Norwegian Meteorological Institute met.no

met.no report

Report no. 22/2008
Meteorology and Oceanography
ISSN: 1503-8017
Oslo, November 25, 2008

A limited-area wave ensemble prediction system for the Nordic Seas and the North Sea

Ana Carrasco and Øyvind Saetra

Norwegian
Meteorological Institute
met.no

report

Number	Subject	Date	Classification	ISSN
22/2008	Meteorology and Oceanography	November 25, 2008	☒ Open ☐ Restricted ☐ Confidential	1503-8017

Title

A limited-area wave ensemble prediction system for the Nordic Seas and the North Sea

Authors

Ana Carrasco and Øyvind Saetra

Client(s)	Client reference
met.no	Validation of wave EPS

Abstract

A regional ensemble prediction system for ocean waves (WAMEPS) has been set in operation at the Norwegian Meteorological Institute (met.no). The system is forced with winds from the atmospheric limited area ensemble prediction system (LAMEPS) also in operation at met.no. The LAMEPS is forced on the lateral boundaries with global ensemble predictions with singular vectors that are targeted for Northern Europe. In this study, the forecast performance of the WAMEPS over one year, 2007, is assessed using available buoy data within the model domain. One of the major finding is that foretasted probabilities are reliable, with small tendency of forecasting probabilities larger that the observed frequencies. It is demonstrated that by tailoring the forecasts to a specific user, finding the optimal probability for decision making, the probability bias can be accounted for. In the context of relative economic value, the ensemble forecasts always beats the deterministic forecasts. The rank histograms reveal a bias as the observations are too seldom observed in the highest ranks. By separating the observations into regions with different variability it is demonstrated that the source for this bias are data from shelter areas and areas close to land. Treating either the ensemble mean or the unperturbed (control) member as a deterministic forecast, the WAMEPS shows a clear spread-skill relation, indicating that the ensemble spread gives a good indication of the uncertainty of a deterministic forecast.

Keywords

Wave forecast, Ensemble forecast, Validation

Disiplinary signature	Responsible signature
Lars Petter Røed, Section for Oceanography	Øystein Hov, Head of Research

Postal address PO Box 43 Blindern N-0313 Oslo Norway
Office Niels Henrik Abels vei 40
Telephone +47 2296 3000
Telefax +47 2296 3050
e-mail: met.inst@met.no **Web:** met.no
Bank account 7695 05 00601
Swift code DNBANOKK

Contents

List of Figures

1 Introduction

The limited predictability of the atmospheric system was first fully recognised by Lorenz (1963). Using a simple model with three non-linearly depended variables, he demonstrated that small errors in the initial state, introduced by very small truncation errors, eventually lead to solutions that are completely different. The errors introduced by imperfections in the observational system is an important factor for the errors growth and limited predictability of weather forecasting systems. As computer capacity has been steadily growing the last decades, Monte-Carlo simulations of the atmospheric state has become a feasible method for dealing with the error growth in forecasts in a probabilistic way. By perturbing the initial state, a number of simulations can be integrated and used to estimate the atmospheric probability distribution.

In 1980, the European Centre for Medium-Range Weather Forecasts (ECMWF) started the first operational ensemble prediction system (EPS), producing 10-day forecasts. Currently, a number of centres provide global EPS forecasts on routine basis. In 1998, the ECMWF atmospheric model was coupled to an ocean wave model. The main reason for this coupling was a positive impact on both wave and atmosphere forecasts (Janssen et al., 2002). A bi-product of the introduction of wave-atmosphere coupling in the forecasting system at ECMWF was a global operational EPS for ocean waves. The potential benefits for marine operations of this system was demonstrated by Saetra & Bidlot (2004). In the ECMWF wave ensemble, all ensemble members uses the unperturbed analysis as initial condition and the spread of the ensemble members is therefore due only to different atmospheric forcing. Farina (2002) investigated the impact of perturbing the initial spectra for wave EPS and concluded that third-generation wave models are essentially insensitive to the spectral initial condition.

At the Norwegian Meteorological Institute (met.no) a 21-member limited-area EPS (LAMEPS) for the atmosphere was set into operational use in February 2005 (Frogner and Iversen, 2002; Frogner et al., 2006), providing 60 hour forecasts on a daily basis. This system is forced on the lateral boundaries with global ensemble predictions that are initiated with targeted singular vectors (TEPS). The targeting area for TEPS is Northern Europe. A limited-area wave ensemble forecasting system (WAMEPS) forced by the LAMEPS was set into operational use in January 2008, producing 60 hour forecasts daily. In contrast to the global wave ensemble from ECMWF, the WAMEPS is run in an uncoupled configuration using the 10-metre wind speed from LAMEPS as external forcing, but provides no information back to the atmospheric model.

The main advantage of using limited area models is that it allows for substantially higher model resolution than what is possible for global models. In the context of wave forecasting this is expected to improve the winds, and hence the waves, particularly in the near coastal areas where orographic effects may significantly alter both the wind speed and the direction. Also the fact that the system is targeted towards Norther Europe is expected to improve the forecasts relative to a global system.

The main objective of this investigation is to study the performance of the limited-area forecasting system by comparison with independent buoy and platform observations over one year. Valid statistics of ensemble prediction systems requires a large number of observations. To obtain this, the system has been re-run for the whole of 2007 to produce daily forecasts in

Hindcast mode, using archived LAMEPS forecasts as wind forcing. We will also demonstrate the forecasting system by looking closer at the performance during two events with large observed significant wave height.

