AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER

10.1002/2014GL060997

Key Points:

- A new data set of unprecedented size yields more reliable wind and wave returns
- The return values are significantly higher than those found with ERA-Interim
- The method is also applicable to parameters like temperature and precipitation

Supporting Information:

- Readme
- Supporting information and Figures S1–S10

Correspondence to:
Ø. Breivik,
oyvind.breivik@ecmwf.int

Citation:
Breivik, Ø., O. J. Aarnes, S. Abdalla, J.-R. Bidlot, and P. A. E. M. Janssen (2014), Wind and wave extremes over the world oceans from very large ensembles, *Geophys. Res. Lett.*, *41*, 5122–5131, doi:10.1002/2014GL060997.

Received 26 JUN 2014
Accepted 11 JUL 2014
Accepted article online 19 JUL 2014
Published online 29 JUL 2014

Wind and wave extremes over the world oceans from very large ensembles

Øyvind Breivik[1], Ole Johan Aarnes[2], Saleh Abdalla[1], Jean-Raymond Bidlot[1], and Peter A.E.M. Janssen[1]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK, [2]Norwegian Meteorological Institute, Bergen, Norway

Abstract Global return values of marine wind speed and significant wave height are estimated from very large aggregates of archived ensemble forecasts at +240 h lead time. Long lead time ensures that the forecasts represent independent draws from the model climate. Compared with ERA-Interim, a reanalysis, the ensemble yields higher return estimates for both wind speed and significant wave height. Confidence intervals are much tighter due to the large size of the data set. The period (9 years) is short enough to be considered stationary even with climate change. Furthermore, the ensemble is large enough for nonparametric 100 year return estimates to be made from order statistics. These direct return estimates compare well with extreme value estimates outside areas with tropical cyclones. Like any method employing modeled fields, it is sensitive to tail biases in the numerical model, but we find that the biases are moderate outside areas with tropical cyclones.

1. The ECMWF Integrated Forecast System Ensemble

Return values for wind and waves are fundamental to assessing the risks associated with human activities at sea, but their computation is complicated by the paucity of observational records. A number of reanalyses and hindcasts have appeared over the past two decades, both global [*Kalnay et al.*, 1996; *Uppala et al.*, 2005; *Onogi et al.*, 2007; *Rienecker et al.*, 2011; *Dee et al.*, 2011; *Saha et al.*, 2013] and regional, e.g., *Wang and Swail* [2001, 2002]; *Weisse and Günther* [2007]; *Breivik et al.* [2009]; *Reistad et al.* [2011]; *Wang et al.* [2012]. However, with few exceptions [*Compo et al.*, 2011; *Hersbach et al.*, 2013], the time series are much shorter than the 100 year return period typically sought and any extremes estimated from them come with wide confidence intervals [see *Aarnes et al.*, 2012; *Breivik et al.*, 2013, the latter hereafter B13].

The Integrated Forecast System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) has been producing daily ensemble forecasts since 1992 [*Buizza et al.*, 2007] and has been coupled to the ECMWF version of the WAM wave model since 1998 [*Janssen*, 2004]. Fifty-one-member ensemble forecasts have been issued twice daily (00 and 12 UTC) since March 2003. Even though the forecast skill has steadily been improving over the years [*Magnusson and Källén*, 2013], ensemble members at +240 h lead time (ENS240 hereafter) still tend to be weakly correlated with one another, and storm events are virtually uncorrelated, as demonstrated by B13. Although the goal of forecasting will always be to raise the skill, perhaps somewhat paradoxically, such weak correlations are a necessity when using ensemble forecasts for extreme value estimation because the entries must be independent [*Coles*, 2001]. That the tail of the ensemble can be considered independent was demonstrated by B13 where memberwise return estimates were shown to give very similar results to return estimates from the entire ensemble. This also showed that the correlation between consecutive +240 h forecasts is negligible.

Although the forecast skill must be low for the ensemble to be used for return value estimation, the empirical forecast distribution must nevertheless closely match the observed distribution all the way to the tail to yield realistic return values. Hence, the tail bias, i.e., the difference between observed and modeled upper percentiles, should be as low as possible. We find very good agreement with the upper percentiles of buoy wave and wind observations and ENVISAT altimeter wind speed measurements (see Figures S1 and S2 in the supporting information). It is also of interest to compare the tail behavior of ENS240 with ERA-Interim (ERA-I) as it is one of the most widely used reanalyses and also because it is a version of IFS

