Final Draft
of the original manuscript:

Behrens, A.:
Development of an ensemble prediction system for ocean surface waves in a coastal area
In: Ocean Dynamics (2015) Springer

DOI: 10.1007/s10236-015-0825-y

Development of an ensemble prediction system for ocean surface waves in a coastal area

The final publication is available at link.springer.com

Arno Behrens
Helmholtz-Zentrum Geesthacht, Institute of Coastal Research,
Max-Planck-Str. 1, 21502 Geesthacht, Germany
Email : arno.behrens@hzg.de, phone : +49 4152-87-1556, fax : +49 4152 87-41556

Abstract: An ensemble prediction system for ocean surface waves has been developed and applied on a local scale to the German Bight and the western Baltic Sea. U_{10}-wind fields generated by the COSMO-DE-EPS upstream forecast chain of the German Met Service (DWD: Deutscher Wetterdienst) have been used as the driving force for the third-generation spectral wave model WAM. The atmospheric chain includes four different global models that provide boundary values for four regional COSMO-EU realisations. Each of those drive five COSMO-DE members respectively with different sets of physical parameterisations, so that finally 20 members are available to run 20 corresponding wave ensemble members of the coastal wave model CWAM (**C**oastal **WA**ve **M**odel) for the German Bight and the western Baltic Sea. It is the first time that in an ensemble prediction system for ocean waves an atmospheric model of such a fine spatial resolution of 2.8 km has been combined with a wave model running on a model grid with a mesh size of 900 m only. Test runs with the wave ensemble prediction system have been executed for two entire months (April 2013 and June 2014) and for an eight days storm case (Xaver) in December 2013 in order to check whether such a system could be a reasonable step to improve the future operational wave forecasts of the DWD. The results computed by the different wave model members agree fairly well with available buoy data. The differences between the results for the integrated wave parameters of the individual members are small only, but more pronounced in extreme storm situations. Finally the statistical analysis of the comparisons with measurements show without exception slightly improved values for the ensemble mean of the wave ensemble members compared with the usual deterministic routine control run.

Keywords: Ensemble prediction · ocean surface waves · COSMO-DE · spectral wave model WAM · wave forecasting · German Bight

1. Introduction

Beside the numerical ensemble prediction which is already a well-accepted method to improve the performance of atmospheric models, a number of promising activities in the field of ensemble prediction for ocean surface waves have been done as well. Global ocean wave forecast ensemble systems are available at the European Center for Medium-Range Weather Forecasts (ECMWF) since 1998 (Hoffschildt et al. 1999, Janssen 1999), at the U.S. National Weather Service (NWS) since 2005 (Chen 2006), at the National Centers for Environmental

Prediction (NCEP) (Cao et al. 2007, 2009), at the US Navy Fleet Numerical Meteorology and Oceanographic Center (FNMOC) and since 2011 in a collaboration between NCEP and FNMOC as a multicenter ensemble at NCEP (Alves et al. 2013). A limited-area wave ensemble prediction system for the Nordic seas and the North Sea is running at the Norwegian Meteorological Institute (NMI) since 2008 (Carrasco and Saetra 2008). Benefits of these wave ensemble systems have been shown in several studies. Hoffschildt et al (1999) reported on the benefits of the probabilistic forecasts of the ECMWF wave ensemble system to shiprouting. Saetra and Bidlot (2004) demonstrated the potential benefit of the ECMWF ensemble prediction system using buoy, platform and satellite altimeter data for verification of waves and marine surface winds. Saetra et al. (2004) discussed the effects of observation errors on the statistics for ensemble spread and reliability. The benefit of the ensemble predictions for safe offshore operations have been discussed by Carrasco et al. (2011) and the advantages of the limited-area wave ensemble prediction system of the NMI for increasing predictability in regional areas are presented in Carrasco and Saetra (2008). Grabemann and Weisse (2008) used an ensemble of wind field data sets to investigate the impact of future climate change scenarios on extreme waves in the North Sea. Bertotti et al. (2011) discussed a deterministic and ensemble-based prediction of five storms in the Adriatic Sea that lead to a flooding of Venice and Breivik et al. (2013) presented a method for estimating return values from ensembles of forecasts at advanced lead times.

In contrast to atmospheric forecast models which represents highly nonlinear dynamic systems that could generate chaotic forecasts due to small perturbations in the initial condition, the models used for ocean wave forecasting are weakly nonlinear systems only that are highly dissipative, losing the information of the initial wave field very fast. Since the source term for nonlinear wave interaction is highly nonlinear, spectral perturbations can be quickly flushed out of the system by transferring those to higher frequencies to be dissipated there. On the other hand, perturbations in the low frequency part of the wave spectrum can persist longer, especially for swell propagating on larger scales. Since in this investigation the focus is targeted on the development of a wave ensemble forecast system for a limited area on a small local scale, only the atmospheric data are perturbed in the ensemble computations for the ocean surface waves in the coastal areas of the German Bight and the western Baltic Sea, concentrating on the variability of the wind forcing rather than on perturbations of the initial state in the wave models which have very small effects on the results only (Farina 2002). This is a different approach compared to other investigations that use for example super-ensemble techniques to merge different forecast or hindcast results, obtained by different wave models for the same area, into a single multi-model prediction system on a local scale. Lenartz et al. (2010) did that for the Adriatic Sea and found that this method performs better than any single forecasting system. Another promising technique, the so-called operational consensus forecast scheme has been discussed in Durrant et al (2008). It was applied to an ensemble of 10 deterministic wave model outputs, gathered from different operational weather centers. The operational consensus forecast scheme uses bias correction and combines model data to provide improved forecasts where recent data are available. It was found that with such a technique it is possible to beat the best individual model with a composite forecast of the worst four.

The idea for this investigation, to set up an ensemble prediction system for ocean surface waves on a local scale for the North Sea and the western Baltic Sea, came up since the DWD run their atmospheric ensemble forecast system operationally for their local COSMO-DE model which covers perfectly the model grid of the wave model CWAM. That offers the opportunity to use the different U_{10}-wind fields generated by the 20 ensemble members of the COSMO-DE-EPS as the driving force for 20 corresponding wave model ensemble members. Although the main focus in the COSMO-DE-EPS is the maximization of the variability in precipitation, sufficient differences in the wind fields are expected that are able to generate an ensemble of wave model results whose ensemble mean could improve the probabilistic wave forecast skills compared with the deterministic routine runs. Other sources of uncertainty affecting the wave forecasts in such a shallow coastal area are the water level variations and the currents which are not included in the wave model setup.

This paper is organized as follows. Section 2 discusses the final set up for the wave ensemble members on the model grid of the coastal wave model CWAM for the German Bight and the western Baltic Sea that are driven by the wind fields obtained by the COSMO-DE-EPS members of the DWD, followed by a short description of the wave model used in section 3. The results of the wave ensemble calculations are discussed in section 4 completed by a summary and conclusions in section 5.

2. Setup of the ensemble members for the coastal wave model CWAM

The atmospheric ensemble system COSMO-DE-EPS is based on the numerical weather forecast model COSMO-DE (Baldauf et al. 2011) and comprises currently 20 ensemble members which are computed with the same horizontal grid size (2.8 km) as the operational routine configuration. The COSMO-DE-EPS provides one-hourly predictions twice a day at 0 and 12 UTC for a forecast period of 21 hours (27 hours since March 2014) respectively on the COSMO-DE model grid that includes Germany and its surroundings. The different ensemble members are generated by variations of the characteristics of the forecast system. That includes the boundary values, initial conditions and model physics (Paralta et al. 2012). Currently methods are applied that are similar to the multi-model ensemble approach. For the variations of the boundary and initial values, global forecasts of four different global atmospheric models are used. Those are the global model GME of the DWD, the integrated forecast system (IFS) of the ECMWF, the global forecast system (GFS) of the NCEP and the global spectral model (GSM) of the Japan Meteorological Agency (JMA). The four different global models are embedded in a chain of ensembles shown in figure 1. The members of the COSMO-DE-EPS are nested in four different predictions of the COSMO-EU model with its mesh size of 7 km. These COSMO-EU predictions, each of those driven by one of the four global models, form the boundary value ensemble BCEPS (boundary condition ensemble prediction system) that provide four different sets of boundary values to drive the members of the COSMO-DE-EPS.

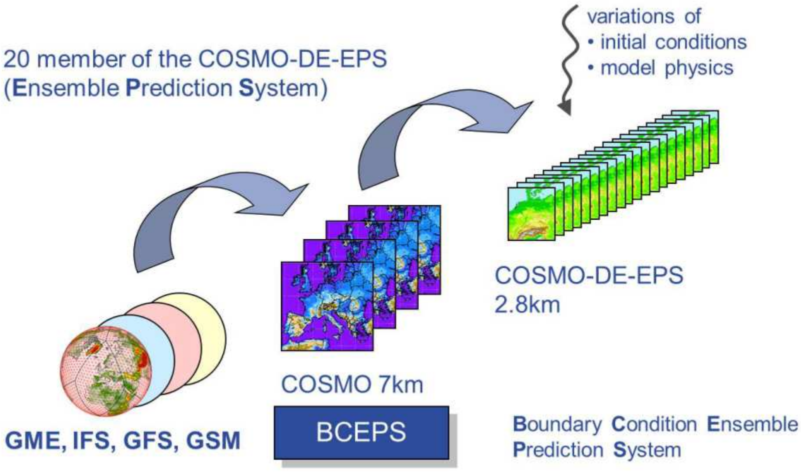


Fig. 1 : Atmospheric forecast chain that drives the COSMO-DE-EPS

For the variation of the model physics different model configurations of the COSMO-DE have been designed and tested (Gebhard et al. 2011). In the different configurations the value of a parameter that is in connection with a certain parameterization has been changed respectively. The changed value will be kept constant during the whole simulations and therefore every member of the current configuration of the COSMO-DE-EPS is clearly characterized with respect to its model physics. The corresponding five parameters with the default and perturbed values are given in table 1. A detailed description of these parameters can be found in the COSMO manual (Schättler et al. 2009).

Table 1 : The 5 different COSMO-DE model configurations in the COSMO-DE-EPS

no.	parameter	relating to parameterization	default	perturbed
1	entr_sc	entrainment rate for shallow convection	0.00003 m^{-1}	0.002 m^{-1}
2	q-crit	critical value for normalized oversaturation	1.6	4.0
3	rlam_heat	scaling factor of the laminar boundary layer for heat	1.0	0.1
4	rlam_heat	scaling factor of the laminar boundary layer for heat	1.0	10.0
5	tur_len	asymptotic mixing length of turbulence	150 m	500.0 m

The five different model configurations of table 1 combined with the four different members in the boundary data ensemble BCEPS generate a total of 20 COSMO-DE-EPS members. Initial and boundary values for one of the individual COSMO-DE-EPS members are provided always by the same BCEPS members. Figure 2 includes an overview for the generation of the 20 COSMO-DE-EPS members. The five columns denote the five different COSMO-DE configurations and the four rows the four BCEPS members identified by its corresponding global model.