One important question to address is whether the targeted limited-area system actually represents an improvement over a lower resolution global system. No direct comparison with a global system is performed, This requires that exactly the same observations and time-span are used. However, the work by Saetra & Bidlot (2004) gives some statistical results for the global ECMWF wave ensemble. We discuss the results obtained in the present study in light of results for the global system. Although no definite conclusions can be drawn due to the lack of a proper intercomparison study, the results may indicate how the limited-area system compares to the global system.

The structure of the presentation is as follows: Section 2 gives a description of the wave ensemble prediction system (WAMEPS) and the set-up for re-running 2007. Section 3 describes the observations used in the analysis. Section 4 study the performance of WAMEPS for to events with waves above 10 metres. In section 5, the forecasts are assessed against observations using standard techniques for EPS verification. Finally, the conclusions are drawn in section 6.

2 Model System

2.1 The Atmospheric Forcing

LAMEPS is a limited-area atmospheric ensemble prediction system for Northern Europe. In addition to continental Northern Europe and the Scandinavian Peninsula the model domain covers the Nordic Seas including the adjacent North Sea and the Barents Sea (Fig. 1). LAMEPS integrates 20 ensemble members from perturbed initial conditions in addition to the control run that uses the unperturbed analysis as the initial state. The model domain is on a rotated spherical grid centred at 65°N and 0°E. The horizontal resolution of the LAMEPS in 2007 was 0.2° in both directions. At the lateral boundaries the LAMEPS is forced by a low-resolution ensemble prediction system based on targeted singular vectors (TEPS). The TEPS is designed to maximise the total energy norm of the perturbations for Northern Europe. A more detailed description of LAMEPS is given by Frogner and Iversen (2002).

2.2 The Wave Component

The wave model is the met.no version of the ocean wave prediction model WAM cycle 4. The WAM model marked the introduction of so-called third generation wave models which explicitly accounts for the non-linear interaction between the wave components (Komen et al., 1994). It solves the wave energy equation without any prior assumption about the shape of the wave spectrum. The model was set into operation at met.no in 1998 and is currently run for a number of limited-area domains with various horizontal resolutions. The model is implemented with deep-water physics and the spectral resolution is 24 directional and 25 frequency bins.

The WAMEPS covers the same model domain as the LAMEPS, (Fig. 1). The horizontal resolution is 0.1°. The hourly 10-metre winds form the LAMEPS are used as external forcing and all ensemble members are started from the same initial condition. The current version of WAMEPS is run without data assimilation and the initial condition is the latest 24-hour forecast. The systems is run daily and the forecast length is 60 hours. On the lateral boundaries all ensemble members are forced with a lower resolution forecast for the North Atlantic Ocean (large domain in Fig. 1). Ideally, a limited-area wave ensemble should be forced by a global, or a larger domain, ensemble prediction system on the boundaries. At met.no no such system is currently available and the use of one deterministic forecast as boundary condition for all members is a compromise that is applied until another option becomes possible. Future plans includes disseminating wave spectra from the ECMWF system for use at the boundaries. This requires huge amount of data to be routinely transferred and is not possible with the present data infrastructure. However, the reader should keep in mind the possible limitations imposed by the use of one single deterministic forecast as boundary value for all ensemble members.

3 Observations

Wave observations from moored buoys, ships and platforms are routinely collected by national organisations in their offshore areas of interest. Hourly wave data are transferred to meteorological centres via the Global Telecommunication System (GTS) and archived together with all other synoptic observations. In the remainder of this paper, the word buoy is used to refer to the selected moored buoys and platforms since most of the reliable observations comes from these observations. Note however, that the observation principle for waves is different for buoys than for platforms. Buoys rely on time series analysis of the buoy motion to derive the spectra whereas platform observations uses radar imaging of the sea surface to derive the spectra.

A number of the available observations are either outside the WAMEPS model grid or located in shallow areas near the coast. Still, about 58 stations report data that are well within the model grid and are located in relatively deep water (depth of 100 m or more). Observations in 100 metres of depth or more are required since the model is set up in deep water mode. Fig. 2 shows the data coverage within the model domain.

From the buoy records, time series are reconstructed to perform a basic quality check on the data (Bidlot et al., 2002). Spatial and temporal scales are made comparable by averaging the hourly observations in a window of 4 h centred around the validation time. Buoys exhibit a high-frequency variability on a time scale of 1 h. Not averaging the data results in a scatter between the model and observations, which can be linked to high frequency variability, not present in the model (Janssen et al., 1997). For a more detailed description of the data treatment, see Bidlot et al. (2002) and Saetra & Bidlot (2004). This investigation covers the period 1 January through 31 December 2007. Roughly 58 000 independent observations are used.

4 WAMEPS forecasts for the storms 11 and 12 January 2007

On the 11 and 12 January 2007 significant wave height around 11 metres were recorded in the buoys west of the British Isles two days in a row. The waves were generated by two different low pressure system that moved in a north-eastward direction over the North Atlantic Ocean towards Northern Europe. In Fig. 3 the mean-sea level pressure (MSLP) from the operational analysis at met.no for 12 UTC 11 January is shown. The position of one of the buoys which recorded waves above 10 metre is marked with a black dot in the figure. The buoy location is 55.40°N and 12.60°W. A low pressure between Scotland and Iceland yield north-westerly winds at the buoy location. The 10-metre wind speed in the operational analysis is about 20 m/s. The second low is seen in this plot outside the Island of Newfoundland. During the next 24 hours this low moves north-eastward and can be seen as the low-pressure centre between Shetland and Faeroe Island in the operational analysis for the next day, 12 UTC 12 January (Fig. 4). In the analysis, the wind at the buoy location is westerly with wind speeds of about 25 m/s.