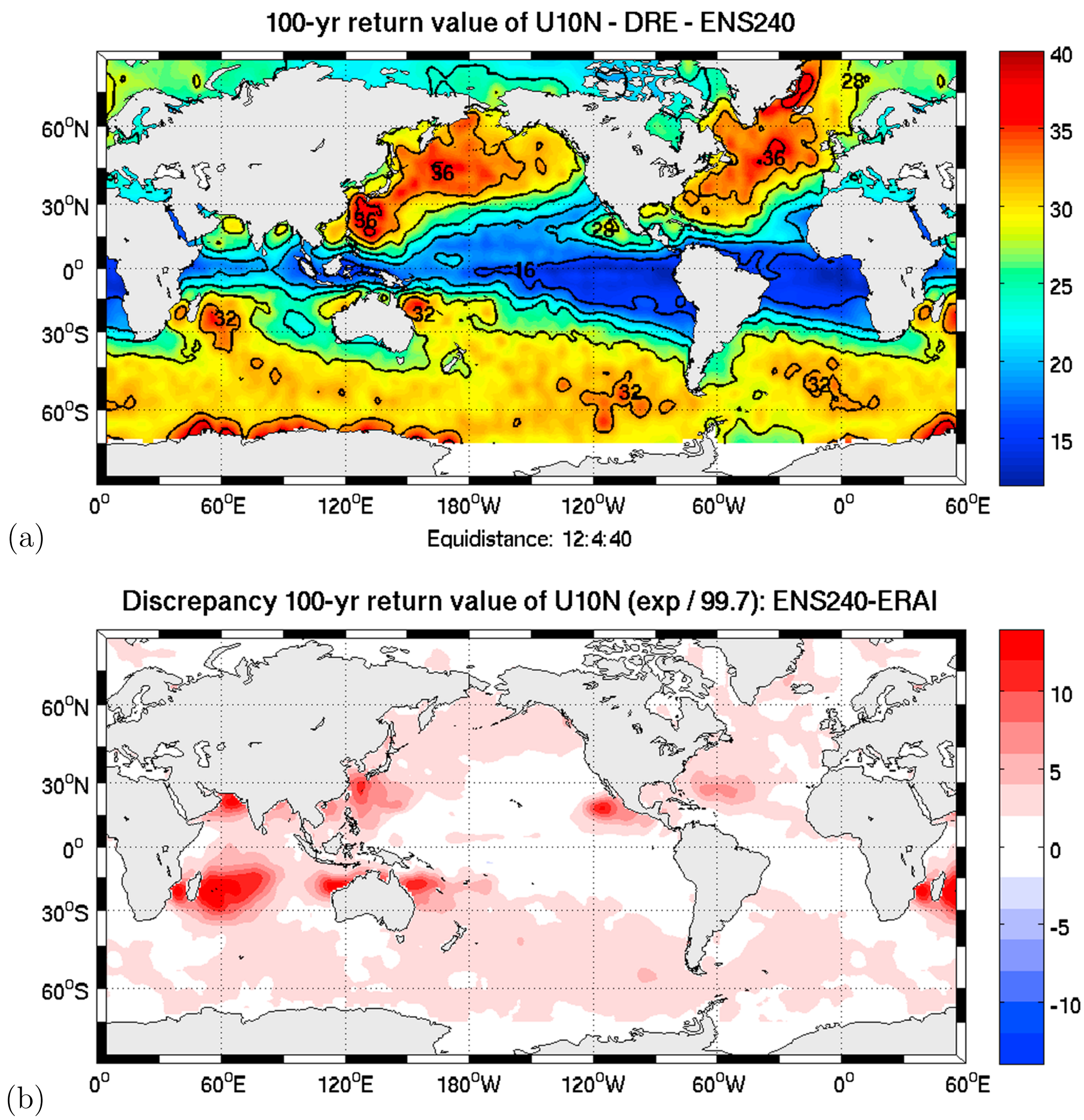


Figure 1. Ten meter wind speed 100 year return values, U_{100}^{DRE} (m s^{-1}). (a) ENS240 estimate direct return estimate. (b) Difference between ENS240 and ERA-I, exponential distribution fit, and threshold 99.7th percentile.

(Cycle 31r2) operational from December 2006 until June 2007 [*Dee et al.*, 2011]. It thus serves as a benchmark against which to compare the ENS240 climatology. We find that ENS240 exhibits smaller biases than ERA-I compared with altimeter wind speed measurements, with ERA-I typically being biased low by more than -1.5 m s^{-1} in certain regions and about -0.5 m s^{-1} globally. This is in contrast to ENS240 which is virtually unbiased when averaged globally and rarely deviates more than $\pm$ 0.5 m s^{-1} locally (see Figure S2b).

2. Estimating Return Values From Forecast Ensembles

So far we have two conditions that must be met for the ensemble to be useful for estimating probabilities of nonexceedance; (i) that the forecasts are uncorrelated and (ii) that the empirical distribution function closely matches the observations all the way to the tail. To these we add that (iii) the model climatology must be statistically stationary. Since the archive spans numerous model cycles, this is not trivially true, but B13 found no spurious trends in the mean and variance over the period compared to a reforecast ensemble from a fixed model cycle [*Hagedorn et al.*, 2012]. One final condition must be met to go from

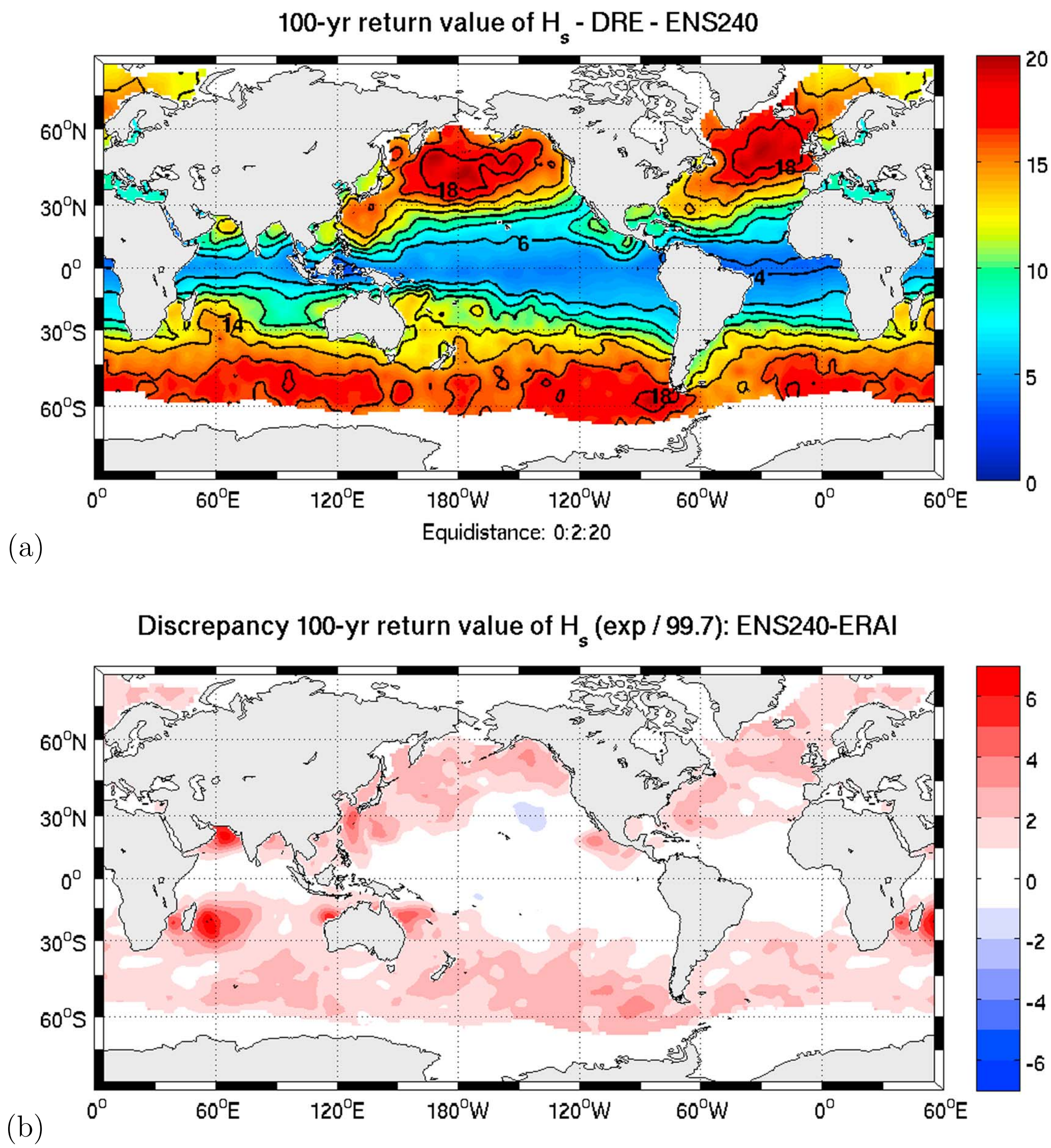


Figure 2. Significant wave height 100 year return values, H_{100} (m). (a) ENS240 direct return estimate. (b) Difference between ENS240 and ERA-I, exponential distribution fit, and threshold 99.7th percentile. Grid points that are ice covered more than 20% of the time are censored.