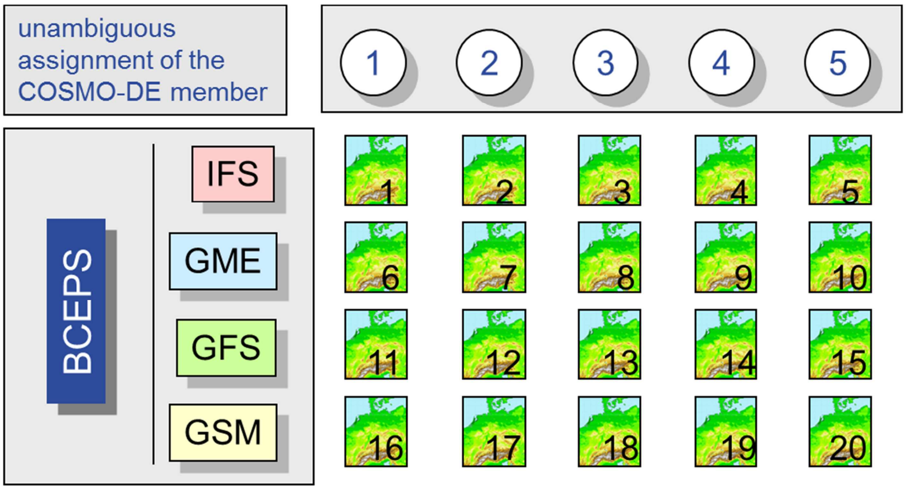


Fig. 2 : Construction of the 20 COSMO-DE-EPS members.

The U_{10}-wind fields generated by the COSMO-DE-EPS upstream forecast chain have been used as the driving force for the spectral wave model WAM. According to the individual members of the atmospheric chain, 20 ensemble members for the ocean surface waves in the German Bight and western Baltic Sea have been computed on the CWAM model grid. The corresponding flow chart of the whole ensemble wave forecast system is shown in figure 3.

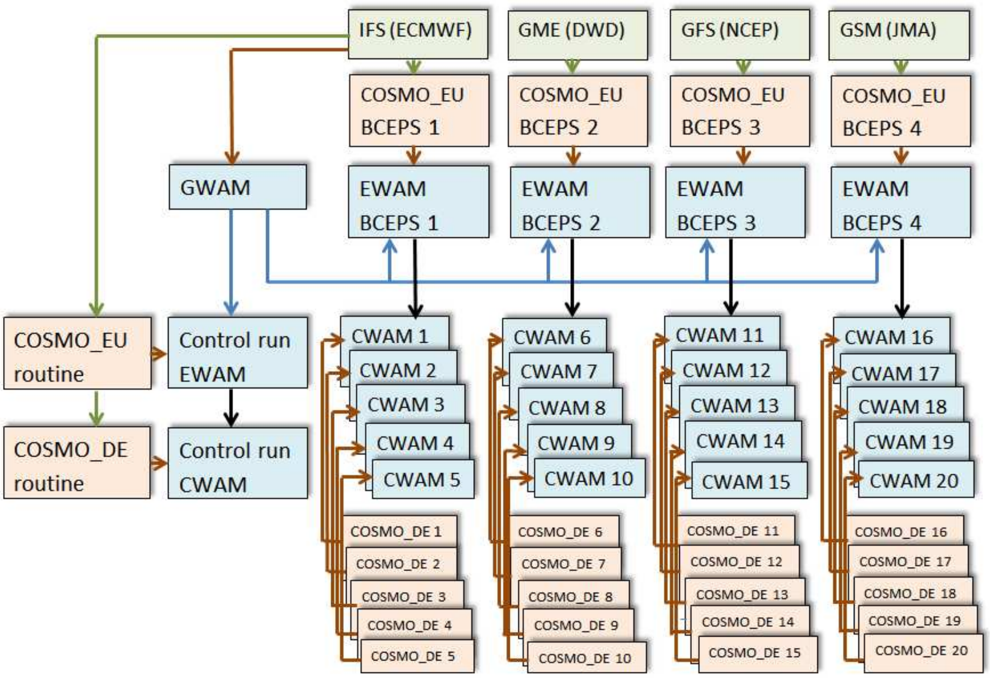


Fig. 3 : Wave ensemble prediction system

Four regional wave models (EWAM) provide boundary values for five CWAM members respectively. Furthermore a control run for all scales has been carried out with the wind fields obtained from the deterministic atmospheric models GME for the global scale, COSMO-EU for the regional scale and COSMO-DE for the local scale as the driving force in order to generate a reference for the results of the different members and the ensemble mean.

A summary of all 27 wave model runs with the source of the driving wind fields are given in table 2 which relates the wave models to the atmospheric models and gives the number of runs that are required to complete the whole system successfully.

Table 2 : different driving forces for 27 wave model runs		
atmospheric models	number of members	wave models
GME	1	GWAM
COSMO-EU	1	EWAM (control run)
COSMO-DE	1	CWAM (control run)
COSMO-BCEPS	4	EWAM-EU
COSMO-LMK	20	CWAM

Furthermore four additional runs for the Xaver time period in December 2013 have been performed to check whether one of the regional COSMO-BCEPS realisations, that are driven by the four different global atmospheric models, generate EWAM-EU boundary values that induces better wave predictions with CWAM when the latter uses the same COSMO-DE wind fields without perturbations. The results of these four control runs have been compared among themselves and each of those with the ensemble mean of the five perturbed member that are driven by the corresponding regional BCEPS-EU member which is identified by its global model. That corresponds in each case to one of the four rows in figure 2. So it becomes apparent whether one of the four global models is the reason for the final CWAM results that are closest to the measurements.
In completion of the setup an ensemble mean of those CWAM runs has been generated for the Xaver storm period that uses different boundary values, but only the wind fields provided by the COSMO-DE-EPS members that are perturbed with the same physical modification. That corresponds in each case to one of the 5 columns in figure 2. The aim was to compare the effect of the 5 selected perturbations in the COSMO-DE-EPS members on the wave results among themselves.

3. Wave model WAM

The wave model that has been used since 1999 by the DWD as the appropriate tool for their numerical wave predictions is the well-established advanced third generation spectral wave model WAM that runs successfully at many institutions worldwide.

WAM is based on the spectral description of the wave conditions in frequency and directional space at each of the active grid points of a chosen model area. The energy balance equation, complemented with a suitable description of the relevant physical processes is used to follow the evolution of each wave spectral component. A full description is given by the WAMDI group (1988), Komen et al. (1994), Günther et al. (1992) and Janssen (2008).

The WAM Cycle 4.5.4 that is used for the wave ensemble computations is an update of the former WAM Cycle 4. The basic physics and numerics are kept in that new release. The source function integration scheme made by Hersbach and Janssen (1999), and the model updates (Bidlot, et al. 2007) are incorporated. Other main improvements introduced in WAM Cycle 4.5.4 are technical improvements, which take into account the new possibilities of Fortran 95 and the MPI (**M**essage **P**assing **I**nterface) for parallelization purposes. On request from the user community a number of additional options are added in the model. A big advantage of the new state-of-the-art version WAM Cycle 4.5.4 is its high-grade modular composition which allows an easy replacement of individual parts of the code. The wave model WAM Cycle 4.5.4 is public and a download of the web-based source code library for users is possible from the GitHub server under http://mywave.github.io/WAM .

The local wave model CWAM runs on a model grid situated between 53°13'45" N to 056°26'45" N and 6°10'25" W to 14°54'35" E, with a spatial resolution of about 900 m, also 30 seconds in latitude, respectively 50 seconds in longitude. The required bathymetry for that model grid has been developed out of different local data sets at the BSH in Hamburg. The bathymetry for the North Sea model EWAM-EU which provides the boundary values for CWAM is generated out of the GEBCO (General Bathymetric Chart of the Oceans, available at http://www.gebco.net) data set, improved along the coastlines manually by the German navy. CWAM calculates the two-dimensional energy density spectrum at each of the 124011 active model grid points in the frequency and directional space. The solution of the energy balance equation is provided for 36 directional bands at 10° each, starting at 5° and measured clockwise with respect to true north, and 30 frequencies logarithmically spaced from 0.042 Hz to 0.66 Hz at intervals of $\Delta f/f = 0.1$. Time dependent depth and current fields are not used in this setup. Results of the CWAM ensemble runs are stored every three hours until the end of the forecast period is reached.

4. Results of the ensemble predictions

The 27 wave model runs listed in table 2 that are required to generate an entire set of 20 local CWAM realisations have been computed for two whole months (April 1013 and June 2014) and for an one week storm period in December 2013 to find out whether such an ensemble system for ocean waves can improve the forecast skills for such a local application. All the hindcasts started already two days before the considered time periods with a cold start to ensure a reasonable initial wave field at the beginning of the validation period. A representative example of the model set up with the global model, the regional model and the local model is given in figure 4 for the first member of the BCEPS and the COSMO-DE-EPS showing the distribution of the significant wave height in the three model areas on the 18 April 2013, 12 UTC.

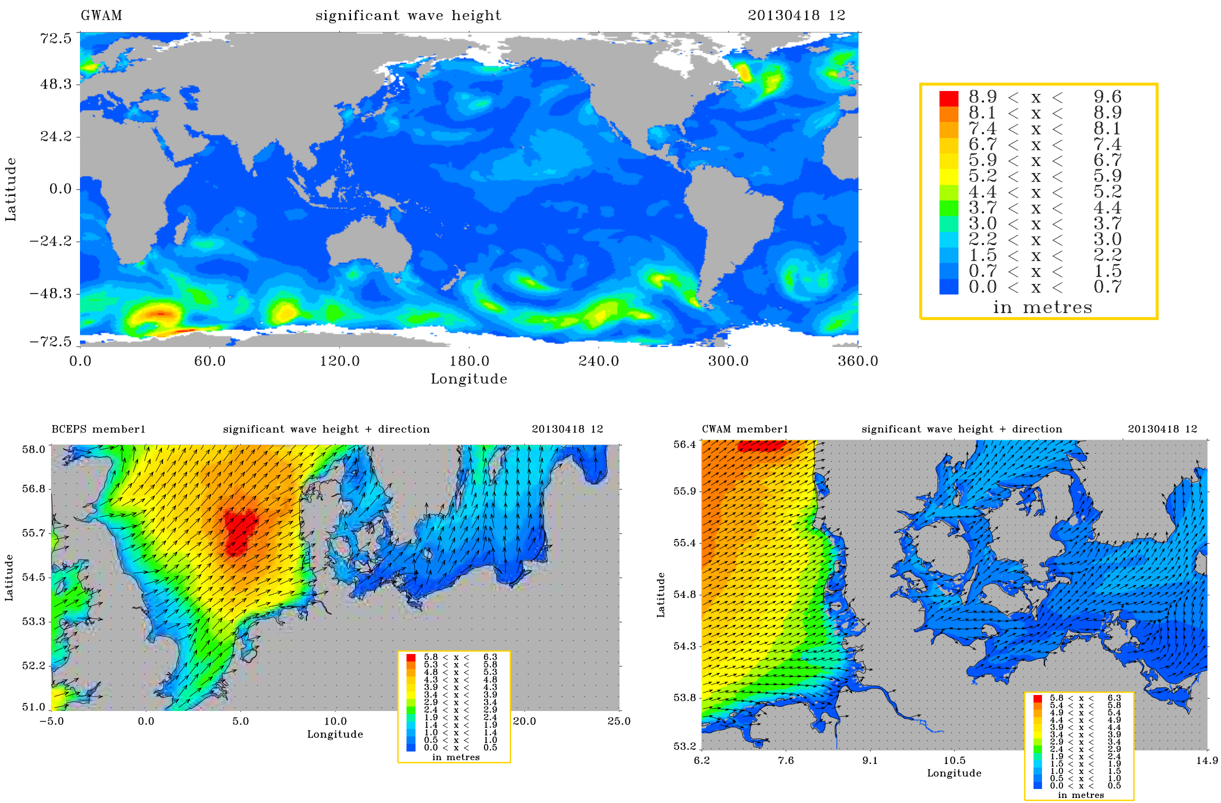


Fig. 4 : Example for the wave ensemble chain (a. global, b. regional, c. local wave model CWAM)

The results of the individual ensemble members and the control run predictions have been compared with measurements recorded at several buoy locations in the German Bight and in the western Baltic Sea. The geographical locations of the available buoys in the CWAM model grid are shown in figure 5.

