate tropical cyclone activity, and hence these equatorial values will be underestimated. The ERAI results are consistently lower than the altimeter data, consistent with previous studies which have shown that this dataset underestimates extremes (Stopa and Cheung 2014). The pooled altimeter data (Fig. 7b) produce very similar values of H_s^{100} to the 2° data (Fig. 7c).

c. Confidence intervals

An examination of Fig. 7c shows that there is clear statistical noise in the estimates of H_s^{100} brought about by the limited number of observations available to estimate the tail of the PDF and the need to extrapolate to the desired extreme probability level. This results in a relatively large confidence interval and the spatial variability evident in Fig. 7c. The spatial ensembles shown in Figs. 7a and 7b are aimed at reducing these CIs.

To calculate the 95% CLs for the resulting estimates of H_s^{100} a bootstrap method (Efron 1979; Caires 2007, 2011; Qi 2008; Breivik and Aarnes 2017; Aarnes et al. 2012; Meucci et al. 2018) was applied. In this approach, we compute a sample of 1000 bootstrapped H_s^{100} estimates taken randomly from the original data sample. For each sample, H_s^{100} was determined and 2.5 percentile and 97.5 percentile values calculated to give the lower and upper 95% confidence limits ($CL^{0.025}$ and $CL^{0.975}$), respectively. The confidence interval is given by $CI^{0.95} = CL^{0.975} - CL^{0.025}$. These values were calculated for each of the spatial ensemble regions and each of the subareas that were pooled to create the ensemble regions. The values of $CL^{0.025}$ and $CL^{0.975}$ are shown for each ensemble region in Figs. 7a and 7b. Table 1 shows values of $CI^{0.95}$ for four selected ensemble areas, as well as the subareas that were pooled to create the ensemble regions. The ensemble areas considered in Table 1 are marked in Fig. 7a for reference. Results are

shown both for ERAI and altimeter data. Numerical values of H_s^{100} , together with the 95% lower and upper confidence limits ($CL^{0.025}$ and $CL^{0.975}$) are shown on Figs. 7a and 7b for each ensemble region.

Table 1 shows that the values of $CI^{0.95}$ for the ERAI subarea data are smaller than for the corresponding altimeter data. This occurs because the ERAI time series is slightly longer (30 years compared to 27 years). In addition, there is less variability in the data from the model compared to the altimeter measurements. This results in more stable estimates of the tail of the PDF with less variability and hence smaller CIs.

The spatial ensemble pooling results in CIs that are between 30% and 60% smaller than the original data. The magnitude of the reduction increases as the number of subareas making up the spatial ensemble increases. The Southern Ocean/southern Pacific (SP) is the area where it was possible to pool the largest number of subareas to create the ensemble and this results in an approximately 60% reduction in the CI. In contrast, in the North Pacific (PN1) and eastern Pacific (PE) it was possible to pool only two subareas, resulting in an approximately 30% reduction in the CI. Farther north in the Pacific (PN2), the spatial correlation scale increases and it is possible to pool four subareas, with a reduction in CI by 40%.

Although the spatial ensemble process can reduce the statistical variability in the extreme value estimates, it has no impact on any tail bias in the PDF of the data used. As noted previously, the ERA-Interim data underestimate extremes and hence the values of H_s^{100} in Fig. 7a are smaller than the corresponding altimeter values in Fig. 7b. Bias correction techniques can be used to address such issues (e.g., Cannon et al. 2015); however, these have not been applied here.

5. Conclusions

The present study investigates whether data from spatial areas can be pooled to create an ensemble data series, the equivalent length of which is longer than that of the individual areas. Such spatial ensembles of data are then subjected to extreme value analysis to determine 100-yr return period significant wave height. Following Breivik et al. (2013, 2014) and Meucci et al. (2018) we show that in order to pool such data, the areas pooled must be independent and identically distributed. In the present context, independence is achieved by only considering regions that are poorly correlated (i.e., influenced by separate storms). The requirement that the data be identically distributed was assessed by requiring that both the monthly means and monthly 99th percentiles between the areas were in good agreement (comparable wave climate).