a probability of nonexceedance to a return period: (iv) that individual forecasts are representative of a synoptic (6 h) time interval. This allows us to treat 330,000 forecasts as the equivalent of 229 yrs of data (see Appendix).

Theoretical extreme value distributions are parametric estimates to be fitted to the modeled or observed maxima that are normally obtained from a continuous time series with a fixed temporal resolution. Annual maxima (AM) belong to the family of blocked maxima [*Coles*, 2001] and should follow the Generalized Extreme Value distribution. Threshold exceedances follow the Generalized Pareto (GP) distribution [*Coles*, 2001], where in the case of correlated time series a peaks-over-threshold (POT) method must be applied to ensure that the maxima are independent of each other (see B13). With the initial distribution method (IDM), all data are used. Usually, the method is employed when data are scarce, but sensitivity to the mean and variance of the series [*Lopatoukhin et al.*, 2000] as well as a lack of theoretical justification for a choice of distribution [*World Meteorological Organization*, 1998] makes the method highly empirical and somewhat difficult to use. The choice of data selection method (AM, threshold exceedances or IDM) and the choice of

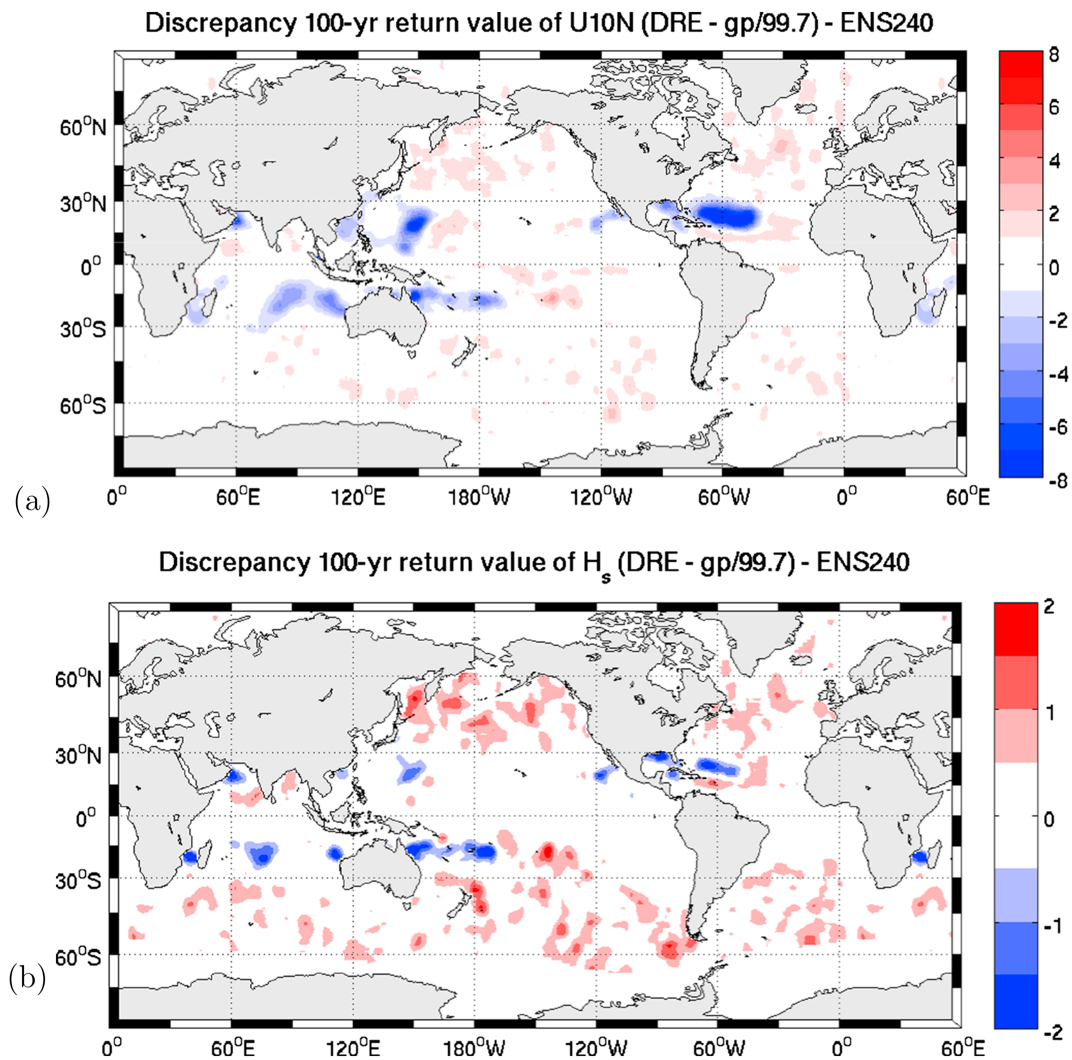


Figure 3. Nonparametric direct return estimates versus parametric GP estimates. (a) The difference between the U_{100}^{DRE} direct return estimate and the U_{100}^{GP} GP estimate for ENS240 (m s^{-1}). The differences are greatest in the regions prone to tropical cyclones where GP has a positive shape parameter. (b) Same as Figure 3a but for H_{100} (m). Note that grid points that are ice covered more than 20% of the time are censored for H_s.

location and shape parameters for the extreme value distribution can have a profound impact on the estimates [*Coles*, 2001; *Aarnes et al.*, 2012], especially with time series significantly shorter than the return period in question.