The 48 hour forecast for the probabilities of waves above 10 metre for 11 January are given in Fig. 5. Fig. 6 shows the corresponding probabilities for 12 January. The differences in forecasted probabilities are striking. For 11 January the probabilities are as high as 80% in certain areas. For the second case, 12 January, the largest probabilities are around 50%. Accordingly, the model system predicts a much larger uncertainty for the latter case. This is further illustrated in the Fig. 7 where time series of all ensemble members for the buoy-location mentioned above are shown. The upper panel is the forecast started 18 UTC 9 January and the lower panel is the forecast started 18 UTC 10 January 2007. The black dots in this figure are the buoy observations. A thick vertical line marks the verification times for the probabilities in Figs. 5 and 6. Note the difference in the EPS distribution between these two cases.

To understand the reason for the large differences in forecast probabilities, the MSLP from all ensemble members of LAMEPS are plotted in Figs. 8 and 9. Fig. 8 shows the 48 hour forecasts valid at 18 UTC 11 January and Fig. 9 the corresponding forecasts for 12 January 2007. For the first case (Fig. 8) almost all members have a pronounced low pressure in, or in the vicinity of, the area north-east of the Iceland-Scotland ridge. Also the depth of the lows are more or less of the same order. Exceptions may be member 1 and 13 with rather weak depressions. For the second case (Fig. 9), the differences in both location and intensity of the low pressures are much larger. Apparently, the high variability and corresponding uncertainty in the wave forecasts is caused by the large spread in MSLP and location of low pressure centres. It is tempting to interpret this as two situations where the flow dependent stability is quite different. The two cases then, nicely illustrate the main motivation for introducing ensemble forecasts where the uncertainty is dependent on the specific weather situation and not only the observation site.

5 Results of the Ensemble Statistics

Ensemble predictions are estimates of the true probability distribution for given weather parameters as nature evolves. The problem in validating ensemble forecasts is the fact that the natural probability distributions are never observed. The observed weather events are one realisation from these distributions. When verifying ensemble forecasts, this has to be taken into consideration. One single observation can say nothing on the quality of the model system as observations are expected to fall in the low probability part of the distribution with a certain frequency for any well tuned system. The basic assumption when using the EPS is that for a perfect system the ensemble members and the observation are random draws from the same probability distribution. If this assumption is true, it is impossible to distinguish the observation from the ensemble members when using statistical methods. A number of consequences of this assumption can be outlined and used to test the EPS against observations. Another consequence of this is that the number of observations needed for proper verification statistics is far greater than needed to validate a deterministic forecast.

Below, this basic hypothesis will be tested in terms of ensemble spread, the relation between spread and skill and the reliability of the foretasted probabilities. We will also test the economic value of the forecasts when used for decision support.

5.1 Ensemble Spread

The ensemble spread will be tested by using rank histograms (Hamill, 2001). The idea is that for a given forecast range, the ensemble members and the observations are pooled and sorted from the lowest to the highest value. For a perfect system. the observation may occur at any rank with the same probability. If this process is repeated for a number of observations, a flat histogram over possible ranks is expected. Bias in the ensemble system will come out as a sloped rank histogram and lack of ensemble spread results in u-shaped histograms.

When using rank histograms, Saetra et al. (2004) demonstrated that observations errors can lead to a false impression of too low spread in the ensemble system. They suggested adding normal distributed noise, with the same standard deviation as the observation errors, to the ensemble prediction system before it is presented to the verification tools. A similar approach was also suggested by Anderson (1996). Janssen et al. (2003) estimated the observation errors for the buoys to be about 10% of the significant wave height. We will use this as a standard deviation for the observation errors when adding normally distributed noise to the ensemble members.

Rank histograms for the 48-hour forecasts are shown in Fig. 10. Fig. 11 shows the similar results for the 60 hour forecasts. Only minor differences between the two forecast range are detected. Clearly, the system is biased as the observed frequency in the highest ranks are too low. By looking at the results for the different groups it is obvious that the bias is most pronounced for group 3, smaller for group 2 and almost absent for group 1. Interestingly, the rank histograms closes to a flat distribution are group 1. Group 1 represents the Norwegian Sea which is exposed to swell and strong winds form the North Atlantic and is well known for its rather rough and highly variable wave climate. In the North Sea, represented by group 2, the wave climate is less harsh as the ocean basin is enclosed by land except in the case of northerly winds. The buoys around the British Isles, group 3, are located closer to land and are those least exposed to strong winds and large waves.

5.2 Spread-Skill Relation

An appealing way of using ensemble predictions is to take the spread as an indication on the expected quality of a deterministic forecast. When the ensemble spread is small forecasters can have confidence in the deterministic forecasts and vice versa. If this is the case, the ensemble prediction can also by used to calculate errors bars for time series of forecasted weather parameters, often referred to as meteograms, displaying the expected uncertainties dependent on the flow regime. The spread is defined as the standard deviation of the ensemble members about the ensemble mean. For a given forecast range, the spread is divided into bins of 0.15 metres. The forecast error is defined as the absolute value of the difference between the ensemble mean forecast and the observed value. The error distributions for the deterministic forecast are then obtained within each bin for the spread. The spread-skill relation is tested by plotting the 90 percentile of the forecasts errors as a function of ensemble spread. If a spread-skill relation is present, the 90 percentile will increase with increasing ensemble spread.

