Ocean Sci. Discuss., 8, C888–C898, 2012
www.ocean-sci-discuss.net/8/C888/2012/

Ocean Science Discussions

Interactive comment on "Wave climatology in the Arkona Basin, the Baltic Sea" *by* T. Soomere et al.

Anonymous Referee #2

Received and published: 15 January 2012

The work is a comprehensive climatological analysis of buoy and numerically simulated wave data in the Southern part of the Baltic Sea. It outlines some aspect of the Baltic wave climate which is not often visible in other areas, like the effect of the sea breeze on wave climate. There is, though, some problem in the methodology used in the statistical analysis of the time series which should be properly addressed. The comparisons between measurements and simulations are also not completely satisfactory. An effort in improving clarity and language would be really appreciated by the reader.

General comments

The time series considered are serially correlated, this affects both the methods used in the estimate of the probability distributions and the extreme event analysis. One consequence is that, in general, probability and frequency are different. Statistical analysis and conclusions should take into account the peculiarity of wind-wave data.

The joint frequency functions (of period and wave height) should be considered on a homogeneous data-set, for example on all the observations at the synoptic times in all the period considered. Another possibility is to consider all the observation taken at 30 min. since 48 observations/day are available.

The probability distribution should be evaluated on independent data, the extreme event analysis is also similarly affected: a solution could be the use of the Peak Over Threshold method (MATHIESEN, M., HAWKES, P., MARTIN, M.J., THOMPSON, E., GODA, Y., MANSARD, E., PELTIER, E., and VAN VLEDDER, G., 1994. Recommended practice for extreme wave analysis. J. Hydraulic Research., IAHR, Vol. 32, pp.803-814.).

Given the importance of the wind in the numerical simulations, some information about the adjustment of the geostrophic wind should be given. A discussion (possibly quantitative) about the reliability of the adjusted wind field at the surface seems also necessary. What information does the use of the adjusted geostrophic wind add to the analysis, by the way?

The analysis of the diurnal variation is affected by 2 problems:
1. wave data are non-homogeneous in time (there is a different number of measurement per day)
2. all meteorological conditions are mixed, from sea breeze to storm.
Both can be easily fixed.

Language and style should be improved.

Specific comments

"During the time interval of 1991-2010 (7305 calendar days altogether), covered in this paper, there were 6198 days with at least one sensible recording per day whereas 1107 days had no data. (15% , 3 years) The gaps in data are distributed unevenly over the calendar year (Fig. 3). There are very few gaps in the summer season, from June to August when usually measurements from only one year are missing. "
** It is really not easy to understand the meaning

The lowest percentage of days covered by measurements is in January and October and, somewhat surprisingly, in April and May when typically 15 yr out of 20 are represented." ???
*** It is really not easy to understand the meaning. The figure does not help much either, unfortunately. Maybe a small table with the annual number of missing data would be more clear. A table with the distribution of missing data by month might be more effective.

"This difference in recordings does not become evident in the diurnal course of the wave height (Fig. 4b). On average, the largest wave heights occur at the Darss Sill between 10:00 and 15:00 whereas a clearly visible minimum exists for late evening and night (21:00-01:00)."
***is it statistically significant?
*** does it change considering different range gates, for example all data $H_{m0} < 2m$ and $H_{m_0} > 2m$?

"The average amplitude of this diurnal variation over all seasons is 2.6 cm, that is, almost 3.5 %. In June and July the amplitude is much larger, about 11% and 14.5%, respectively. The diurnal variation is reversed, with higher waves present in the evening time in November and December"
***How is calculated the average amplitude of the diurnal variation??

OSD
8, C888–C898, 2012

"The almost perfect match of the long-term wave height at the Darss Sill indicates, however, that the resulting driving wind speeds are adequate for the SW Baltic Sea."
*** 8 cm in the average wave height does not represent a 'perfect match'.
*** A comparison should not be based on the use of the global average of H_{m_0} only. It would be advisable to apply some other statistical operator (mse, bias...)

"The quite substantial diurnal variation in wave heights suggests that the changing number of wave measurements per day and during different seasons or longer time intervals may affect the estimates of the climatological properties of wave fields. "
*** which is exactly why the data-set must be homogeneous.