Because the ENS240 data set is larger than the return period sought, it is possible to avoid extreme value analysis altogether and compute a nonparametric direct return estimate (DRE) from the tail of the empirical distribution function (see Appendix). Figures 1a and 2a show maps of the 100 year return values of U_{100}^{DRE} and H_{100}^{DRE}, respectively. Estimates of U_{100} and H_{100} based on GP and the special case of the exponential distribution (EXP) closely resemble DRE in the extratropics (see Figure 3 and Figure S10), with U_{100}^{GP} estimates about 1.0 m s^{-1} below U_{100}^{DRE}. Similarly, H_{100}^{GP} estimates are found to be within ± 0.5 m of H_{100}^{DRE} for most of the extratropics (Figure 3b). However, in the tropics the parametric estimates deviate significantly from DRE. The differences match closely areas visited by tropical cyclones, see, e.g., *Oouchi et al.* [2006]. EXP will yield significantly lower estimates compared with DRE, while GP does the opposite. This is illustrated by Figure 4 where the confidence interval of U_{100}^{EXP} fails to encompass the extremal values for a location in the subtropics. Tropical cyclones are underestimated in global numerical weather prediction models [*Magnusson et al.*, 2014], thus the GP estimates come closer to the real return values in the affected areas, albeit still with large

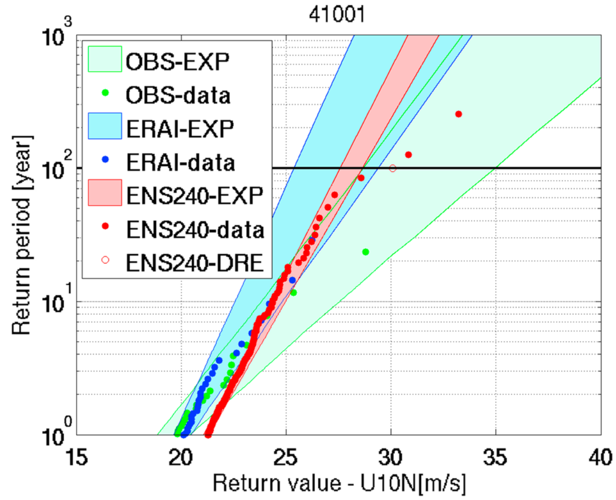


Figure 4. A comparison of the exponential distribution confidence intervals of U_{100} for buoy 41001, located east of Florida at 34.7°N, 072.7°W. Bootstrapped (500 resamples) confidence intervals are shown for in situ measurements from 1980 to 2012 with some gaps (green), ERA-I (blue), and ENS240 (red). The direct return estimate is shown as a red open circle. The horizontal black line indicates the 100 year return level. As expected the confidence intervals decrease with magnitude of the data sets, but fail to include the highest data points. A GP distribution with a shape parameter ξ different from zero would fit the data better. This is a feature particularly prominent in areas where the tail of the wind speed distribution is influenced by tropical cyclones.

uncertainties and a large negative bias. It is interesting to note in passing that the highest wind speed found throughout the entire ensemble is close to 50 m s^{-1} while H_s comes close to 25 m (see Figure S5).

3. Comparison With Traditional Return Estimates

Because ERA-I and the in situ observational time series are substantially shorter than the return period sought, we compare with ENS240 in terms of parametric return value estimates. ERA-I EXP estimates tend to exhibit a more stable behavior than GP estimates and will be used in the following (see also supporting information). In the extratropics the ENS240 U_{100} estimates are 2–4 m s^{-1} higher than ERA-I (Figure 1b), while H_{100} estimates are 1–3 m higher (Figure 2b). The situation is quite different in areas with tropical cyclones. Here we find much larger differences (more than 14 m s^{-1} for U_{100}, see Figure 1b, and more than 6 m for H_{100}, see Figure 2b). However, it is clear that the coarse resolution of the atmospheric model still renders tropical cyclones much too weak [*Magnusson et al.*, 2014] compared to the maximum sustained wind speed observed under such weather systems [*Elsner et al.*, 2008]. This means that our wind speed estimates in the subtropics and the tropics should be considered lower bounds. It is worth remembering at this point, though, that the winds reported here must be considered averages over much longer periods (6 h) than the 10 min typically investigated in relation to tropical cyclones [*Bidlot et al.*, 2002] (see supporting information).

ENS240 is assumed to represent the equivalent of approximately 229 years. The size of the data set reduces the confidence intervals to less than one third the width of ERA-I confidence intervals (34 year data set), see Figures 5 and 6. In general, ENS240 U_{100}^{EXP} and H_{100}^{EXP} compare reasonably well with POT estimates (see Appendix) of in situ observations (less than 10% difference except in areas with tropical cyclones, see supporting information). These differences are within the expected uncertainties since return value estimates from observational time series of relatively short length (10–30 years) come with large confidence intervals. This is clearly seen for buoy 41001 east of Cape Hatteras (Figure 4), where U_{100}^{EXP} estimates from ENS240 reveal yields 95% confidence intervals of 1 m s^{-1} while ERA-I has a width of about 4 m s^{-1} and the observations (1980–2012 with some gaps) in excess of 6 m s^{-1}.

How do our return values compare with previous investigations? U_{100}^{DRE} and H_{100}^{DRE} exhibit the same geographical features as found by *Caires and Sterl* [2005] who calibrated the earlier ERA-40 reanalysis [*Uppala et al.*, 2005] to in situ measurements of U_{10m} and H_s. However, their H_{100} fields are up to 6 m higher in the storm tracks in the North Atlantic and the North Pacific compared to what we find. No confidence intervals are provided, but a comparison of decadalwise return values shows approximately 10% spread in U_{10m} and H_s in the extratropical storm tracks. Estimates of U_{100} and H_{100} based on aggregated altimeter measurements from satellite missions over the past 30 years [*Vinoth and Young*, 2011; *Young et al.*, 2012] also yield much higher return values for wind speed in the extratropics. Wind speeds in excess of 52 m s^{-1} in the North Atlantic and the North Pacific are reported [*Vinoth and Young*, 2011; *Young et al.*, 2012] using IDM. This is almost 50% higher than what we find and is likely to be an artifact of IDM, something which is also hinted at by the fact that the GP estimates [*Vinoth and Young*, 2011] fail to reproduce these extremes. For H_{100} the results are much closer to our estimates. Since in the extratropics ENS240 compares well with in