Fig. 12 shows the spread-skill for the 48 hour forecasts. In this figure, the ensemble mean is taken as the deterministic forecast. Similar figures were produced with the control forecast as the deterministic forecasts and the results were more or less identical. The groups represents results for the areas described in section 5.1. All plots reveal a clear spread-skill relation with a 90 percentile more or less following the 45° line. Similar results for the 60 hour forecasts are given in Fig. 13, which also yields a 90 percentile line close to the diagonal. These results strongly supports the idea that the ensemble spread can be interpreted as a measure of the expected accuracy of the deterministic forecast. The results are also in close agreement with those obtained by Saetra & Bidlot (2004) for the global wave ensemble.

5.3 Reliability

An important and very useful aspect of ensembles is that the probability of binary events can be forecasted, for instance the probability of waves exceeding a given threshold value. The simplest way of doing this is by counting the number of individual members that are larger than the threshold. More sophisticated methods have been suggested such as creating a continuous probability density function by dressing each member of the ensemble with its own statistical error distribution (Roulston and Smith, 2003).

To test how the forecasted probabilities corresponds to the observed frequency of the binary event, reliability diagrams are useful tools (Wilks, 1995). The forecasted probabilities are split into discrete bins from zero to one. For each forecast the probability class is determined and the observation is used to decide whether the binary event is true of false. By repeating this process over all observations the observed frequency in each probability bin is calculated. The observed frequency is then plotted against the forecast probabilities. The expected observed frequency is then 0.25 for the 25% probability class, 0.5 for the 50% probability class and so on. A perfect system then yields a diagonal line in the reliability diagram. A line below the diagonal indicates too low observed frequency and consequently too large forecast probabilities and vice versa.

The reliability diagrams for 48 and 60 hour forecasts are given in Figs. 14 and 15. Again the difference between the two forecasted ranges are minor. A small tendency of too high probabilities are seen as the observed frequencies are slightly below the diagonal line. This tendency is found for all threshold levels and for all sub-regions. This probability bias may actually not limit the value of the forecasts when used in decision support systems as will be discussed in the next section.

5.4 Relative Economic Value

Many important applications of weather and wave forecasts are often linked to some sort of decision making. Based on forecasts, the decision management may have to choose either to take action to protect against a given weather event or do nothing if the event does not occur. A user with access to probability forecasts must also decide on the probability threshold at which to take action. Should it be at 50%, or perhaps 80%? Is it possible to determine optimum probability on which to take action? By doing this, the probability forecast is actually converted to a deterministic forecast. An optimal system for the decision management is one that has a maximum number of hits and as few false alarms as possible.

Relative Operating Characteristic, or ROC diagrams, are constructed by calculating the hit rate H and false alarm rate F for a discrete number of probability classes from 0 to 1. The hit rate is defined as the fraction of occurrences of the event which were correctly forecasted, while the false alarm rate is the fraction of non-occurrences for which the event was incorrectly forecasted. The ROC diagram is then constructed by plotting the hit rate against the false alarm rate. The point $(0,0)$ corresponds to never forecasting the event, the point $(1,1)$ corresponds to always forecasting the event and the point $(0,1)$ represents perfect forecasts. Random forecasts with the sample climatological probabilities will have a curve along the 45° diagonal line. It is convenient to summarise a ROC diagram using a single scalar variable. The usual choice for this is the area under the ROC curve, say A. A perfect forecasting system have a ROC curve that includes the entire unit square, $A_{perf} = 1$. ROC curves for random forecasts along the diagonal line have the area $A_{rand} = 0.5$.

Here, the ROC areas have been calculated for the threshold values 2, 4, 6 and 8 metre. The results are listed in Tables 1 and 2. The tables also list the number of cases when the observed wave height exceeded the threshold value and the total number of observations used. The ROC area values are surprisingly large. One can not exclude the possibility that too few observations have been considered. In particular for the higher threshold classes this may be a problem as number of observed cases always will be relatively few. Next we will demonstrate how these ROC area scores affect the potential value of the ensemble forecasting system.

Richardson (2000) proposed a method to assess the economic value of deterministic and probability forecasts. The idea is to consider a hypothetical decision maker who must choose to take action or do nothing. If the events take place and no action to protect is taken, a loss L is incurred. Taking action to protect involve a cost, C. An example could be a supply ship that operated between land and an offshore platform. For waves exceeding a certain value, the operator may experience a loss due to damage on the vessel. Taking action to protect could involve postponing the operation which incurs a cost due to delay etc. The operator wishes to minimise the expenses over a large number of cases. For a given choice of L and C the observations could be used to calculate the expenses experienced when a certain forecast system is used to take the decisions. If the event is forecasted the cost C is added to the expenses. If the event occurs and is not forecasted the cost L is added to the expenses. This process is then repeated for a discrete number of cost-loss ratios, $\alpha = C/L$. The relative economic value of a given forecasting system is defined as

$$V = \frac{E_{climate} - E_{forecast}}{E_{climate} - E_{perfect}}. \tag{1}$$

Here, $E_{climate}$ is the economic value of using the sample climate to decide whether to take action or not, $E_{forecast}$ is the economic value of the forecasting system considered and $E_{perfect}$ is the economic value of a hypothetical perfect forecasting system. A perfect system has relative economic value 1 and the climatological system has value 0. The relative economic value is closely related to the ROC curves discussed above as it can be expressed in terms of the hit and false alarm rates. For more details on this the reader is referred to Richardson (2000).