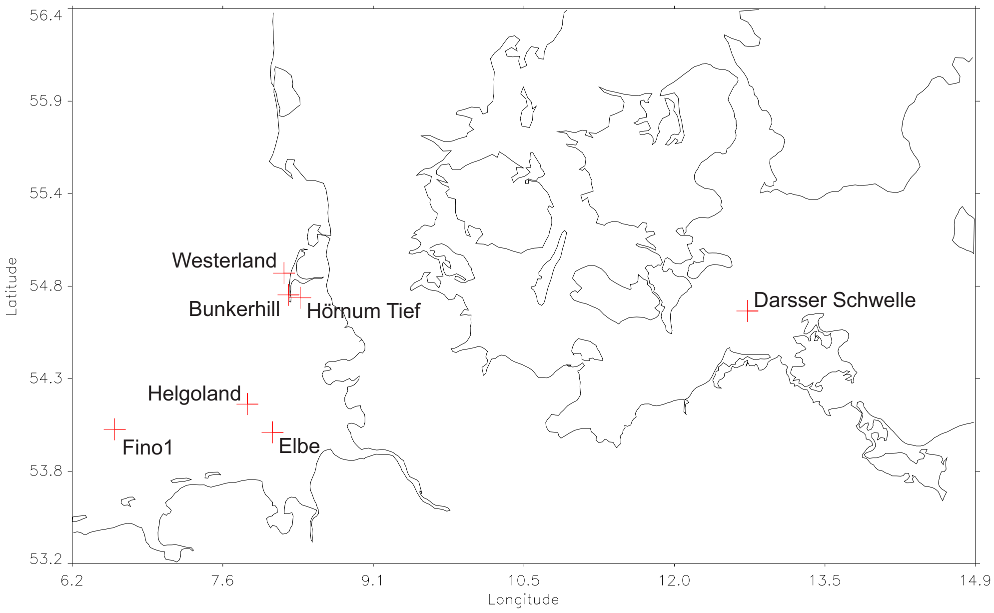
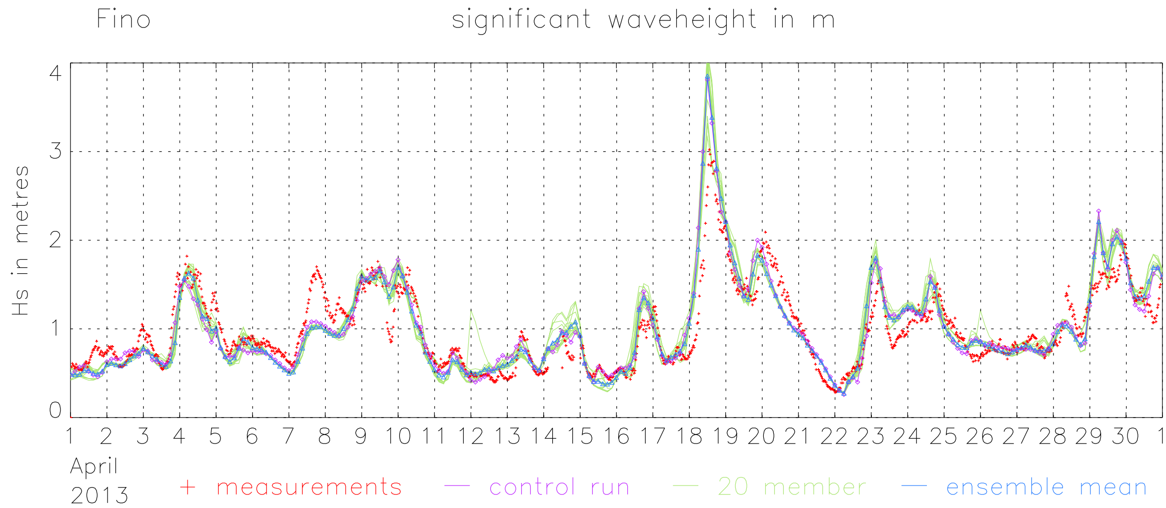
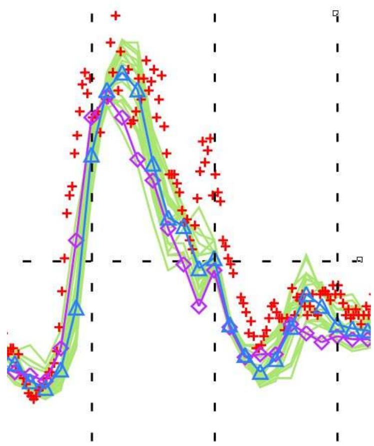


Fig. 5 : Locations in the CWAM model area where measurements are available

The buoys Helgoland, Westerland, Elbe and Fino are operated by the BSH (Bundesamt für Seeschifffahrt und Hydrographie, Hamburg) and the buoys Hörnum, Bunkerhill and Darsser Schwelle by the HZG (Helmholtz-Zentrum Geesthacht). All of those are datawell directional waverider buoys. Spatial and temporal scales of measurements and model results are made comparable for the statistical analysis by averaging the half-hourly observations in time windows of one hour centered on the three-hourly model output time steps.
At the buoy locations time series plots of measured and computed integrated wave parameters have been generated. In addition to the time series of the 20 individual ensemble members, an ensemble mean has been calculated and is included in the validation plots. A representative example for the measured and computed significant wave heights H_s for April 2013 at the location Fino is given in figure 6.





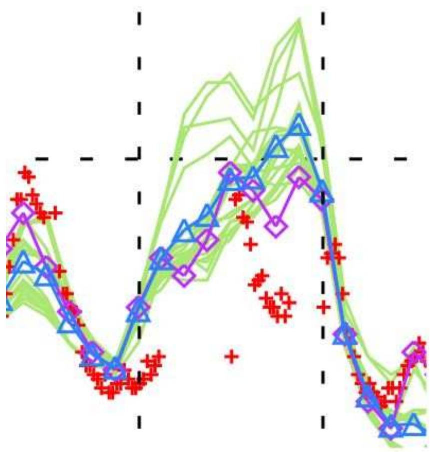


Fig. 6 : time series of significant wave heights at the location Station Fino (a. above) and zoom in for the time periods 3-6 April (b. below, left) and 13-16 April (c. below, right)

The overall agreement between the measured and the computed wave heights is good and the differences between the results of the different ensemble members are small. However to get a better illustration of the conditions at Fino it makes sense to zoom in the comparisons for two small time periods. The two pictures below the complete time series in figure 6 enable a

closer look into the behaviour of the curves of the different members, the control run and the ensemble mean compared with the measurements. In the left picture the results of nearly all of the ensemble members for the time period 3-6 April 2013 show higher values than the control run and therefore the ensemble mean is definitely in better agreement with the measurements than the control run. But this is not always the case as shown in the right picture for the time period 13-16 April 2013. A couple of the individual members underestimate the first and overestimate the second peak and therefore the results of the control run match the measurements better in this case. But in general the agreement between ensemble mean and measurements is slightly better than that one between control run and measurements.
Comparisons between measured data and the results of the wave ensemble members have not been done for significant wave heights only, but for other integrated wave parameter as well. The peak period T_p which is more volatile than the significant wave height H_s due to its dependency from the resolution of the frequency space is considered in a time series plot in figure 7 that includes measured and computed peak periods at the buoy location Elbe. An interesting feature occurs here in the afternoon of the 22 April with measured peak periods around 16 s. A zoom in, shown in the left picture of figure 7 illustrates the characteristics of the behaviour of the different members at that time. It seemed to be that three of those predict the swell peak periods quite well, whereas the control run misses that event. But maybe that this is an artefact only since the total peak period gives always the value in the frequency bin with the maximum of the wave energy. The values for wind sea and swell could be very close together and the decision whether it is wind sea or swell can be a fortuity. Anyway, the ensemble mean as a result of this is definitely nearer to the measurements as the peak periods of the control run.

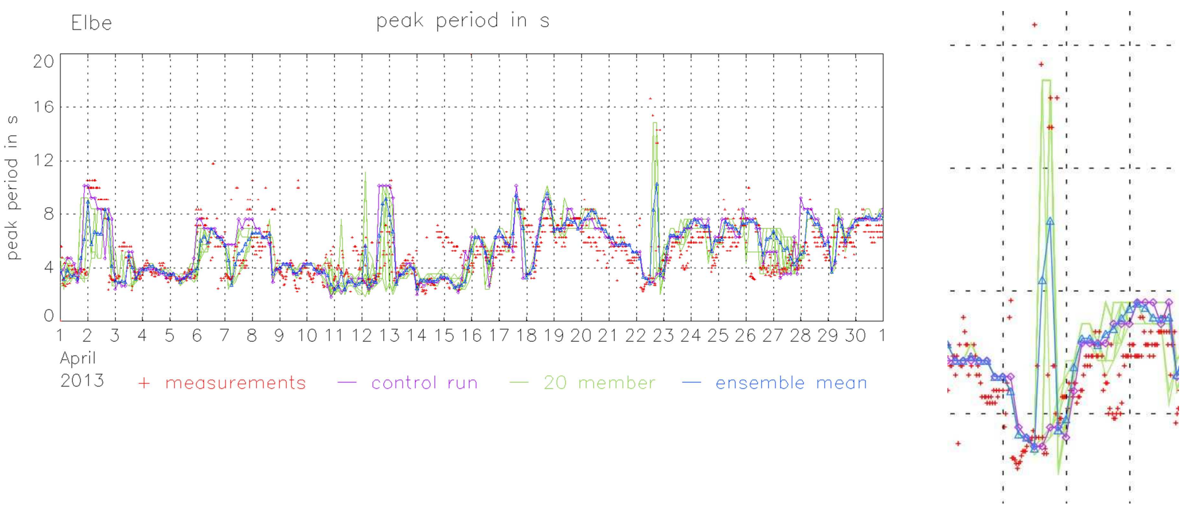


Fig. 7 : time series of the peak period at location Elbe (a), zoom 22-23 April (b)

For the Tm_2 periods and the total wave directions is the agreement between the measured data and all the computed values excellent. Representative examples plotted for the buoy location Fino are included in figure 8. The differences between the results of the 20 different wave ensemble members are very small for these two wave parameters. That has been proved by the determination of maximum, minimum and spread of those. The spread of the ensemble mean corresponds to the standard deviation of the individual members from the ensemble mean (the mathematical definition is given in A9 in the appendix). The computed Tm_2 periods have been compared with the corresponding measured T_z periods (mean zero up-crossing periods).