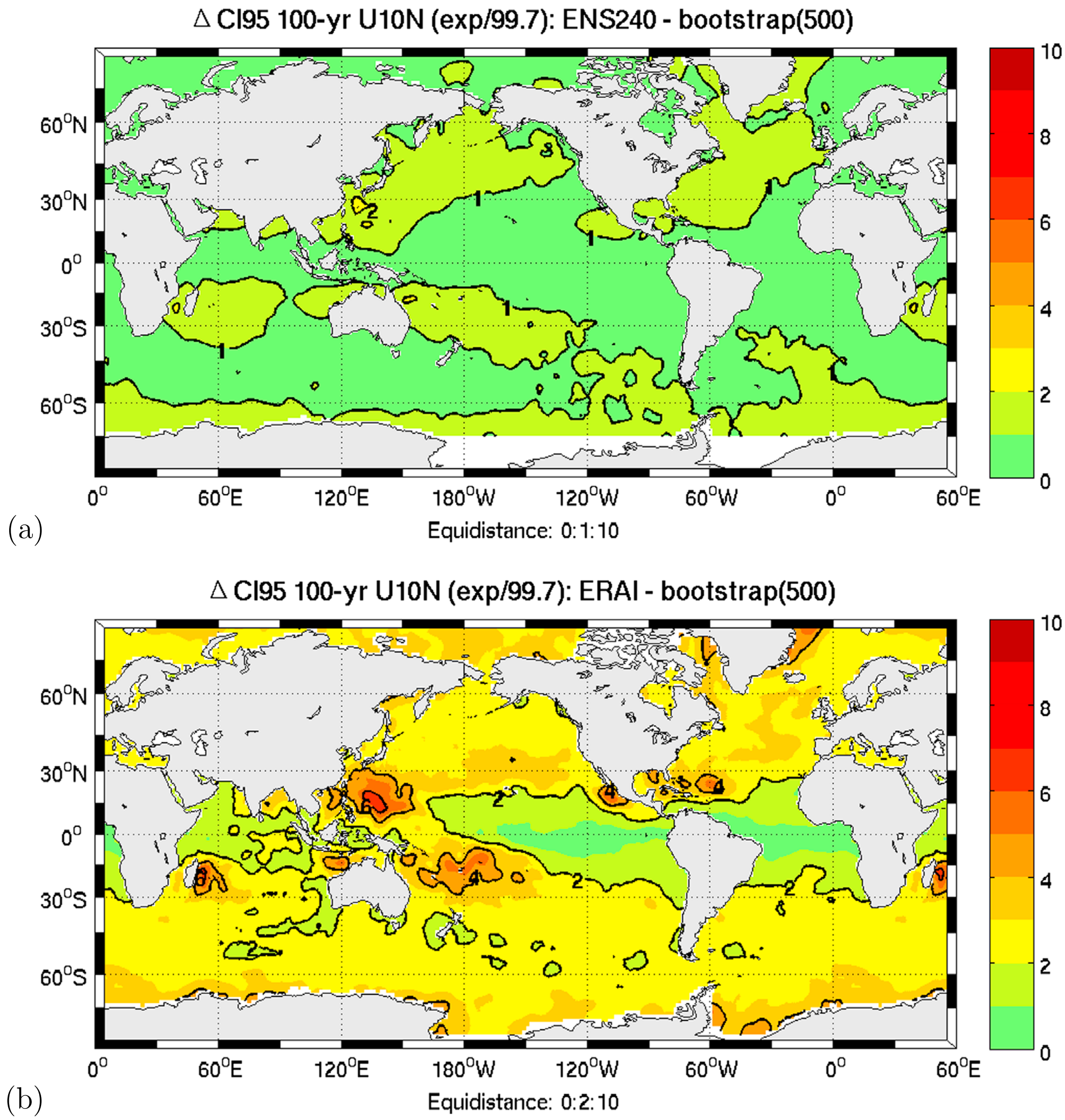
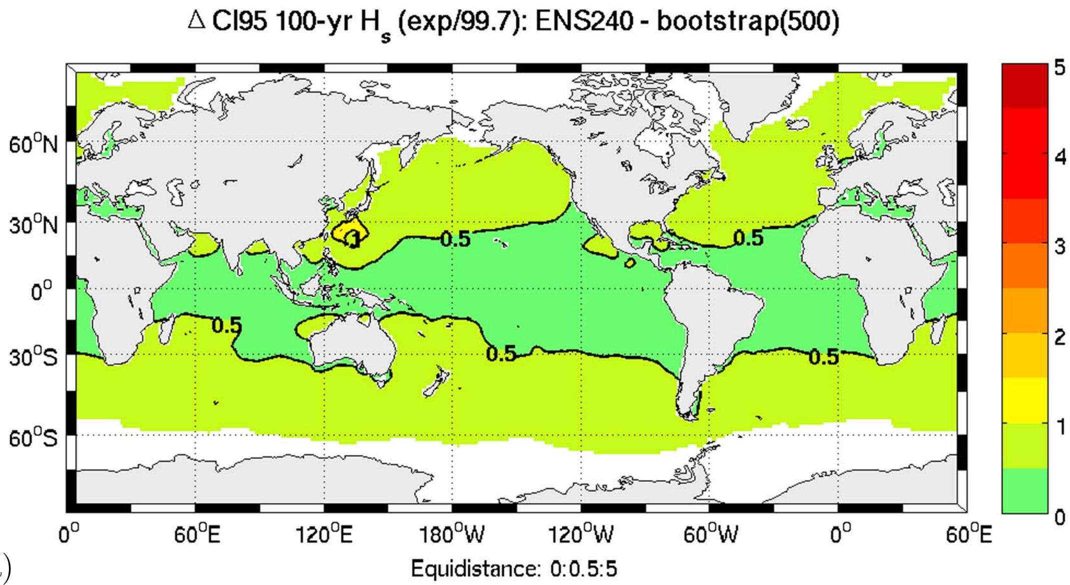


Figure 5. ENS240 versus ERA-I width of U_{100}^{EXP} confidence intervals. Estimates are fitted to the EXP distribution and confidence intervals are computed using bootstrapping of 500 resamples. ENS240 has confidence intervals about one third the width of ERA-I due to the much larger data set. (a) Width of the ENS240 U_{100}^{EXP} 95% confidence interval. All data exceeding $P_{99.7}$ were used. (b) Same as Figure 5a but for ERA-I. A peaks-over-threshold method was used to select data exceeding a threshold of $P_{99.7}$.