"The long-term average wave height, however, is almost the same when calculated from all measurements in a straightforward way (0.7603 m) and from daily mean wave heights (0.7608 m)."
*** but the mse would change, see also fig 5 and 6

"For all measurements the waves are most frequent in the range of 0.25-0.375m whereas wave situations with the daily mean heights of 0.25-0.625m occur with an almost equal frequency. The difference in the shapes of these two distributions becomes evident in terms of the parameters of the relevant Weibull distribution: k = 1.6550, b = 0.8505 for the entire data set and k = 1.8112, b = 0.8552 for the daily mean wave heights. This dissimilarity would lead to obvious differences in estimates for the probabilities of occurrence of severe seas. Interestingly, the parameter b is almost the same for the AW (1958-2002) and RS (1970-2007) simulations (b = 0.8416 and b = 0.8320, respectively) whereas the parameter k is much smaller (k =1.4906 and k =1.3774, respectively)."
*** data are serially correlated: POT would be of help (see General Comments)

"A comparison of the distributions of the occurrence of different wave heights estimated from measurements with those extracted from numerical reconstructions for the time interval of 1991$--$2002, which is covered by all the data sets (Fig. 6), shows certain deviations of the modelled wave statistics from the measured data. The modelled data sets adequately reproduce the frequency of moderately high wave fields (0.5-2 m). 25 Simulations by RS overestimate low wave heights (0.125-0.375 m) whereas simulations by AWunderestimate the frequency of such wave conditions. The largest discrepancy between the model results occurs for the range of wave heights of 0.25-0.375 m. Still, both the models adequately reproduce the overall proportion of wave fields with HS < 0.5 m: the measured data contain 49% of such fields while the RS hindcast has 52% of waves with HS < 0.5 m. The AW hindcast result for such wave fields (44 %) is close to the value in terms of daily mean wave heights (48 %). Both models tend to overestimate the proportion of relatively rough wave fields with Hs> 1.5m. This feature may partly reflect several longer gaps in the measured data pool (see Sect. 2)."

** data must be homogeneous

"caption Fig. 9. 'Variations on a weekly scale of (a) measured and modelled (AW) wave heights at the Darss Sill for 1991-2002 and (b) of all measurements in 1991-2010 and model results for 1957-2002.'

**monthly scale?? daily scale?

"Comparison of the annual course of daily mean wave heights over a longer time scale (Fig. 9) demonstrates that most of the variations on weekly scales do not persist over longer time intervals. For example, extensive variations at the end of January – beginning of February in 1991-2002 are completely lost in data covering four decades and the amplitude of the variations during the calm period is much smaller. As in many weeks this course of wave height is already in antiphase, the variations on weekly

scales apparently are smoothed out over a longer time series of observations. Still, an interesting feature is the relatively low period of wave activity in December, which becomes also evident in visually observed data sets from the NE Baltic Sea (Soomere et al., 2011). Also, Fig. 9 demonstrates that while the monthly mean wave height evidently well represents the wave activity during windy (November-February) and calm (May-July and partially August) months, there exist systematic variations in the wave activity within the transitional months (March-April and September-October)."
*** not clear: in fig 9 there are more than 22 points in a month, so they probably are 1 point every day or so. Does every point represent the average of the daily H_{m_0} of the same calendar day (i.e. avg of 20 values of 1st Jan, avg of 20 values of 2th Jan...)?
*** it is not surprising that averaging smooths out the variance, but the analysis should consider also the variance (or mse)
***a climatological analysis of wind speed/direction (wind rose) would help to understand the variability

"There has been a very weak increase (about 9 mm/decade) in the modelled wave heights. The increase was somewhat larger since about 1970 until 2002 (14 mm/decade) but is still much lower than reported in (Soomere, 2011). "
*** mse or variance analysis seems necessary. statistical significativity might help.

"The RS simulations using adjusted geostrophic winds show a considerable increase (about 50 mm/decade) in the wave heights for 1970-1990 and a comparable decrease since then. Although both the simulations reproduce the annual mean wave heights with an error of < 15 %, neither of them adequately represents the interannual variations in the measured wave heights. In general, the AW simulation appears to reproduce the large interannual variations somewhat more exactly than the RS simula tion. The latter seems to be in antiphase with respect to measured data for a few years

(e.g., 1996 and 1997). "
** given the data distribution, comparison shoud be made at least on percentiles

"The temporal behaviour of the basic annual characteristics of modelled and measured wave data (annual mean wave heights and the higher quantiles of the wave heights – thresholds for 10 % and 1 % of the highest waves; also called 90 % and 99 % percentiles, respectively)"
** simply percentiles

"is mostly very similar except for the year 1995 when their course was somewhat different (Fig. 11). The measured wave heights seem to undergo a quasiperiodic change, with relatively high waves around 1995 and 2005, and with a lower wave activity at the turn of the century. In essence, such a manner of variations matches the variation derived from visual obser vations at the Lithuanian coast (Zaitseva-P ÌĹarnaste et al., 2011)."
** Conclusion seems hardly adequately supported: at very least a scatterplot should be considered. Apparently the variance seems to be so high that any trend could fit the available data.
** Among all possible parameters, the annual average height hardly represent the better choice

" The formal trend for measured wave heights is decreasing (18 mm/decade). The behaviour of the 90 % and 99 % percentiles is, not surprisingly, qualitatively very similar to that of the annual average wave height (Fig. 11b). These quantities only show a larger interannual variability and relatively higher peaks with respect to their average values. While the threshold for the highest 10 % of wave heights shows a very weak increase (14 mm/decade), the threshold for the highest 1 % of wave heights reveals a comparably weak decrease with a comparable rate (−18 mm/decade). Interestingly, the 90 % and 99 % percentiles show a different trend for both modelled

and measured waves."
*** I understand that WAM overpredicts the waves by approx. half a meter: here a measure of the bias of the H_{m_0} at the peak of the storm episodes would help.