The relative economic value for the 48 and 60 hour forecasts are presented in Figs. 16 and 17. In this case, the deterministic forecasts are represented by the control runs, i.e. the ensemble member based on the unperturbed atmospheric analysis. Note that the relative economic values of the ensemble forecasts are larger that those for the deterministic forecasts for all threshold values and cost-loss ratios. As for the ROC areas, the relative economic values are relatively high. Maximum values are around 0.8 which is approximately 0.15 larger than the values obtained by Saetra & Bidlot (2004). However, they presented results for forecast day 5 and can not be directly compared to those obtained here.

In section 5.3 it was shown that the forecasted probabilities were generally higher than the observed frequencies. The high scores for the relative economic values demonstrated that the forecasted probabilities may have a bias and still be of value to end-users. It is however, necessary to tailor the forecasts for specific end users.

6 Conclusions

A limited-area wave ensemble prediction system, WAMEPS, is implemented and set into operation at the Norwegian Meteorological Institute. The WAMEPS is forced with winds form the limited-area atmospheric ensemble system (LAMEPS) which dynamically downscale a global ensemble that is initiated with targeted singular vectors for Northern Europe. The system integrates 21 members to produce 60 hours probability forecasts for ocean waves daily.

On the 11 and 12 January 2007 waves above 10 metres were observed in the buoys west of the British Isles two successive days. The waves were generated by two different low pressure systems moving over the area. This gave us a nice opportunity to study how the model system handles forecast variability in two different flow regimes. In the first case, the system forecasted waves above 10 metres with a probability of almost 80%. The second case was far more uncertain with probabilities of about 50%. In both cases the observation was inside the range spanned by the ensemble members. We take this variability as evidence of flow dependent instability in the forecasting system.

The ensemble forecasts is assessed against buoy and platform observations available within the model domain. To obtain sufficient amount of data, the system was re-run for the whole 2007 which is the period covered in this study.

The ensemble spread is tested using rank histograms (Hamill, 2001). To avoid spurious over-population of the lowest and highest ranks due to observation errors (Saetra et al., 2004), normal distributed noise with a standard deviation of 10% of the wave height is added to the ensemble members before ranking the data. The rank histograms reveals a rather strong bias with too few observations in the higher ranks. Separating the data into regions of different expected variability, indicates that the bias is most pronounced for sheltered regions and areas closest to land. The Norwegian Sea, the area with the roughest wave climate, displays more even rank histograms.

Ensemble spread is often interpreted as the flow dependent uncertainty of a deterministic forecast. To test this the spread-skill relation was calculated as 90 percentile of the absolute error of the deterministic system as a function of the ensemble spread. Here, the control forecast and the ensemble mean are taken to represent the deterministic forecasts. The deterministic error distribution showed a clear dependence on the ensemble spread. In this context, the results of the global system found by Saetra & Bidlot (2004) are very similar to those obtained here. This is indeed strong evidence of a real spread-skill relation in the model system and justifies the use of the ensemble spread as an indicator of the expected error range in deterministic forecasts.

Reliability diagrams have been used to test how the forecast probabilities correspond to the observed frequency of predefined binary events such as wave height above a given threshold value. Generally, the reliability is high, but there is a slight tendency for the system to forecast too high probabilities. This bias was obtained for all wave thresholds and forecast ranges and appears to be a persistent feature of the forecasting system.

The biased probability forecasts does not reduce the potential value of the WAMEPS in a decision process. By tailoring the forecast to specific users, i.e. finding the optimal probability threshold given the users cost-loss ratio, the bias in the probabilities can be accounted for. This is clearly demonstrated in the very high ROC areas and relative economic values obtained in this study. Our main conclusion is therefore that limited-area wave ensemble forecast positively contributes to improve probabilistic wave forecasts. As wave ensembles are comparatively cheep to run on computers, it is indeed a worthwhile approach that provides valuable information to risk management and decision support for users such as the offshore industry.

7 acknowledgements

We like to thank Jean Bidlot at the ECMWF for providing us with quality controlled observations for this investigation.

References

Anderson, J. L., A Method for Producing and Evaluating Probabilistic Forecasts from Ensemble Model Integrations, 1996, J. Climate, 9, 1518-1530.

Bidlot, J. R., D. J. Holmes, P. A. Wittmann, R. L. Lalbeharry, & H. S. Chen, Intercomparison of the performance of operational ocean wave forecasting systems with buoy data, 2002, Wea. Forecasting, 17, 287-310.

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

Frogner, I-L., & T. Iversen, High-resolution limited-area ensemble predictions based on low-resolution targeted singular vectors, 2002, Q. J. R. Meteorol. Soc., 128, 1321-1341.

Frogner, I-L., H. Haakenstad & T. Iversen, Limited-area ensemble predictions at the Norwegian Meteorological Institute, 2006, Q. J. R. Meteorol. Soc., 132, 2785-2808.