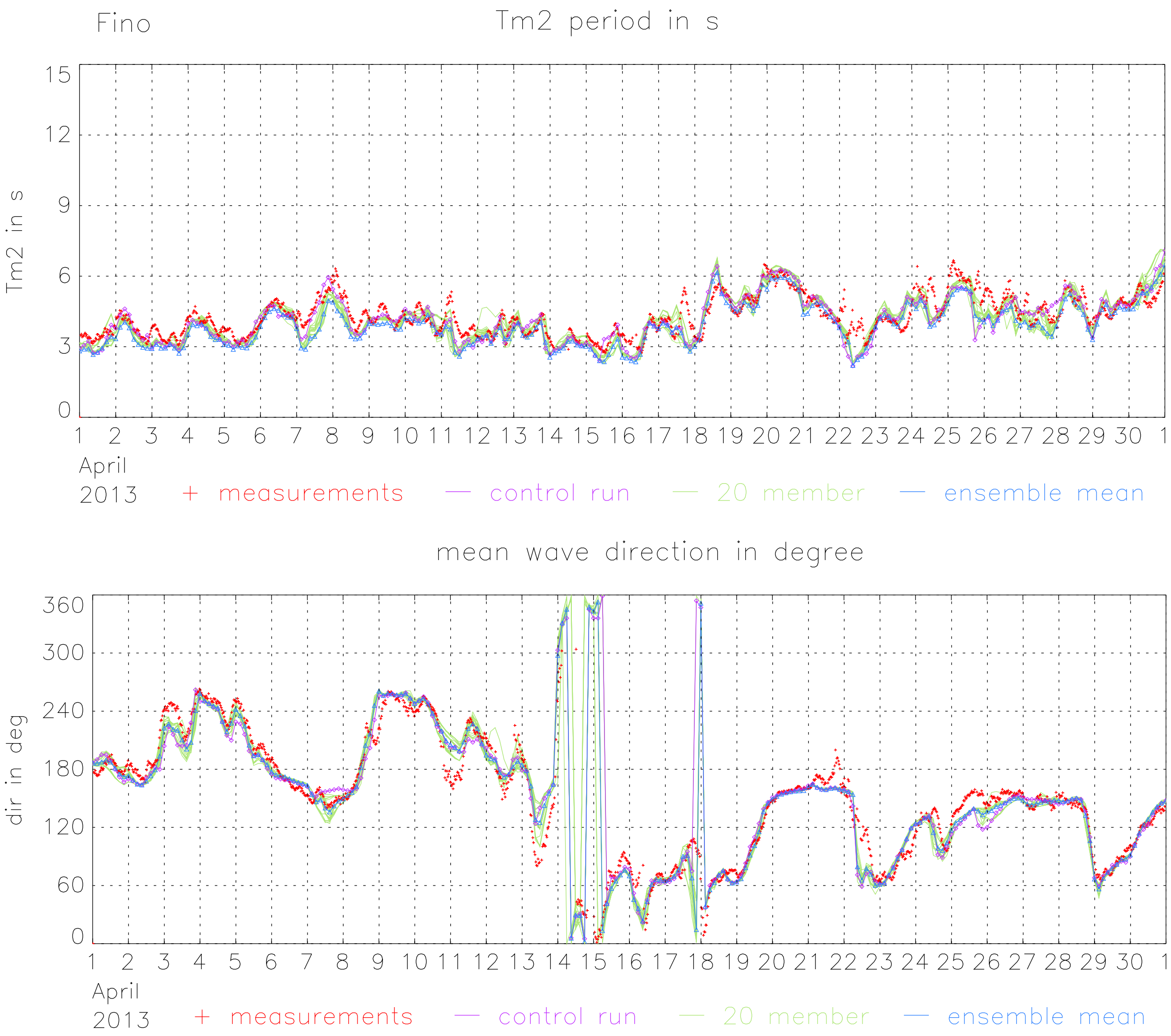
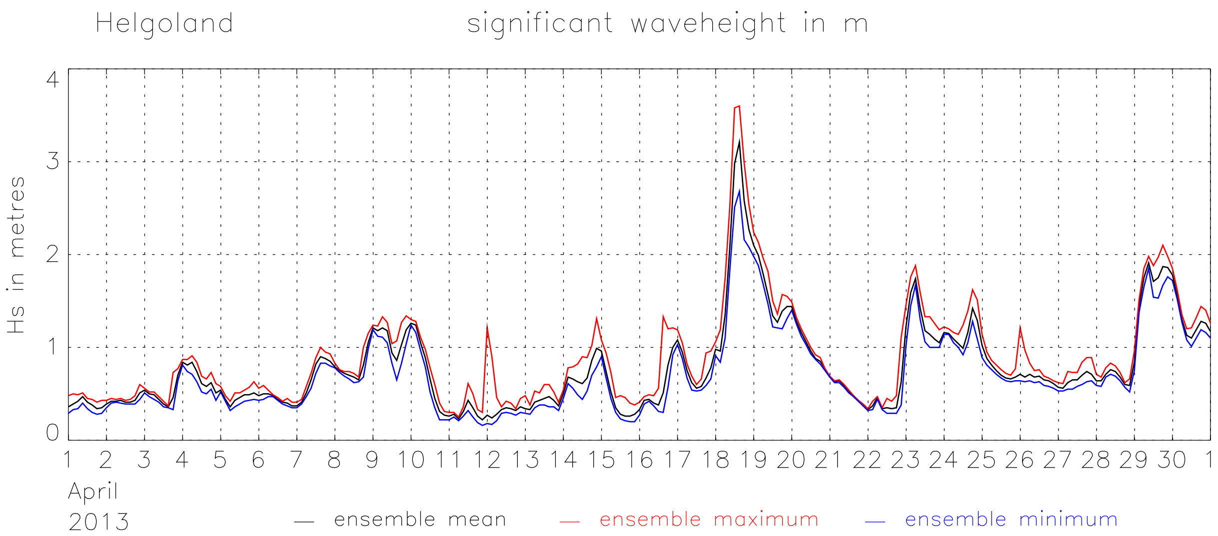


Fig. 8 : time series of the Tm_2-periods (a. above) and the total wave direction (b. below) at Fino

The variability in the results of the different wave ensemble members are more pronounced in the significant wave heights and therefore plots have been generated that show the ensemble mean of the 20 individual members together with the maximum and the minimum of those. Furthermore the spread of the ensemble mean has been calculated and plotted as ensemble mean ± spread as well. Figure 9 shows representative examples for these time series at the buoy location Fino in the German Bight.



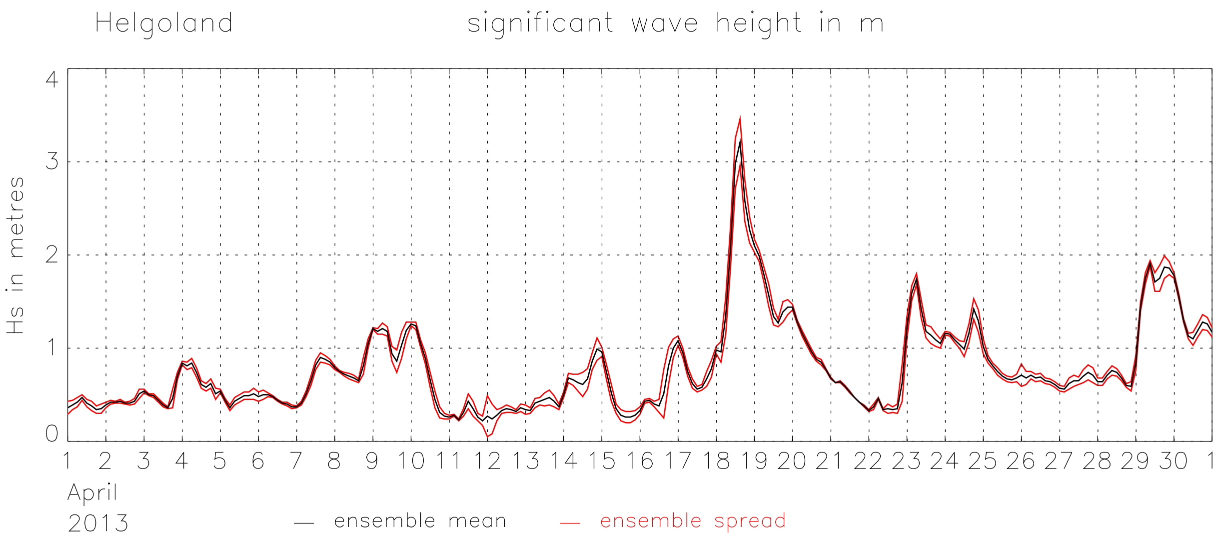


Fig. 9 : H_s-maximum and minimum of the 20 members at buoy location Helgoland (a. above) and the ensemble spread (b. below)

These plots support the former statement that the differences in H_s between the results of the different wave ensemble members are small. The maximum of the members is generally more pronounced than the minimum and differences can be actually more than 50 cm. The ensemble spread is definitely very small during April 2013 and don't exceed 20 cm. A disadvantage in this case is definitely the short forecast period of the atmospheric members of 21 hours only.

Finally a statistical analysis has been done for the significant wave heights. The statistical parameters calculated for the comparisons between the measurements and the results of the control run are included in table 3 and those obtained for the comparisons between the measured data and the ensemble mean for April 2013 are given in table 4. The mathematical definitions of all the statistical parameters are given in the appendix (equations A1 – A7). The uncertainty of the estimation of the skill measures for the reduction of variance and the scatter index is discussed later for the Xaver time period.

Table 3 : H_s statistics for the comparisons between the measurements and the control run							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	236	1.00	0.02	0.24	0.92	0.80	24
Elb	240	0.69	0.04	0.18	0.84	0.83	25
Wes	139	0.53	-0.03	0.17	0.98	0.69	31
Hel	178	0.76	0.08	0.20	0.78	0.86	24
Bun	214	0.58	0.06	0.21	0.71	0.83	35
Hoe	240	0.22	0.03	0.10	0.93	0.59	45
Dar	240	0.85	-0.25	0.34	1.70	0.28	27

Table 4 : H_s statistics for the comparisons between the measurements and the ensemble mean							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	236	1.00	0.01	0.23	0.92	0.81	23
Elb	240	0.69	0.05	0.17	0.83	0.85	23
Wes	139	0.53	-0.02	0.15	0.96	0.75	27
Hel	178	0.76	0.08	0.20	0.78	0.86	24
Bun	214	0.58	0.08	0.21	0.69	0.84	33
Hoe	240	0.22	0.04	0.10	0.91	0.59	42
Dar	240	0.84	-0.25	0.33	1.71	0.31	26

Looking especially at the two parameters skill (reduction of variance) and scatter index which give a measure for the quality of the numerical results, it becomes obvious that the values for the ensemble mean comparisons are slightly better than those obtained with the control run. It is expected in this context that by constructing the rms-based scores, the ensemble mean will on average fare better than any single forecast. The comparisons that have been done in this investigation do not constitute a validation of the probabilistic ability of the ensemble as was done in Saetra et al. (2004) or Cao et al. (2009).