" While these characteristics saw almost no changes in the modelled data over 1958–2002, the 90 %-tile decreased at a rate of 57 mm/decade and the 99 %-tile even at a rate of 72 mm/decade. "
*** conclusions should really consider also mse/variance

" Notice that the latter rate would have been considerably larger without the evident peak in 2010 (Fig. 11b)."
**which is exactly why the conclusion about any trend is not supported by the analysis

" In general, the modelled thresholds have almost the same long-term average while the relevant values for particular years differ considerably. The modelled thresholds for 10 % of the highest waves very slightly exceed the values estimated from the measurements. Both models tend to generally overestimate the 99 %-tile (by up to 40 % for some years) although in some other years (e.g. 1993 or 1995) the modelled and measured estimates of this threshold almost coincide.However, the most typical (0.64 m)"
**why the median should be most typical?? there can be no observation at all

" Analysis of the modelled and measured wave data from the Darss Sill measurement site in SW Baltic Sea reveals that there have been no substantial changes in the long- term wave climatology in this area since the 1950s."
**buoy measurements were not available in all the period

" Somewhat surprisingly, the wave measurement site is almost perfectly sheltered from storm waves excited in the Baltic Proper and the role of remote swells is

OSD
8, C888–C898, 2012

negligible in this area. This feature together with the small size of the sea area allows for considerable simplification of the repro- duction of wave properties for engineering applications: in many cases the wave field properties are determined by the local wind and fetch length."
**the analysis is based on average heights only. No directional information is given, so conclusions tends to be qualitative. Maybe a wind rose might better support the claim. How many episodes are determined only by local wind? And how many are not? and if they are not, what would be their wave height at the peak?

" The described combination of wave properties explains why both numerical simu- lations and in situ wave recordings obviously distort the shape of the joint distribution of wave heights and periods by systematically overestimating periods if they are < 3 *s."*
***if observations are wrong and simulations are wrong too, it wouldn't seem there is much space for a scientific discussion.

" This shortcoming can be removed by using an extended range of frequencies, up to about 1 Hz, to properly resolve shor t-wave situations in fur ther simulations. The limi- tations of the waverider are apparently deeper and some other means should be used for wave measurements in order to properly resolve the shor t-wave domain."
***Are waves with periods less than 2s really relevant as 'pure' wind waves?

" This feature by no means contradicts the phenomenon of increasing wind speed."
***it would, if true

" More importantly, it demonstrates that wave proper ties in semi-sheltered sea areas such as the Baltic Sea depend on the wind properties in a complicated manner and frequently are more strongly affected by (changes to) the wind direction (Soomere

and R ÌĹa ÌĹamet, 2011b)."
***but in the present discussion no information is given about directionality of wind and waves

Technical corrections

"increase in SW winds" —> increase in South-Westerly winds

*** it would improve clarity to use percentiles in all the discussion.

****"changes to"* —> changes in

"The number of measurements per day varies over time. It is 9–10 during the first years of measurements, decreases to about 8 by the end of 1991, increases to about 48 from April 1999, fluctuates very strongly between almost zero and 50 in 2001-2002 and up to March 2003, after which it stabilises on the level of 48 measurements per day. As each measurement cycle covers 1600 s, each day can only host maximally 54 independent measurements. For this reason, measurements with overlapping time interval are left out from the further analysis in which maximally 50 measurements a day are accounted for (Fig. 2). The total amount of recorded wave conditions is 608 in 1991-2010, that is, about 26 measurements per day."
*** I don't really think that programming 54 measurements/day on a Waverider would be a sensible (or possible) practice, as some time is needed in order to elaborate and transmit the data. In my experience this kind of buoys are programmed to operate at fixed times. In the eighties 8 measurements at synoptic hours were a common choice. Then there was the option to switch (automatically) to a 30 min recording when waves exceeded a given threshold. Generally, buoys have been operating at 30 min recording (48 measurement/day) since calculations are made onboard. May be this can be the reason of the inhomogeneity in the data-set. The bottom line here is that the measurement are always taken at 'fixed' times, while from the text it seems that

OSD
8, C888–C898, 2012

what is fixed is the number of measurements/day.

OSD

8, C888–C898, 2012

Interactive Comment