Spatial correlation and climate were assessed globally using ERAI reanalysis data. This showed that spatial regions with a long axis in the zonal direction could be pooled to form spatial ensembles. The size of these regions varies by geographic region, with the largest (longest) regions being in the Southern Ocean.

This technique of forming spatial ensembles was applied to both ERAI and altimeter data. The resulting 100-yr return period significant wave heights were similar in magnitude to conventional analyses but have confidence intervals that are reduced by between 30% and 60%. That is, there is greater statistical confidence in the resulting extreme value estimates.

Acknowledgments. IRY gratefully acknowledges the support of the Australian Research Council through Grants DP130100215 and DP160100738. This support has been invaluable in completing this extensive study. The raw altimeter datasets used in the study were supplied by *Globwave* (altimeter and buoy) and are archived on the Australian Ocean Data Network (AODN) (<https://portal.aodn.org.au/>).

REFERENCES

- Aarnes, O. J., Ø. Breivik, and M. Reistad, 2012: Wave extremes in the northeast Atlantic from ensemble forecasts. *J. Climate*, **25**, 1529–1543, <https://doi.org/10.1175/JCLI-D-11-00132.1>.
- , S. Abdalla, J. R. Bidlot, and Ø. Breivik, 2015: Marine wind and wave height trends at different ERA-Interim forecast ranges. *J. Climate*, **28**, 819–837, <https://doi.org/10.1175/JCLI-D-14-00470.1>.
- Alves, J. H. G. M., and I. R. Young, 2003: On estimating extreme wave heights using combined Geosat, TOPEX/Poseidon and ERS-1 altimeter data. *Appl. Ocean Res.*, **25**, 167–186, <https://doi.org/10.1016/j.apor.2004.01.002>.
- Breivik, Ø., and O. J. Aarnes, 2017: Efficient bootstrap estimates for tail statistics. *Nat. Hazards Earth Syst. Sci.*, **17**, 357–366, <https://doi.org/10.5194/nhess-17-357-2017>.
- , —, J. R. Bidlot, A. Carrasco, and Ø. Saetra, 2013: Wave extremes in the northeast Atlantic from ensemble forecasts. *J. Climate*, **26**, 7525–7540, <https://doi.org/10.1175/JCLI-D-12-00738.1>.
- , —, S. Abdalla, J.-R. Bidlot, and P. A. E. M. Janssen, 2014: Wind and wave extremes over the world oceans from very large forecast ensembles. *Geophys. Res. Lett.*, **41**, 5122–5131, <https://doi.org/10.1002/2014GL060997>.
- Bulgakov, K., Kuzmin, V., and Shilov, D., 2018, Evaluation of extreme wave probability on the basis of long-term data analysis. *Ocean Sci.*, **14**, 1321–1327, <https://doi.org/10.5194/OS-14-1321-2018>.
- Caires, S., 2007: Extreme wave statistics: Confidence intervals. Tech rep., Delft Hydraulics, prepared for Rijkswaterstaat, Rijksinstituut voor Kust en Zee, 32 pp., <http://resolver.tudelft.nl/uuid:8d38ef9c-ea44-4b9d-850c-d4dd2e71a34f>.
- , 2011: Extreme value analysis: Wave data. JCOMM Tech. Rep. 57, 33 pp., <http://hdl.handle.net/11329/367>.