Beyond the investigations for the entire month April 2013, a shorter one week period has been computed with the wave ensemble members to check the behaviour of the individual members in extreme storm situations. In the beginning of December 2013, the storm Xaver afflicted the German coasts with maximum significant wave heights of more than 12 m. The distribution of H_s at the storm peak on 05 December 2013, 18 UTC, is shown in figure 10, together with the driving wind field at that time. The north-westerly winds drive the waves directly to the coasts with wind speeds up to 30 m/s and generates waves higher than 10 m in large parts of the German Bight and actually more than 12 m in some small parts directly in front of the coastline due to shoaling effects.

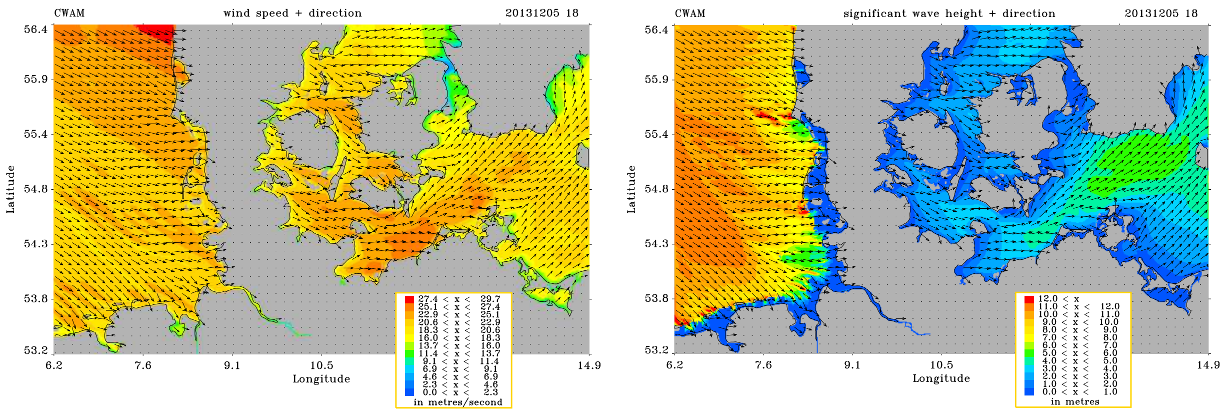


Fig. 10 : distribution of wind speed and wind direction (a. left) and significant wave height H_s and total wave direction (b. right) at the peak of storm Xaver computed by the control run (18 hours forecast)

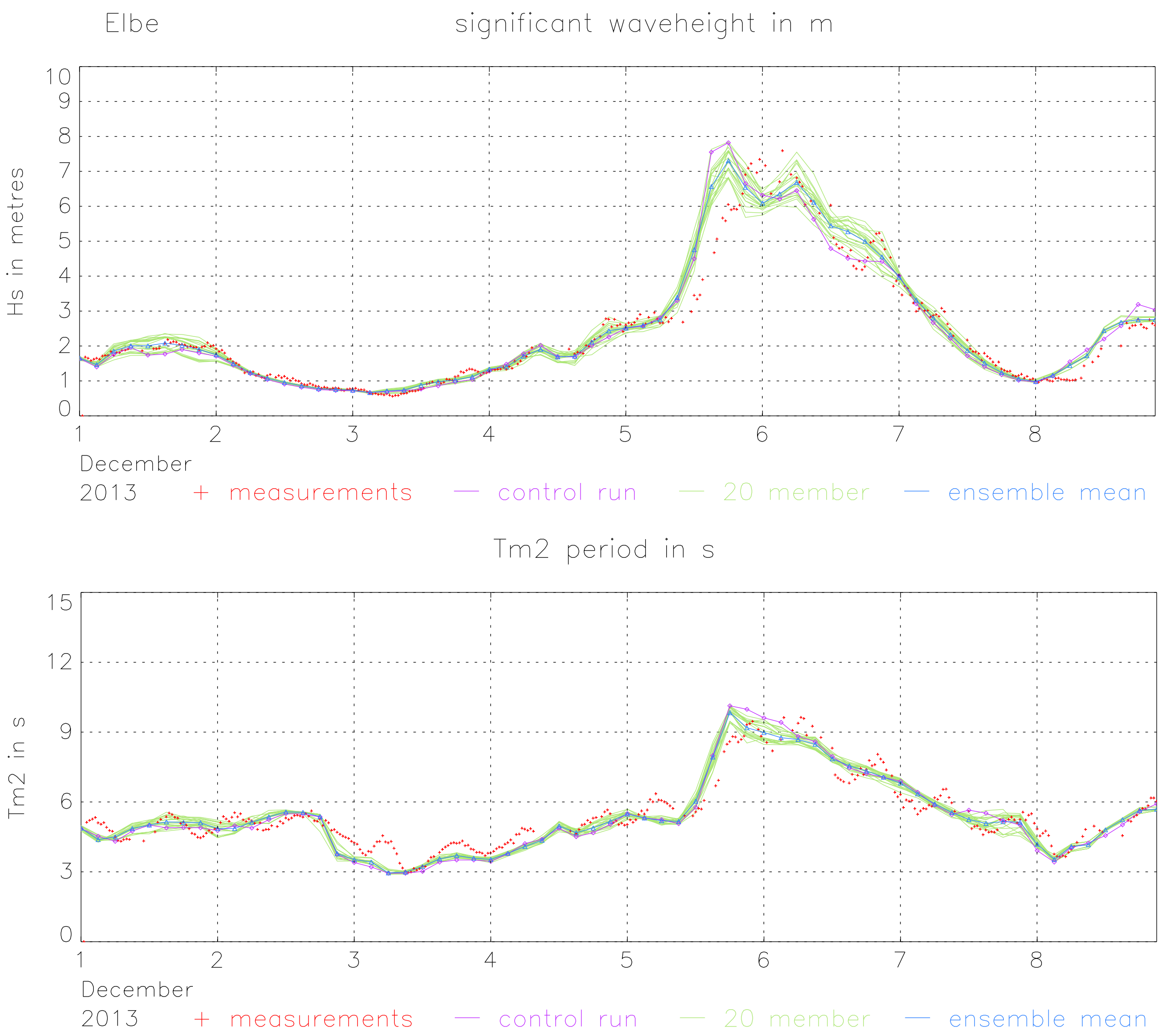


Fig. 11 : time series of H_s (a. above) and Tm_2/T_z-periods (b. below) at buoy Elbe during storm Xaver

Time series plots for the eight days including the Xaver storm peak show some interesting features. Figure 11 includes the comparisons between measured and all computed values for the significant wave heights and the Tm_2/T_z-periods at the buoy Elbe. Whereas the agreement of all the curves is very good during the first five days, there is a time shift at the peak in all the model predictions for this location; the peak in all the model results is predicted too early in this case due to the corresponding wind fields. The agreement at this buoy between the periods is fairly good although the measurements show a clear tidal modulation which can't be found in the model results because the water level deviation are not taken into account in the wave model.

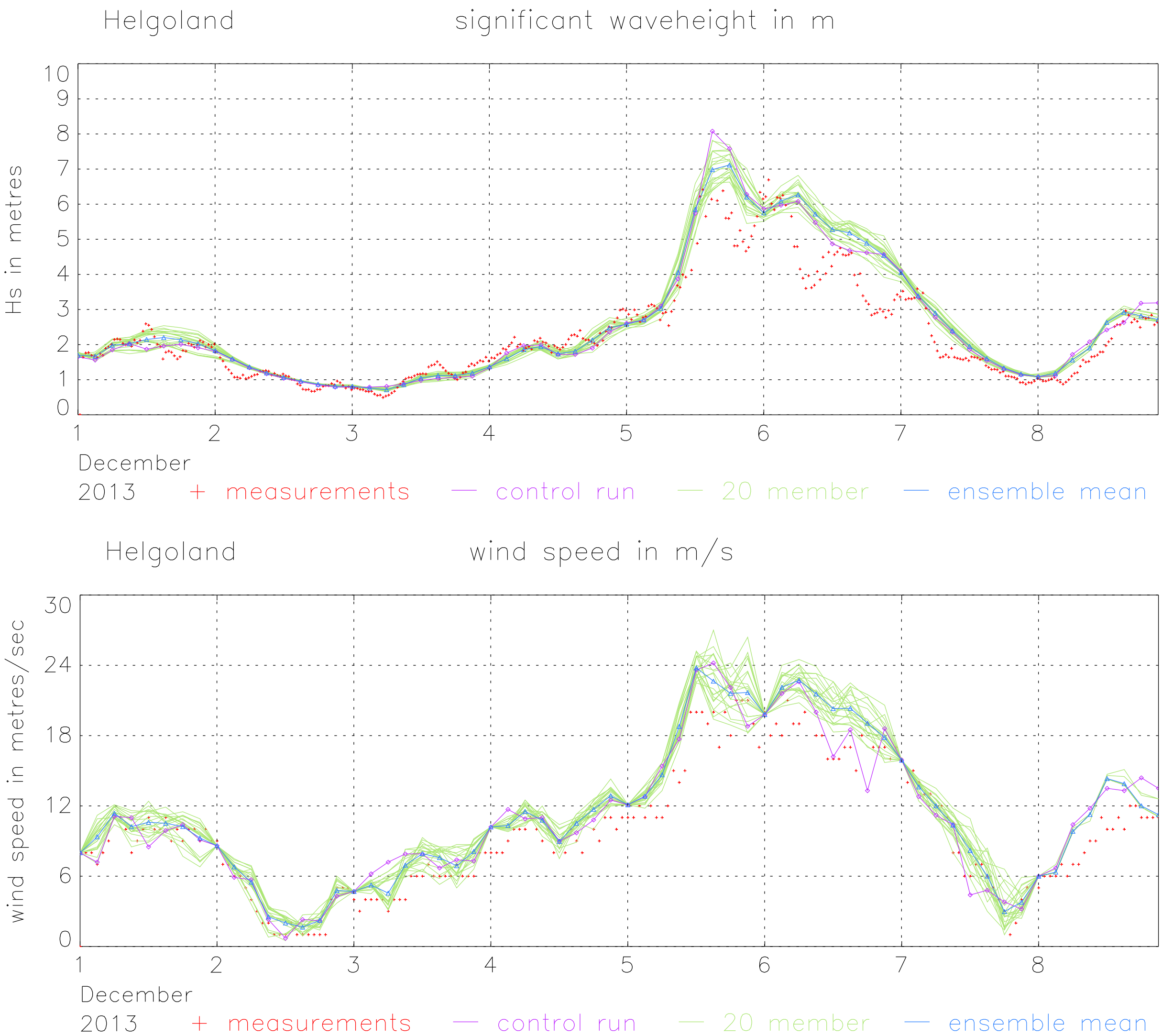


Fig. 12: time series of H_s (a. above) and wind speeds (b. below) at buoy Helgoland during storm Xaver

This tidal signal observed already in the comparisons of the periods at the buoy Elbe are included clearly in the time series for the measured significant wave height at the buoy Helgoland as well. Figure 12 shows the corresponding time series of measured and computed significant wave heights and wind speeds. The computed values matches well the measured wave heights during the first five days, the peak is overestimated by the ensemble mean by 60 cm. This is due to the significant overestimation of the wind speeds at the peak detected in all COSMO-DE realisations with values up to 3 m/s. The wind speed measurements at Helgoland are recorded in integer steps of one meter. During the rest of the time period, the agreement for the significant wave height is satisfactory at flood tide and the model values overestimates the measured data always at ebb tide. These results are definitely a distinct hint to run such a coastal wave model coupled to a hydrodynamic model in order to take care of the depth variations in the wave model grid due to the tidal water level deviations.