Hamill, T. M., Interpretation of Rank Histograms for Verifying Ensemble Forecasts, 2001, Mon. Wea. Rev, 129, 550-560.

Komen, G., J. L. Cavaleri, M. Donelan, K. Hasselman, S. Hasselman, & P. A. E. M. Janssen, Eds., Dynamics and Modelling of Ocean Waves, Cambridge University Press, 533 pp.

Janssen, P., B. Hansen, & J.-R. Bidlot, Verification of the ECMWF forecasting system against buoy and altimeter data, 1997, Wea. Forecasting, 12, 763-784.

Janssen, P., J. D. Doyle, J. Bidlot, B. Hansen, L. Isaksen & P. Viterbo, Impact and feedback of ocean waves on the atmosphere, 2002, Atmosphere-Ocean Interaction, N. Perrie, Ed., Advances in Fluid Mechanics, Vol. I, WIT Press, 155-197.

Janssen, P., S. Abdalla, & H. Hersbach, Error estimation of buoy, satellite, and model wave height data, 2003, ECMWF Research Dept. Tech Memo. 402. 17 pp.

Lorenz, E. N., Deterministic nonperiodic flow, 1963, J. Atmos. Sci., 20, 130-141.

Richarson, D. S., Skill and relative economic value of the ECMWF ensemble prediction system, 2000, Q. J. R. Meteorol. Soc., 126, 649-667.

Roulston, M. R., & L. A. Smith, Combining dynamical and statistical ensembles, 2003, Tellus, 55A, 16-30.

Saetra, Ø., H. Hersbach, J-R. Bidlot, & D. S. Richardson, Effects of Observation Errors on the Statistics for Ensemble Spread and Reliability 2004, Mon. Wea. Rev, 132, 1487-1501.

Saetra, Ø., & Bidlot J-R. Potential Benefit of Sig Probabilistic Forecasts for Waves and Marine Winds Based on the ECMWF Ensemble Prediction System, 2004, Wea. Forecasting, 19, 673-689.

Wilks, D. S., Statistical Methods in the Atmospheric Sciences: An Introduction, Academic Press, 467 pp.

8 Figures

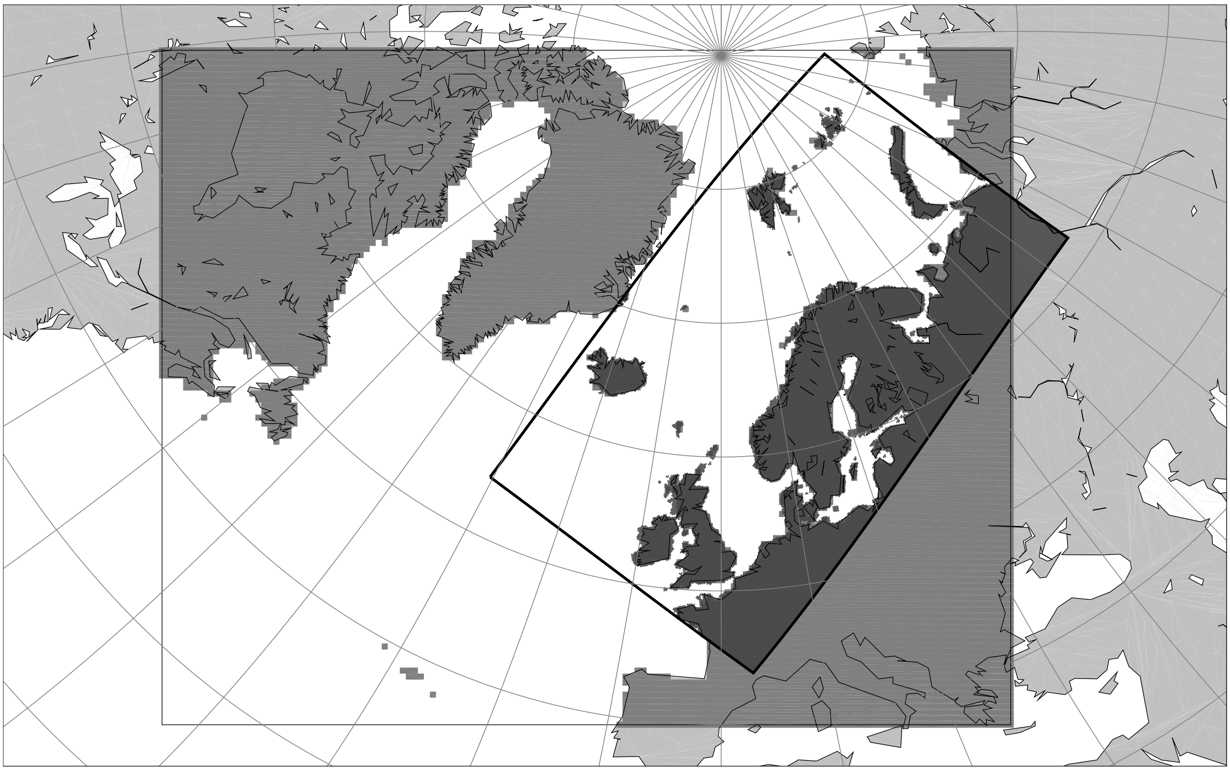

Figure 1: The small domain covering the Nordic Seas is the model domain of WAMEPS. The large area is the model domain of the coarse resolution wave model providing lateral boundary values for the the nested WAMEPS.