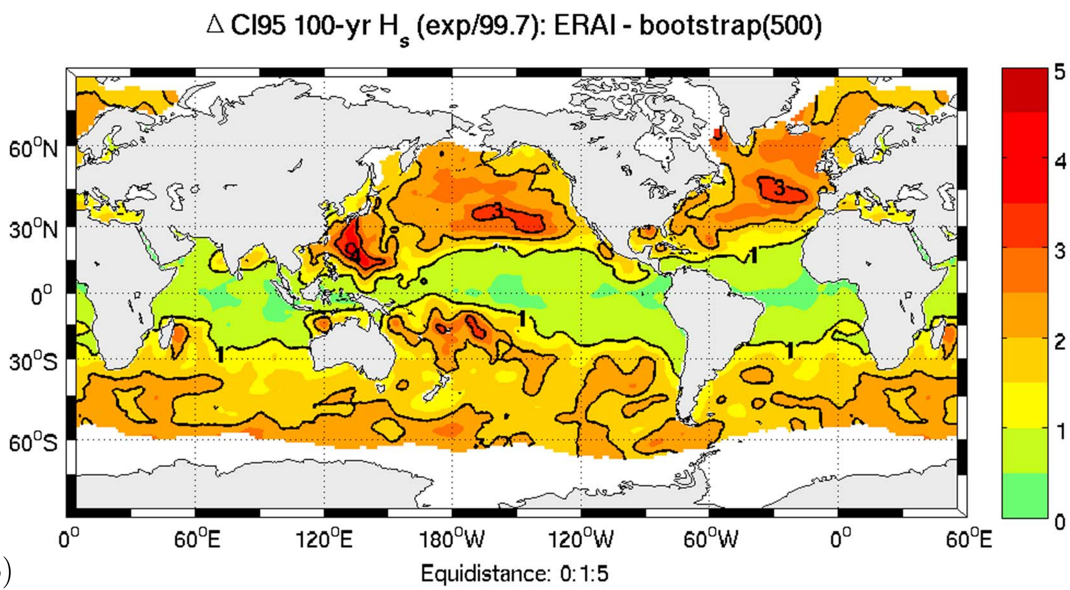
situ observations (both $U_{10\mathrm{m}}$ and H_s) and altimeter wind speed, we conclude that previous estimates [*Caires and Sterl*, 2005; *Vinoth and Young*, 2011; *Young et al.*, 2012] are probably too high outside the subtropics and much too low in the tropics and subtropics (as are our estimates). Note that *Caires and Sterl* [2005] calibrated against 3 h averaged observations, whereas *Vinoth and Young* [2011] and *Young et al.* [2012] assumed that the altimeter observations represented 3 h averages. This will lead to somewhat higher return values since we assume that the values are representative of 6 h interval. However, this can only explain a small fraction of the discrepancy (see Figure S4).

We note also that the investigation of 20 year return values of H_s from altimeter measurements by *Izaguirre et al.* [2011] yielded a positive shape parameter (unbounded extreme values) in areas with tropical cyclones, but a negative shape parameter (bounded extremes) in the extratropics. Their Figure 1 yields a map which is qualitatively very similar to Figure 3a.

 10.1002/2014GL060997



(a)



(b)

Figure 6. ENS240 versus ERA-I width of H_{100}^{EXP} confidence intervals. (a) Width of the ENS240 H_{100}^{EXP} 95% confidence interval. All data exceeding $P_{99.7}$ were used. (b) Same as Figure 6a but for ERA-I. A peaks-over-threshold method was used to select data exceeding a threshold of $P_{99.7}$.

4. Discussion

Implicit in the term "return value" lies the tacit assumption that the probability of nonexceedance is drawn from a long, stationary time series [*Coles*, 2001]. In practice, observed and modeled time series vary in length from 30 to 100 years and are thus much shorter than the period over which stationarity is assumed, and with the exceptions of *Compo et al.* [2011], *Wang et al.* [2012], and *Hersbach et al.* [2013], the time series are also substantially shorter than the 100 year return period. Decadal variations in storminess may affect these estimates [*Caires and Sterl*, 2005; *Young et al.*, 2012], and climate projections of future wind and wave climate [*Wang and Swail*, 2006; *Debernard and Røed*, 2008; *Hemer et al.*, 2013] suggest that nonstationary methods must be applied to assess future return values (see *Kharin and Zwiers* [2005]; *Izaguirre et al.*, [2010, 2011], for examples of nonstationary extreme value analysis). There is disagreement about the presence of trends in wave extremes in the extratropics [*Wang and Swail*, 2002; *Caires and Sterl*, 2005; *Young et al.*, 2012], but there is some evidence [*Young et al.*, 2012] that over the past three decades there has been a positive trend in extreme wind speed globally as well as in the intensity of tropical cyclones [*Elsner et al.*, 2008]. The trends in extreme wind speed [*Young et al.*, 2012] are, however, based on 10 year chunks of altimeter measurements, and the uncertainties are very high. This is one of the strengths of using large aggregates of ensemble forecasts. These will, by construct, not exhibit long-term trends and low-frequency oscillations since the initial conditions cover a period of 9 years. Stationarity is thus not an issue, suggesting that a similar approach could be used on short time slices (about 10 years) from ensembles covering several decades to avoid the complicating factors of using nonstationary methods. Candidate data sets that approach the size required would be reforecast ensembles [*Hagedorn et al.*, 2012] and the recently completed 20th century model 10-member ensemble integrations [*Hersbach et al.*, 2013]. For climate projections, multimodel ensembles [*Hemer et al.*, 2013] could provide the amount of data necessary for extreme trend estimates.

Very large ensembles allow us to sidestep the choice of parametric distribution as we can compute direct 100 year nonparametric return estimates. In principle, by looking at forecast ranges (briefly investigated by B13) instead of a fixed lead time, it would be possible to perform a peaks-over-threshold analysis and increase the data set by an order of magnitude. This would also remove any ambiguity concerning the representative interval of forecasts as the forecast range would determine the equivalent length of the data set. Candidates for such an analysis would be, for example, the ECMWF monthly forecast system [*Vitart et al.*, 2008], the extended forecast range of IFS from day 10 to day 15, and the seasonal forecast system [*Stockdale et al.*, 2011; *Molteni et al.*, 2011].