As expected, the differences in the driving wind fields of the individual COSMO-DE-EPS members are more pronounced in extreme situations. That can be shown also in the time series of the ensemble mean of the significant wave height, together with the maximum and the minimum of the individual members as given in figure 13 for the buoy Helgoland.

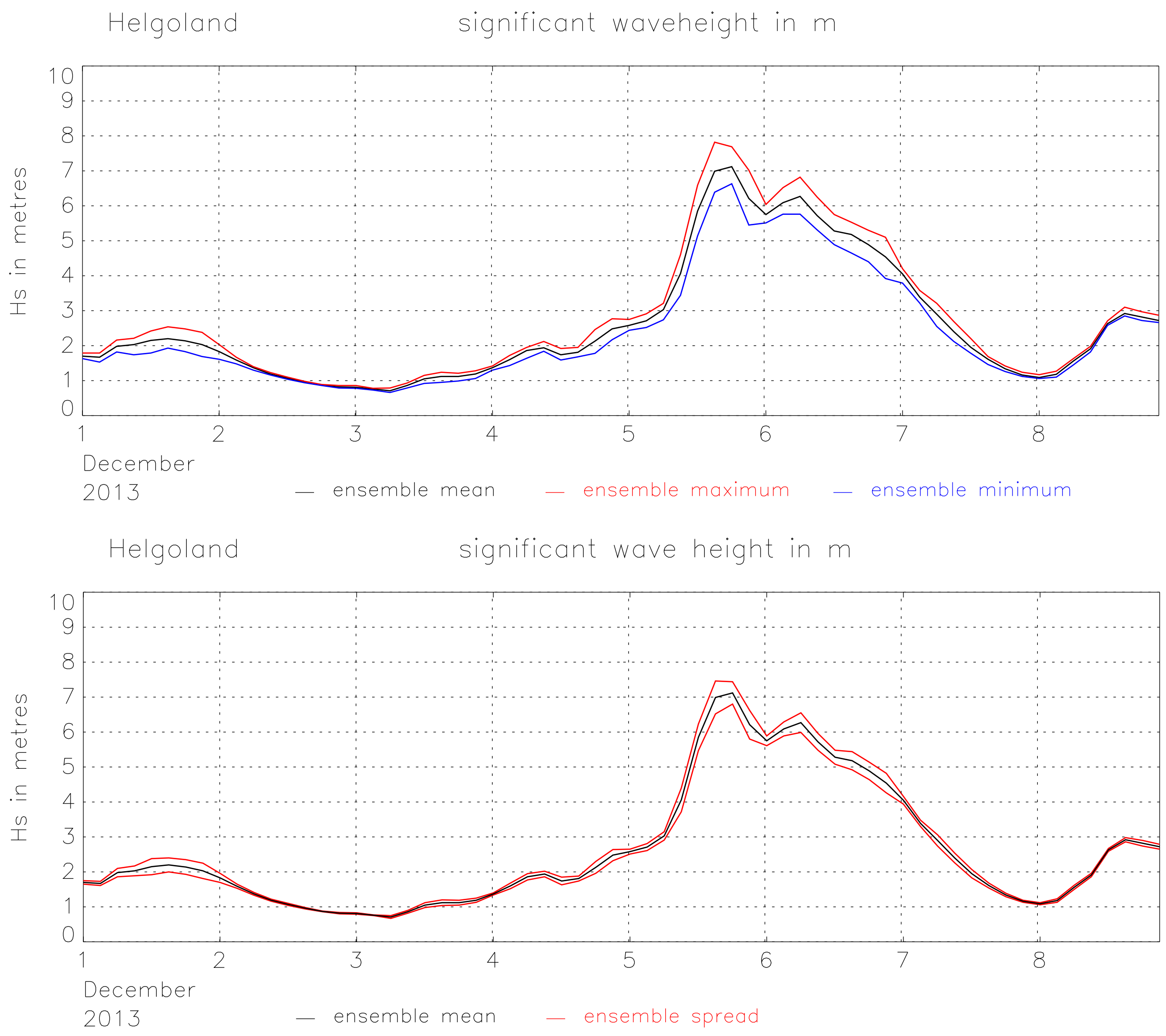


Fig. 13: H_s-maximum and minimum of the 20 members at buoy Helgoland (a. above) and the ensemble spread (b. below) during storm Xaver.

During the storm Xaver on 5-6 December, the differences are clearly enhanced up to about 1 m at the peak. The spread of the ensemble mean in the lower picture of figure 13 shows maximum values up to 50 cm during the Xaver storm period.

The biases in the statistics that have been done for the short eight days' time period prove that the model results overestimate the significant wave heights in extreme storm situations. This is valid both for the results of the control run in comparison with the measurements (table 5) and the results of the ensemble mean with measured data (table 6). The two buoys Hörnum Tief and Darsser Schwelle unfortunately did not measure during that time period. The buoy at Bunkerhill failed to work when Xaver started in the afternoon of the 5th of December. Checking the statistical parameters skill and scatter index which are a measure for the quality of the model runs, it becomes obvious that the ensemble mean performs better than the control run. To get a feeling for the uncertainty of the estimation of skill measures, the bootstrapping method (Efron and Tibshirani 1993) has been applied to calculate the standard error for the statistics of skill and scatter index for the comparisons of measurements and ensemble mean results for the Xaver storm period. A number of bootstrap samples are generated from the original Xaver data set followed by 1000 bootstrap replicates that are obtained by calculating the corresponding statistical value on each bootstrap sample. Finally the standard deviation of that value is the estimation of the standard error of it. The results of that calculations are included in table 6.

Table 5 : Xaver - one week statistics for the comparisons between measurements and control run

buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	58	2.97	0.15	0.82	0.79	0.88	27
Elb	61	2.48	0.02	0.52	0.96	0.92	21
Wes	63	2.34	0.49	0.77	0.64	0.85	25
Hel	63	2.34	0.25	0.67	0.77	0.86	27
Bun	36	1.48	0.25	0.33	0.74	0.86	14

Table 6 : Xaver - one week statistics for the comparisons between measurements and ensemble mean

							bootstrap			bootstrap	
buoy	no.	mean	bias	rmse	regr.	skill	skill	stde	scatter	scatter	stde
		(m)	(m)	(m)					(%)	(%)	(%)
Fin	58	2.97	0.14	0.66	0.82	0.91	0.914	0.030	22	21.03	5.07
Elb	61	2.48	0.08	0.42	0.93	0.95	0.947	0.020	17	16.41	2.63
Wes	63	2.34	0.51	0.72	0.65	0.86	0.861	0.020	22	21.46	2.44
Hel	63	2.34	0.27	0.64	0.77	0.87	0.865	0.033	25	24.61	3.01
Bun	36	1.48	0.30	0.37	0.70	0.84	0.820	0.060	14	13.79	1.94

stde : standard error of the mean of the bootstrap samples (1000 bootstrap replicates)

For the Xaver time period four additional control runs have been performed by running CWAM with the four different sets of EWAM-EU boundary values, but with identical wind fields without perturbations. The aim was to check whether one of the global models would be the reason for CWAM results that are closer to the measurements. Figure 14 gives a comparison of the measured significant wave heights at the location Helgoland with the control runs depending on the four different global models (the fifth control run has been generated by the routine EWAM and not by EWAM-EU). The control runs driven by the wind fields provided by the routine COSMO-DE model show a remarkable overestimation of the significant wave heights during Xaver due to the overestimation in the wind speeds as shown already in figure 12b.

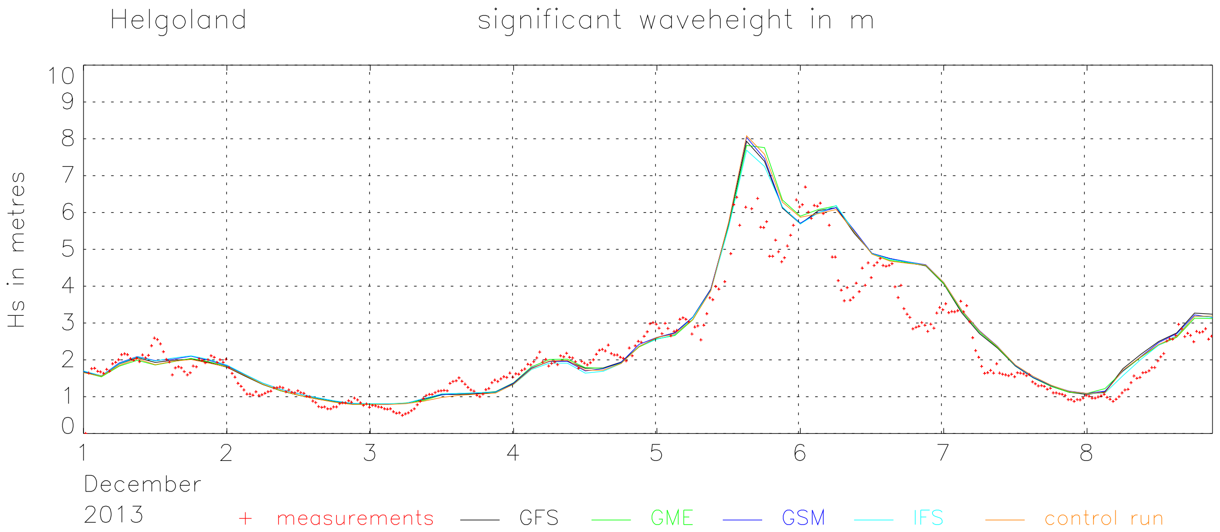


Fig. 14 : time series of H_s (model : control runs with different sets of boundary values but identical wind fields) at buoy Helgoland

Although the differences between the control runs are small only, the influence of the boundary values provided by the four different EWAM-EU models can be seen, especially at the peak of Xaver in the afternoon of the 5th of December. That is due to the strong north-westerly winds generating low frequency energy in the four EWAM-EU realisations that propagates as swell into the CWAM area. Further investigations taking into account only the five perturbed members of one of the rows shown in figure 2 that depend on one of the global models, show that an ensemble mean for these five members respectively leads actually to different statistics. Comparing the four ensemble means among themselves and to those computed for the significant wave height for all 20 members for Xaver in table 5 and 6, it becomes obvious that the forecast chain that depends on the global model IFS of the ECMWF generates CWAM results that are closest to the measurements. This can be proved by the corresponding statistics included in table 7.