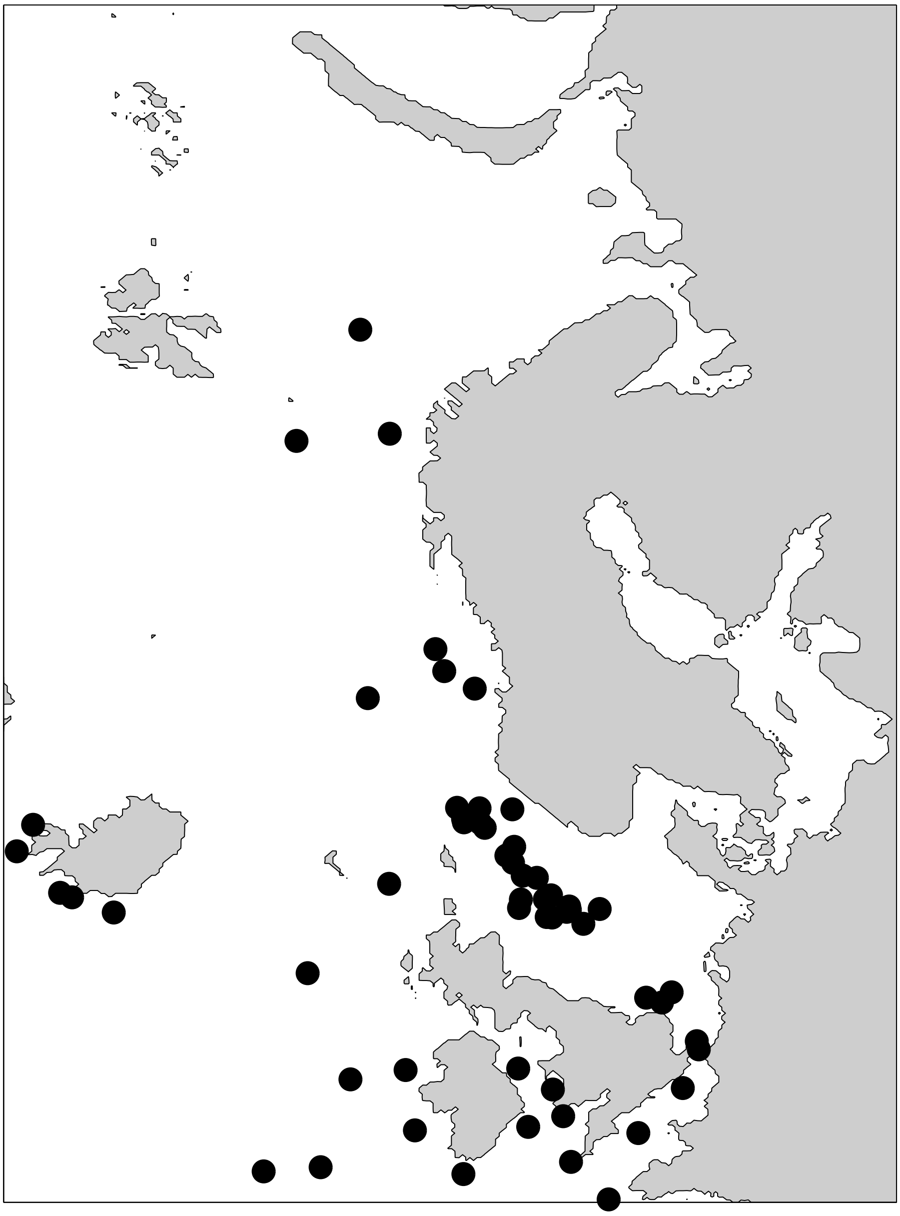

Figure 2: Positions of the buoy and platform observations used in this investigation marked as black triangles.