Using very large ensembles at advanced lead times is a novel method previously only explored for H_s in the Northeast Atlantic (see B13). The method, like any method employing modeled fields, is sensitive to tail biases in the numerical model [*Magnusson et al.*, 2014]. These biases must be assessed for each new parameter, but in our case we find that they are moderate outside areas with tropical cyclones and significantly lower than those found for ERA-I. On the other hand, the unsystematic error associated with the parameter fitting to an extreme value distribution [*Coles*, 2001] is directly related to the size of the data set and thus benefits immediately from an increase in size. The method is general and holds the promise of making much larger data sets available for a wide range of oceanographic, hydrological, and meteorological parameters that are routinely forecast and archived.

Appendix A: Methods Summary

The ENS240 forecasts were interpolated onto a regular 1° × 1° grid. All +240 h forecasts (two per day, 00, and 12 UTC, 51 ensemble members in each) between 26 March 2003 and 25 March 2012 were used. ERA-I analyses (1979–2012) were interpolated onto the same grid. The current spectral truncation of ENS240 is T639 for the atmospheric model, corresponding to approximately 32 km, whereas the wave model ECWAM is run at approximately 55 km. Previous model cycles had coarser resolution, see supporting information and B13. The 10 m neutral wind speed was extracted for both ENS240 and ERA-I. This is the field used to force the wave model and is thus consistent with the H_s fields investigated. Grid points that are ice covered in more than 20% of the forecasts were censored from the analysis of H_s. No such censoring is required for U_{10m}.

We compared GP and EXP estimates and decided to use EXP for our comparison of ENS240 return values with ERA-I and in situ observations (see supporting information). The threshold was set to the 1000th highest forecast, corresponding to $P_{99.7}$. This choice (see supporting information) was made after investigating the stability of EXP and GP estimates as a function of the threshold. Since ensemble forecasts are assumed uncorrelated, all values exceeding the threshold were used. For ERA-I and the in situ observations, the threshold was also set to $P_{99.7}$ and a POT technique applied where peaks must be separated by 48 h to ensure that all entries are independent [*Lopatoukhin et al.*, 2000; *Coles*, 2001]. The distribution fitting was done with the maximum likelihood method, and confidence intervals were estimated by bootstrapping. We found that 500 resamples were sufficient to yield a stable mean and a stable 95% confidence interval (see supporting information). The threshold was set high enough to avoid summertime values. This means that in locations with strong seasonality in wind speed and direction, our estimates should be interpreted as wintertime return values. We make no attempt to stratify the data set by season or direction.

Under the assumption that the collection of +240 h forecast members are equivalent to a temporal period (each forecast representative of a 6 h interval, see supporting information) we may convert our collection of ensemble forecasts into an equivalent time series. The two daily ensembles of +240 h forecasts from 50 perturbed ensemble members plus the unperturbed control member aggregated over 9 years can be considered to represent

$$9\,\mathrm{yr} \times 365.25\,\frac{\mathrm{days}}{\mathrm{yr}} \times 2\,\frac{\mathrm{forecasts}}{\mathrm{day}} \times 6\,\mathrm{h} \times 51\,\mathrm{mem} = 2,011,797\,\mathrm{h} = 229.5\,\mathrm{yr}.$$

We can therefore make nonparametric direct return estimates from the ensemble of the 100 year return value, x_{100}^{DRE}, without invoking a theoretical extreme value distribution (see supporting information).

The assumption that the modeled wind and wave fields are representative of a 6 h interval is based on considerations of the horizontal model resolution [*Bidlot et al.*, 2002]. To test the sensitivity of the return value estimates to this assumption, we have reduced the interval to 4 h, which is close to the averaging period used for the observations ($\pm$2 h). This reduction leads to an increase of at most 0.6 m for H_{100} and 1.2 m s^{-1} for U_{100} using the EXP distribution (Figure S4). This suggests that the method is reasonably insensitive to the choice of interval but a little conservative.

All return estimates were computed marginally, i.e., per grid point. After that all fields were smoothed using a 5° × 5° two-dimensional box car filter after applying a smaller (3° × 3°) MAX filter which selects the highest value from the nearest neighbors. This counters the tendency of the smoothing to lower the highest return values.

Acknowledgments

This work has been supported by the Research Council of Norway through the project “Wave Ensemble Prediction for Offshore Operations” (WEPO, grant 200641) and through the European Union FP7 project MyWave (grant 284455). S.A. was funded by the European Space Agency (contract 21519/08/I-OL, “Technical support for global validation of Envisat data products”). This study has also been part of a PhD program for O.J.A., partially funded by the Norwegian Centre for Offshore Wind Energy (NORCOWE). All data sets presented in this study are archived in ECMWF’s MARS database.

The Editor thanks Philip Jonathan and an anonymous reviewer for their assistance in evaluating this paper.

References

Aarnes, O. J., Ø. Breivik, and M. Reistad (2012), Wave extremes in the Northeast Atlantic, *J. Clim.*, *25*, 1529–1543, doi:10.1175/JCLI-D-11-00132.1.

Bidlot, J., D. Holmes, P. Wittmann, R. Lalbeharry, and H. Chen (2002), Intercomparison of the performance of operational ocean wave forecasting systems with buoy data, *Weather Forecasting*, *17*(2), 287–310, doi:10.1175/1520-0434(2002)017<0287:IOTPOO>2.0.CO;2.

Breivik, Ø., Y. Gusdal, B. R. Furevik, O. J. Aarnes, and M. Reistad (2009), Nearshore wave forecasting and hindcasting by dynamical and statistical downscaling, *J. Mar. Syst.*, *78*(2), S235–S243, doi:10.1016/j.jmarsys.2009.01.025.

Breivik, Ø., O. J. Aarnes, J.-R. Bidlot, A. Carrasco, and Ø. Saetra (2013), Wave extremes in the North East Atlantic from ensemble forecasts, *J. Clim.*, *26*, 7525–7540, doi:10.1175/JCLI-D-12-00738.1.

Buizza, R., J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart (2007), The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System), *Q. J. R. Meteorol Soc*, *133*(624), 681–695, doi:10.1002/qj.75.

Caires, S., and A. Sterl (2005), 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data, *J. Clim.*, *18*(7), 1032–1048, doi:10.1175/JCLI-3312.1.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, 210 pp., Springer Verlag, London, U. K.

Compo, G. P., et al. (2011), The twentieth century reanalysis project, *Q. J. R. Meteorol. Soc.*, *137*(654), 1–28, doi:10.1002/qj.776.