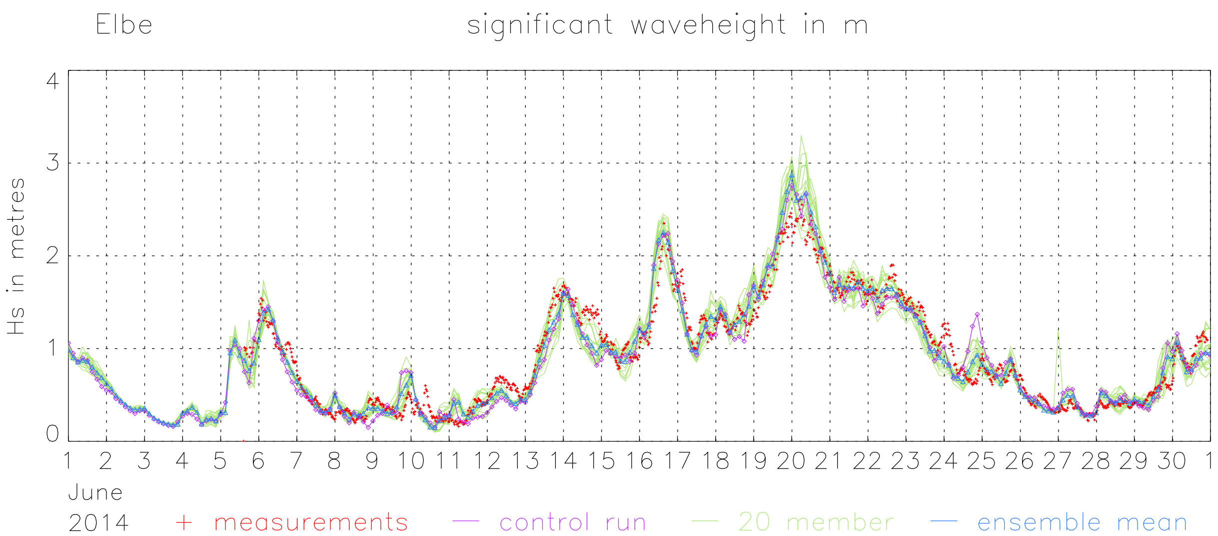
Table 7 : Xaver - one week statistics for the comparisons between measurements and ensemble mean of the five members only that depend on the global model IFS of the ECMWF							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	58	2.97	0.10	0.52	0.86	0.94	17
Elb	61	2.48	0.08	0.41	0.95	0.95	17
Wes	63	2.34	0.51	0.67	0.66	0.87	19
Hel	63	2.34	0.27	0.63	0.78	0.86	25
Bun	36	1.48	0.32	0.40	0.70	0.81	17

To investigate finally the effect of the five selected perturbations in the COSMO-DE-EPS members on the wave model results, an ensemble mean of those CWAM members has been generated, that use different boundary values, but in each case the driving wind fields perturbed by the same modification. The statistical analysis of the mean of the five corresponding CWAM members compared with the measurements make obvious that the effect due to the change of the default value of the asymptotic mixing length of turbulence from 150 m to 500 m (no. 5 in table 1) induces the highest wind speeds over sea. Therefore the overestimation of the wave heights in this case is more pronounced, the bias is much higher than for the other perturbations and the statistical parameters are worst as well. The corresponding values are given in table 8. Definitely the best statistics could be achieved using the COSMO-DE-EPS members perturbed by an increase of the default scaling factor of the laminar boundary layer by a factor of 10. That reduces the wind speed and induces a much better bias for the significant wave heights and improves the other statistical parameters as shown in table 9. The differences in the statistics generated for the other 3 perturbations are small only and the values are comparable to those obtained for the ensemble mean of all 20 members. For significant changes in the wind fields above sea, the perturbations in the COSMO-DE-EPS due to the entrainment rate for shallow convection and the critical value for normalized oversaturation are less useful than changes in the asymptotic mixing length of turbulence and in the scaling factor of the laminar boundary layer for heat.

Table 8 : Xaver - one week statistics for the comparisons between measurements and ensemble mean of the five members only that depend on the same perturbation no.5 (table 1) of the COSMO-DE-EPS							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	58	2.97	0.18	0.70	0.79	0.91	23
Elb	61	2.48	0.16	0.47	0.87	0.94	18
Wes	63	2.34	0.56	0.76	0.63	0.85	22
Hel	63	2.34	0.35	0.72	0.72	0.84	27
Bun	36	1.48	0.36	0.43	0.66	0.80	17

Table 9 : Xaver - one week statistics for the comparisons between measurements and ensemble mean of the five members only that depend on the same perturbation no.4 (table 1) of the COSMO-DE-EPS							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Fin	58	2.97	0.07	0.62	0.87	0.92	21
Elb	61	2.48	-0.03	0.42	1.03	0.94	17
Wes	63	2.34	0.47	0.69	0.67	0.87	22
Hel	63	2.34	0.17	0.55	0.84	0.89	23
Bun	36	1.48	0.27	0.33	0.74	0.85	13

Since March 2014 the forecast time period of the atmospheric ensemble member has been extended by 6 hours from 21 to 27 forecast hours. A side effect of the new configuration is the possibility to use one-hourly wind fields as the driving force for the wave ensembles instead of the three-hourly winds that have been used before. To check whether the new configuration has an important influence on the results of the wave ensemble computations, additional wave model results have been generated for June 2014. Figure 15 shows as a representative example the time series of measured and computed significant wave heights, together with the corresponding driving wind of all members, including ensemble mean and control run at the location Elbe. Remarkable differences compared with the previous results for April 2013 and Xaver can't be found for the new configuration, although the maximum, minimum and the spread given in figure 16 at the location Elbe indicate, that the differences in the wave model results for the 20 wave ensemble member are a little bit more pronounced in June 2014. The statistics at the five locations included in table 10 and 11 shows definitely a better agreement between measurements and ensemble mean than between measurements and control run.



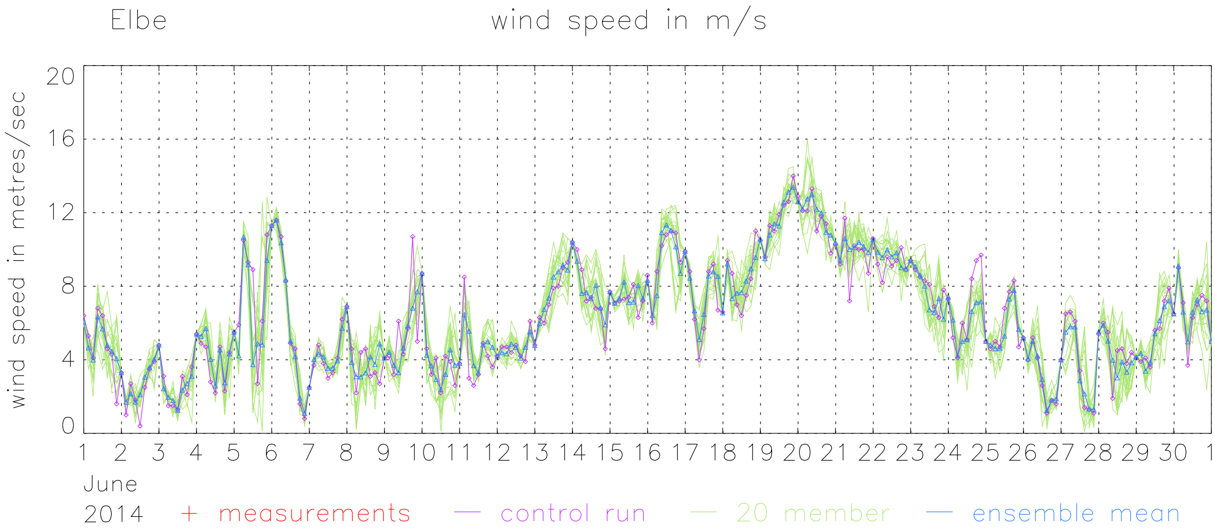


Fig. 15 : time series of H_s (a. above) and wind speeds (b. below) at buoy Elbe

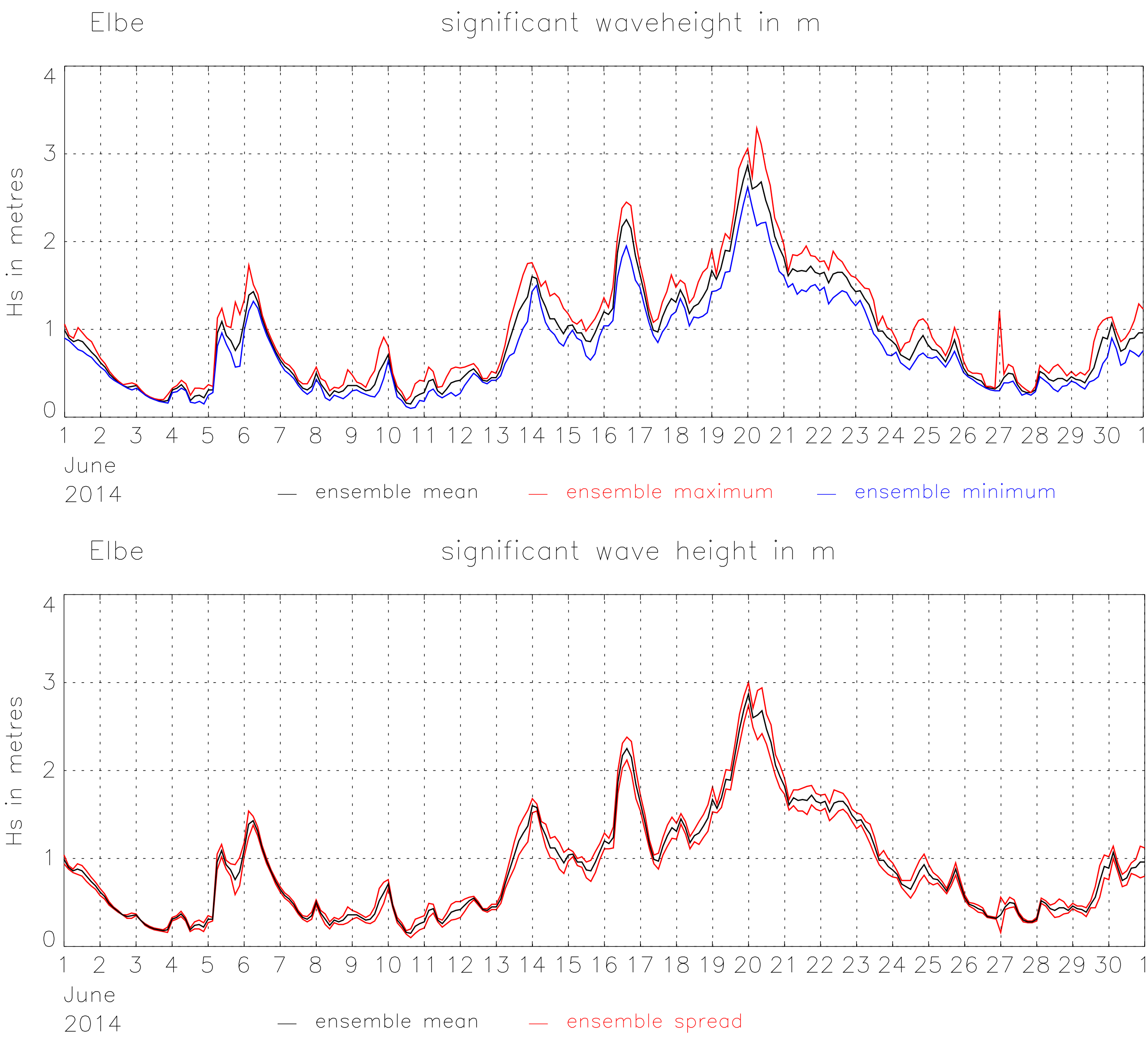


Fig. 16 : H_s-maximum and minimum of the 20 members at buoy Elbe (a. above) and the ensemble spread (below)

Table 10 : H_s statistics for the comparisons between the measurements and the 'control run'							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)	(m)			(%)
Elb	203	0.96	-0.01	0.19	0.99	0.90	20
Wes	239	0.90	0.05	0.23	0.81	0.90	25
Bun	127	0.73	0.01	0.16	0.94	0.92	22
Hoe	240	0.22	0.06	0.11	0.72	0.63	41
Dar	240	0.49	0.00	0.11	0.99	0.91	22

Table 11 : H_s statistics for the comparisons between the measurements and the 'ensemble mean'							
buoy	number	mean	bias	rmse	regr.	skill	scatter
		(m)	(m)				(%)
Elb	203	0.96	0.00	0.16	0.97	0.93	16
Wes	239	0.90	0.07	0.22	0.79	0.91	23
Bun	127	0.73	0.03	0.15	0.89	0.93	20
Hoe	240	0.22	0.06	0.11	0.72	0.63	38
Dar	240	0.49	0.00	0.10	1.00	0.91	21

At the two locations Fino and Helgoland no wave data has been recorded during June 2014.