- , and A. Sterl, 2005: 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. *J. Climate*, **18**, 1032–1048, <https://doi.org/10.1175/JCLI-3312.1>.
- Cannon, A. J., S. R. Sobie, and T. Q. Murdock, 2015: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *J. Climate*, **28**, 6938–6959, <https://doi.org/10.1175/JCLI-D-14-00754.1>.
- Challenor, P. G., W. Wimmer, and I. Ashton, 2005: Climate change and extreme wave heights in the North Atlantic. *Proc. 2004 Envisat and ERS Symp.*, Salzburg, Austria, European Space Agency, SP-572.
- Chen, G., S. W. Bi, and R. Ezraty, 2004: Global structure of extreme wind and wave climate derived from TOPEX altimeter data. *Int. J. Remote Sens.*, **25**, 1005–1018, <https://doi.org/10.1080/01431160310001598980>.
- Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, 208 pp.
- Cooper, C. K., and G. Z. Forristall, 1997: The use of satellite altimeter data to estimate the extreme wave climate. *J. Atmos. Oceanic Technol.*, **14**, 254–266, [https://doi.org/10.1175/1520-0426\(1997\)014<0254:TUOSAD>2.0.CO;2](https://doi.org/10.1175/1520-0426(1997)014<0254:TUOSAD>2.0.CO;2).
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Efron, B., 1979: Bootstrap methods: Another look at the jackknife. *Ann. Stat.*, **7**, 1–26, <https://doi.org/10.1214/aos/1176344552>.
- Evans, D., C. Conrad, and F. Paul, 2003: Handbook of automated data quality control checks and procedures of the National Data Buoy Center. NOAA National Data Buoy Center Tech. Document 03–02, 44 pp.
- Gibson, J., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF, 72 pp., <https://www.ecmwf.int/en/elibrary/9584-era-description>.
- Goda, Y., 1988: On the methodology of selecting design wave height. *Proc. 21st Conf. on Coastal Engineering*, Malaga, Spain, ASCE, 899–913.
- Greenslade, D. J. M., and I. R. Young, 2005: The impact of altimeter sampling patterns on estimates of background errors in a global wave model. *J. Atmos. Oceanic Technol.*, **22**, 1895–1917, <https://doi.org/10.1175/JTECH1811.1>.
- Kumar, P., S. K. Min, E. Weller, H. Lee, and X. L. Wang, 2016: Influence of climate variability on extreme ocean surface wave heights assessed from ERA-Interim and ERA-20C. *J. Climate*, **29**, 4031–4046, <https://doi.org/10.1175/JCLI-D-15-0580.1>.
- Lewis, J. M., 2005: Roots of ensemble forecasting. *Mon. Wea. Rev.*, **133**, 1865–1885, <https://doi.org/10.1175/MWR2949.1>.
- Mathiesen, M., Y. Goda, P. J. Hawkes, E. Mansard, M. J. Martín, E. Peltier, E. F. Thompson, and G. Van Vledder, 1994: Recommended practice for extreme wave analysis. *J. Hydraul. Res.*, **32**, 803–814, <https://doi.org/10.1080/00221689409498691>.
- Mazas, F., and L. Hamm, 2011: A multi-distribution approach to POT methods for determining extreme wave heights. *Coast. Eng.*, **58**, 385–394, <https://doi.org/10.1016/j.coastaleng.2010.12.003>.
- Meucci, A., I. R. Young, and Ø. Breivik, 2018: Wind and wave extremes from atmosphere and wave model ensembles. *J. Climate*, **31**, 8819–8842, <https://doi.org/10.1175/JCLI-D-18-0217.1>.
- Muir, L. R., and A. H. El-Shaarawi, 1986: On the calculation of extreme wave heights: A review. *Ocean Eng.*, **13**, 93–118, [https://doi.org/10.1016/0029-8018\(86\)90006-5](https://doi.org/10.1016/0029-8018(86)90006-5).
- Pearson, K., 1895: Note on regression and inheritance in the case of two parents. *Proc. Roy. Soc. London*, **58**, 240–242, <https://doi.org/10.1098/rspl.1895.0041>.
- Qi, Y., 2008: Bootstrap and empirical likelihood methods in extremes. *Extremes*, **11**, 81–97, <https://doi.org/10.1007/s10687-007-0049-8>.
- Ribal, A., and I. R. Young, 2019: 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Sci. Data*, **6**, 77, <https://doi.org/10.1038/s41597-019-0083-9>.
- Rodgers, J. L., and W. A. Nicewander, 1988: Thirteen ways to look at the correlation coefficient. *Amer. Stat.*, **42**, 59–66, <https://doi.org/10.1080/00031305.1988.10475524>.
- Semedo, A., K. Suselj, A. Rutgersson, and A. Sterl, 2011: A global view on the wind sea and swell climate and variability from ERA-40. *J. Climate*, **24**, 1461–1479, <https://doi.org/10.1175/2010JCLI3718.1>.
- Shanas, P. R., and V. S. Kumar, 2014: Temporal variations in the wind and wave climate at a location in the eastern Arabian Sea based on ERA-Interim reanalysis data. *Nat. Hazards Earth Syst. Sci.*, **14**, 1371–1381, <https://doi.org/10.5194/nhess-14-1371-2014>.
- , and —, 2015: Trends in surface wind speed and significant wave height as revealed by ERA-Interim wind wave hindcast in the central Bay of Bengal. *Int. J. Climatol.*, **35**, 2654–2663, <https://doi.org/10.1002/joc.4164>.
- Sterl, A., and S. Caires, 2005: Climatology, variability and extrema of ocean waves: The web-based KNMI/ERA-40 wave atlas. *Int. J. Climatol.*, **25**, 963–977, <https://doi.org/10.1002/joc.1175>.
- Stopa, J. E., and K. F. Cheung, 2014: Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis. *Ocean Modell.*, **75**, 65–83, <https://doi.org/10.1016/j.ocemod.2013.12.006>.
- Takbashi, A., I. R. Young, and O. Breivik, 2019: Global wind speed and wave height extremes derived from long-duration satellite records. *J. Climate*, **32**, 109–126, <https://doi.org/10.1175/JCLI-D-18-0520.1>.
- Teng, C. C., 1998: Long-term and extreme waves in the Gulf of Mexico. *Proc. Conf. on Ocean Wave Kinematics and Loads on Structures*, Houston, TX, ASME, 342–349.
- Tucker, M. J., 1991: *Waves in Ocean Engineering*. Ellis Horwood, 431 pp.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, <https://doi.org/10.1256/qj.04.176>.
- Vanem, E., A. B. Huseby, and B. Natvig, 2012a: A Bayesian hierarchical spatio-temporal model for significant wave height in the North Atlantic. *Stochastic Environ. Res. Risk Assess.*, **26**, 609–632, <https://doi.org/10.1007/s00477-011-0522-4>.
- , B. Natvig, and A. B. Huseby, 2012b: Modelling the effect of climate change on the wave climate of the world's oceans. *Ocean Sci. J.*, **47**, 123–145, <https://doi.org/10.1007/s12601-012-0013-7>.
- Vinoth, J., and I. R. Young, 2011: Global estimates of extreme wind speed and wave height. *J. Climate*, **24**, 1647–1665, <https://doi.org/10.1175/2010JCLI3680.1>.
- Wikle, C. K., L. M. Berliner, and N. Cressie, 1998: Hierarchical Bayesian space-time models. *Environ. Ecol. Stat.*, **5**, 117–154, <https://doi.org/10.1023/A:1009662704779>.
- Wimmer, W., P. Challenor, and C. Retzler, 2006: Extreme waveheights in the North Atlantic from altimeter data. *Renew. Energy*, **31**, 241–248, <https://doi.org/10.1016/j.renene.2005.08.019>.