Debernard, J., and L. P. Røed (2008), Future wind, wave and storm surge climate in the Northern Seas: A revisit, *Tellus A*, *60*, 427–438, doi:10.1111/j.1600-0870.2008.00312.x.

Dee, D., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597, doi:10.1002/qj.828.

Elsner, J. B., J. P. Kossin, and T. H. Jagger (2008), The increasing intensity of the strongest tropical cyclones, *Nature*, *455*(7209), 92–95, doi:10.1038/nature07234.

Hagedorn, R., R. Buizza, T. Hamill, M. Leutbecher, and T. Palmer (2012), Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts, *Q. J. R. Meteorol. Soc.*, *138*, 1814–1827, doi:10.1002/qj.1895.

Hemer, M. A., Y. Fan, N. Mori, A. Semedo, and X. L. Wang (2013), Projected changes in wave climate from a multi-model ensemble, *Nat. Clim. Change*, *3*(5), 471–476, doi:10.1038/nclimate1791.

Hersbach, H., C. Peubey, A. Simmons, P. Poli, D. Dee, and P. Berrisford (2013), ERA-20CM: A twentieth century atmospheric model ensemble, *ERA Report Series 16*, European Centre for Medium-Range Weather Forecasts, U. K., http://www.ecmwf.int/publications.

Izaguirre, C., F. J. Mendez, M. Menendez, A. Luceño, and I. J. Losada (2010), Extreme wave climate variability in southern Europe using satellite data, *J. Geophys. Res.*, *115*, C04009, doi:10.1029/2009JC005802.

Izaguirre, C., F. J. Méndez, M. Menéndez, and I. J. Losada (2011), Global extreme wave height variability based on satellite data, *Geophys. Res. Lett.*, *38*, L10607, doi:10.1029/2011GL047302.

Janssen, P. (2004), *The Interaction of Ocean Waves and Wind*, 300 pp., Cambridge Univ. Press, Cambridge, U. K.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–472, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kharin, V., and F. Zwiers (2005), Estimating extremes in transient climate change simulations, *J. Clim.*, *18*(8), 1156–1173, doi:10.1175/JCLI3320.1.

Lopatoukhin, L., V. Rozhkov, V. Ryabinin, V. Swail, A. Boukhanovsky, and A. Degtyarev (2000), Estimation of extreme wind wave heights, *JCOMM Tech. Rep. No 9*, World Meteorological Organization.

Magnusson, L., and E. Källén (2013), Factors influencing skill improvements in the ECMWF forecasting system, *Mon. Weather Rev.*, *141*, 3142–3153, doi:10.1175/MWR-D-12-00318.1.

Magnusson, L., J.-R. Bidlot, S. Lang, A. Thorpe, N. Wedi, and M. Yamaguchi (2014), Evaluation of medium-range forecasts for hurricane Sandy, *Mon. Weather Rev.*, *142*, 1962–1981, doi:10.1175/MWR-D-13-00228.1.

Molteni, F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, L. Magnusson, K. Mogensen, T. Palmer, and F. Vitart (2011), The new ECMWF seasonal forecast system (System 4), *ECMWF Tech. Memorandum 656*, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2, U. K.

Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, *85*(3), 369–432, doi:10.2151/jmsj.85.369.

Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda (2006), Tropical cyclone climatology in a global-warming climate as simulated in a 20 km-mesh global atmospheric model: Frequency and wind intensity analyses, *J. Meteorol. Soc. Jpn.*, *84*(2), 259–276, doi:10.2151/jmsj.84.259.

Reistad, M., Ø. Breivik, H. Haakenstad, O. J. Aarnes, B. R. Furevik, and J.-R. Bidlot (2011), A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea, *J. Geophys. Res.*, *116*, C05019, doi:10.1029/2010JC006402.

Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, *24*, 3624–2648, doi:10.1175/JCLI-D-11-00015.1.

Saha, S. et al. (2013), The NCEP climate forecast system version 2, *J. Clim.*, *27*, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Stockdale, T., D. Anderson, M. Balmaseda, F. Doblas-Reyes, L. Ferranti, K. Mogensen, T. Palmer, F. Molteni, and F. Vitart (2011), ECMWF seasonal forecast system 3 and its prediction of sea surface temperature, *Clim. Dyn.*, *37*, 455–471, doi:10.1007/s00382-010-0947-3.

Uppala, S., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, doi:10.1256/qj.04.176.

Vinoth, J., and I. Young (2011), Global estimates of extreme wind speed and wave height, *J. Clim.*, *24*(6), 1647–1665, doi:10.1175/2010JCLI3680.1.

Vitart, F., R. B. M. A. Balmaseda, G. Balsamo, J.-R. Bidlot, A. Bonet, M. Fuentes, A. Hofstadler, F. Molteni, and T. N. Palmer (2008), The new VarEPS-monthly forecasting system: A first step towards seamless prediction, *Q. J. R. Meteorol. Soc.*, *134*(636), 1789–1799, doi:10.1002/qj.322.

Wang, X., and V. Swail (2001), Changes of extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes, *J. Clim.*, *14*(10), 2204–2221, doi:10.1175/1520-0442(2001)014<2204:COEWHI>2.0.CO;2.

Wang, X., and V. Swail (2002), Trends of Atlantic wave extremes as simulated in a 40-yr wave hindcast using kinematically reanalyzed wind fields, *J. Clim.*, *15*, 1020–1035, doi:10.1175/1520-0442(2002)015<1020:TOAWEA>2.0.CO;2.

Wang, X., and V. Swail (2006), Climate change signal and uncertainty in projections of ocean wave heights, *Clim. Dyn.*, *26*(2), 109–126, doi:10.1007/s00382-005-0080-x.

Wang, X., Y. Feng, and V. Swail (2012), North Atlantic wave height trends as reconstructed from the 20th century reanalysis, *Geophys. Res. Lett.*, *39*, L18705, doi:10.1029/2012GL053381.

Weisse, R., and H. Günther (2007), Wave climate and long-term changes for the Southern North Sea obtained from a high-resolution hindcast 1958–2002, *Ocean Dyn.*, *57*, 161–172, doi:10.1007/s10236-006-0094-x.

World Meteorological Organization (1998), *Guide to Wave Analysis and Forecasting,* 2nd ed., World Meteorol. Org., Geneva, Switzerland.

Young, I., J. Vinoth, S. Zieger, and A. Babanin (2012), Investigation of trends in extreme value wave height and wind speed, *J. Geophys. Res.*, *117*, C00J06, doi:10.1029/2011JC007753.