5. Summary and conclusions

A local wave ensemble system with 20 individual members has been developed and applied to the German Bight and western Baltic Sea. The model used for the investigation is the third generation spectral wave model WAM. The driving wind fields are computed by the atmospheric members of the COSMO-DE-EPS that are available at the DWD operationally. Since a wave model loses the information of the initial wave field very fast, in this investigation the wave ensemble computations have been done concentrating on the variability of the wind forcing only. The complete wave ensemble chain including the global model, the four regional models and the 20 members for the local scale have been calculated for April 2013, June 2014 and a one week period around the extreme storm Xaver in December 2013. The aim was to check whether such a system could improve the operational wave forecast routine at the DWD. The results of all the wave model predictions have been compared with buoy measurements and showed a good agreement for the selected wave parameters in moderate wind situations. During storm events like Xaver in December 2013 it became obvious, that in the shallow parts near the coast the tidal modulations in the measurements are not predicted by the wave model because the water level variations and the currents are not included in the CWAM setup. To achieve an improvement on this, a coupled system of a circulation model and a wave model is required.

The differences between the results of the individual wave ensemble members are generally small with a tendency to enhanced values in storm situations as discussed for the extreme storm Xaver. The same is valid for the spread of the ensemble mean which usually is definitely very small during the one day prediction period of the wave model runs. Unfortunately the forecast period of the atmospheric COSMO-DE-EPS members was 21 hours only up to March 2014 and 27 hours afterwards and therefore currently too short for a reasonable investigation of the spread development in time for the different wave members. An extension of those from 27 to 45 hours is planned already.

The results of the wave ensemble computations show with the ensemble mean of the 20 individual members in all cases slightly improved results compared with those obtained by the control run. In further tests it was found that one of the four global models (the IFS of the ECMWF) is responsible for CWAM results that are closer to the measurements and that one of the five physical modifications in the COSMO-DE-EPS is more useful than the others. The perturbation by an increase of the default scaling factor of the laminar boundary layer generates reduced wind speeds, reduced bias in the wave forecasts and therefore better statistics.

Although this investigation finally proves that such an ensemble prediction system for ocean surface waves could improve the operational wave forecast system of the DWD, the wave ensemble calculations should be tested continuously for a longer time period to enable an authoritative answer whether such an ensemble mean is always better than a single routine forecast. Since the consumption of computer resources is very high for such a wave ensemble system it is necessary to balance costs and benefit carefully.

Acknowledgements The computed wind fields of all the atmospheric models that are used to drive the wave models of the ensemble prediction system and the wind measurements as well are kindly made available by the German Met Service (DWD: Deutscher Wetterdienst). I thank Thomas Bruns, Thomas Hanisch and Susanne Theis from the DWD for support and valuable discussions. I would also like to thank the two anonymous reviewers. Their constructive comments helped to improve the manuscript.

Appendix

Definition of the statistical parameters computed for the comparisons between measured data and wave model results :

mean of measurements and model values	$\bar{x} = \frac{1}{n}\sum x_i \;,\; \bar{y} = \frac{1}{n}\sum y_i$		(A1)
Bias	$bias = \bar{y} - \bar{x}$		(A2)
root mean square error (rmse)	$rmse = \left[\frac{1}{n}\sum (y_i - x_i)^2\right]^{0.5}$		(A3)
slope of regression line	$regr. = \left[\frac{\sum y_i^2}{\sum x_i^2}\right]^{0.5}$		(A4)
reduction of variance (skill)	$rv = 1 - \frac{\sum (x_i - y_i)^2}{\sum (x_i - \bar{x})^2}$		(A5)
standard deviation	$\text{std} = \left[\frac{1}{n-1}\left((x_i - \bar{x}) - (y_i - \bar{y})\right)^2\right]^{0.5}$		(A6)
scatter index	$sci = \frac{std}{\bar{x}} * 100$		(A7)
ensemble mean	$\bar{y} = \frac{1}{n}\sum y_i$		(A8)
spread	$\text{spread} = \left[\frac{1}{n-1}(y_i - \bar{y})^2\right]^{0.5}$		(A9)

References

Alves J-EGM, Wittman P, Sestak M, Schauer J, Stripling S, Bernier NB, McLean J, Chao Y, Chawla A, Tolman H, Nelson G, Klotz S (2013) The ncep-fnmoc combined wave ensemble product: expanding benefits of interagency probabilistic forecasts to the ocean environment. Bull Amer Meteor Soc 94: 1893-1905.
DOI: http://dx.doi.org/10.1175/BAMS-D-12-00032.1

Baldauf M, Seifert A, Förstner J, Majewski D, Raschendörfer M, Reinhardt T (2011) Operational convective-scale numerical weather prediction with the COSMO model. Monthly Weather Review 139: 3887-3905.

Bertotti L, Bidlot J-R, Buizza R, Cavaleri L, Janousek M (2011) Deterministic and ensemble-based predictions of Adriatic Sea sirocco storms leading to 'acqua alta' in Venice. Quarterly Journal of the Royal Meteorological Society 137: 1446-1466. DOI:10.1002/qj:861.

Bidlot J-R, Janssen PAEM, Abdalla S (2007) A revised formulation for ocean wave dissipation and its model impact. ECMWF, Technical Memorandum Nr. 509.

Breivik O, Aarnes OJ, Bidlot JR, Carrasco A, Saetra O (2013) Wave Extremes in the Northeast Atlantic from Ensemble Forecasts. J Climate 26: 7525-7540.
DOI: http://dx.doi.org/10.1175/JCLI-D-12-00738.1.

Cao D, Chen HS, Tolman H. (2007) Verification of Ocean Wave Ensemble Forecast at NCEP. Technical Note Nr. 261.

Cao D, Tolman H, Chen, HS, Chawla A, Wittmann P (2009) Performance of the Ocean wave ensemble forecast at NCEP. Technical Note Nr. 279.

Carrasco A, Saetra O (2008) A limited-area wave ensemble prediction system for the Nordic Seas and the North Sea. Norwegian Meteorological Institute, Report 22/2008. 29 pages.

Carrasco A, Saetra O, Bidlot J-R (2011) Wave Ensemble Predictions for safe Offshore Operations. Proc. 12th International Workshop on Wave Hindcasting and Forecasting and 3rd Coastal Hazard Symposium. Kohala Coast Hawaii, October 30 – November 4. Available at http://www.waveworkshop.org/12thWaves/index.htm.

Chen, HS (2006) Ensemble prediction of ocean waves at NCEP. Proc 28th Ocean Engineering Conf, Taipei, Taiwan, NSYSU: 25-37

Durrant TH, Woodcock F, Greenslade DJM (2009) Consensus forecasts of modelled wave parameters. Weather Forecasting 24: 492-503.

Efron, B and Tibshirani, RJ (1993) An introduction to the bootstrap. Boca Raton: Chapman & Hall/CRC.

Farina L (2002) On ensemble prediction of ocean waves. Tellus 54A: 148-158.

Gebhardt C, Theis SE, Paulat M, Ben Bouallègue Z (2011) Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. Atmospheric Research 100, 168-177.

Grabemann I, Weisse R (2008) Climate Change impact on extreme wave conditions in the North Sea: An ensemble study. Ocean Dynamics 58: 199-212.

Günther H, Hasselmann S, Janssen PAEM (1992) The WAM Model Cycle 4.0. User Manual. Technical Report No. 4, Deutsches Klimarechenzentrum, Hamburg, Germany. 102 pages.

Hersbach H, Janssen PAEM (1999) Improvements of the short fetch behaviour in the WAM model. J Atmos Oceanic Tech 16: 884-892.

Hoffschildt M, Bidlot J-R, Hansen B, Janssen (1999) Potential benefits of ensemble forecasting for ship routing. ECMWF, Technical Memorandum 287.

Janssen PAEM (1999) Wave modelling and altimeter wave height data. ECMWF, Technical Memorandum Nr. 269.

Janssen PAEM (2008) Progress in ocean wave forecasting. J Comput Phys 227:3572-3594.

Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM (1994) Dynamics and Modelling of Ocean Waves, Cambridge University Press.

Lenartz F, Beckers J-M, Chiggiato J, Mourre B, Troupin C, Vandenbuicke L, Rixen M (2010) Super-ensemble techniques applied to wave forecast : performance and limitations. Ocean Science 6: 595-604.

Peralta C, Ben Bouallègue Z , Theis SE , Gebhardt C, Buchhold M. (2012) Accounting for initial condition uncertainties in COSMO-DE-EPS, J Geophys Res 117 (D7). DOI:10.1029/2011JD016581

Saetra O, Bidlot J-R (2004) On the potential benefit of using probabilistic forecast for waves and marine winds based on the ECMWF ensemble prediction system. Weather Forecasting 19: 673-689.

Saetra O, Hersbach H, Bidlot, J-R, Richardson D (2004) Effects of observation errors on the statistics for ensemble spread and reliability. Monthly Weather review 132: 1487-1501.

Schättler U, Doms G, Schraff C (2009) A description of the nonhydrostatic COSMO-Model. Part VII: User's guide, report: 147 pp., Consort for Small-Scale Modell., Dtsch Wetterdienst, Offenbach, Germany. Available at http://www.cosmo-model.org/content/model/documentation/core/cosmoUserGuide.pdf.

WAMDI group: Hasselmann S, Hasselmann K, Bauer E, Janssen PAEM, Komen GJ, Bertotti L, Guillaume A, Cardone VC, Greenwood JA, Reistad M, Zambreski L, Ewing J (1988) The WAM model – a third generation ocean wave prediction model, J Phys Oceanogr 18: 1775-1810.