- Young, I. R., 1994: Global ocean wave statistics obtained from satellite observations. *Appl. Ocean Res.*, **16**, 235–248, [https://doi.org/10.1016/0141-1187\(94\)90023-X](https://doi.org/10.1016/0141-1187(94)90023-X).
- , 1999: Seasonal variability of the global ocean wind and wave climate. *Int. J. Climatol.*, **19**, 931–950, [https://doi.org/10.1002/\(SICI\)1097-0088\(199907\)19:9<931::AID-JOC412>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0088(199907)19:9<931::AID-JOC412>3.0.CO;2-O).
- , and M. A. Donelan, 2018: On the determination of global ocean wind and wave climate from satellite observations. *Remote Sens. Environ.*, **215**, 228–241, <https://doi.org/10.1016/j.rse.2018.06.006>.
- , and A. Ribal, 2019: Multi-platform evaluation of global trends in wind speed and wave height. *Science*, **364**, 548–552, <https://doi.org/10.1126/science.aav9527>.
- , E. Sanina, and A. V. Babanin, 2017: Calibration and cross validation of a global wind and wave database of altimeter, radiometer, and scatterometer measurements. *J. Atmos. Oceanic Technol.*, **34**, 1285–1306, <https://doi.org/10.1175/JTECH-D-16-0145.1>.