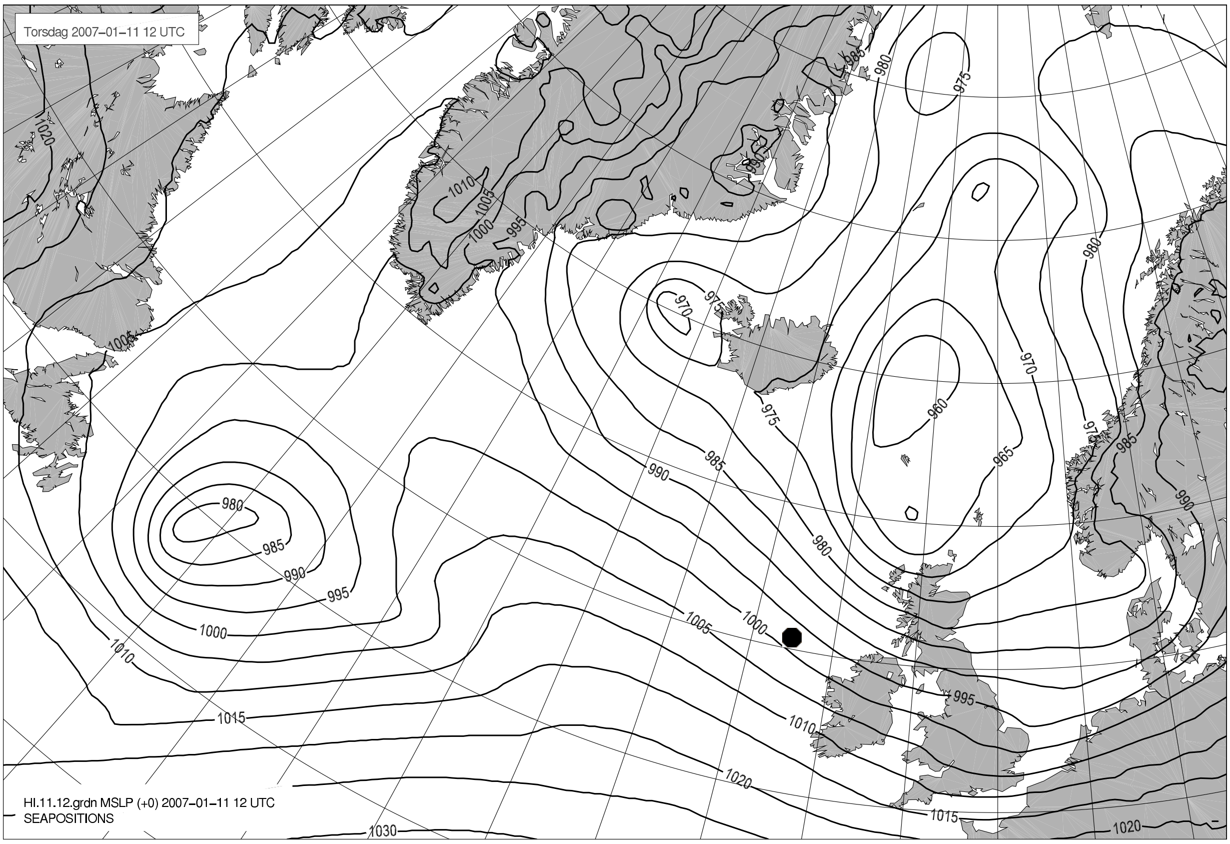

Figure 3: Mean-sea level pressure from met.no's operational analysis for 12 UTC 11 January 2007. The black dot marks the position of the buoy referred to in section 4.

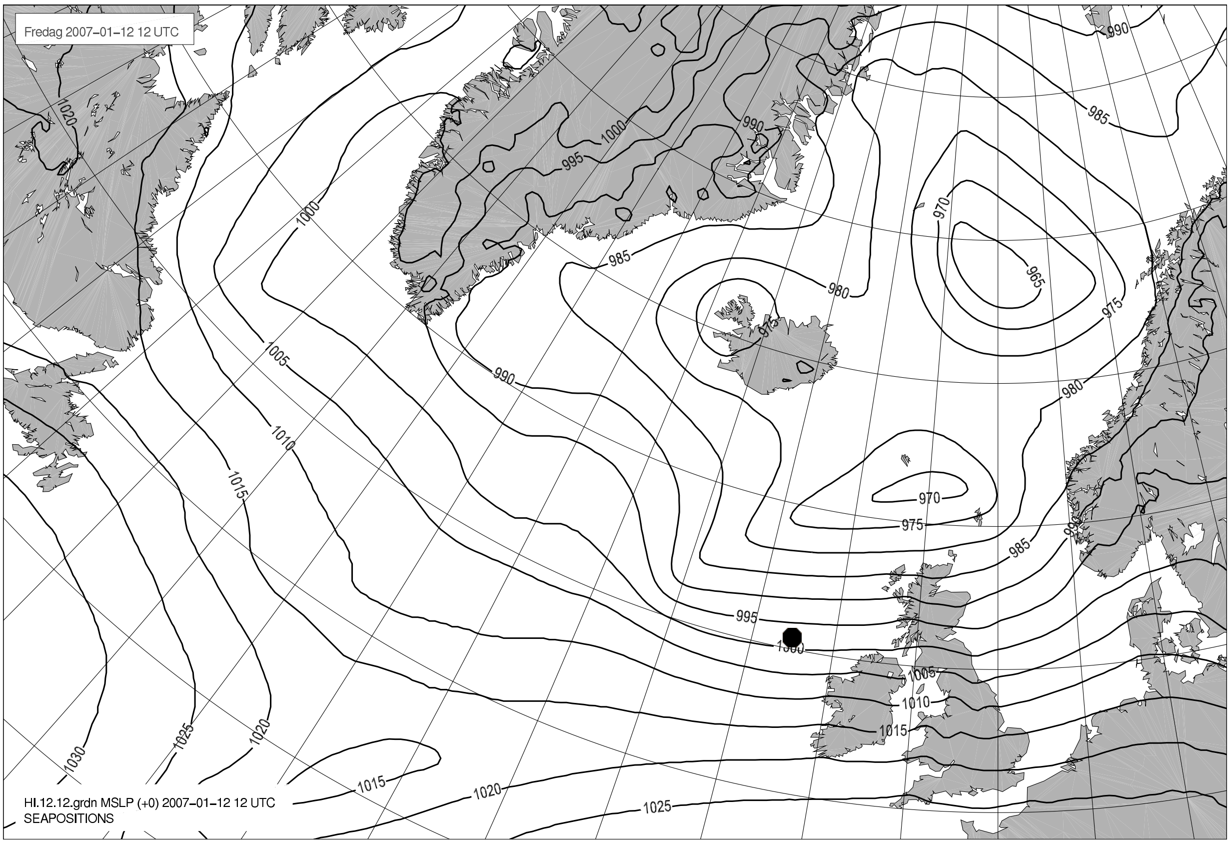

Figure 4: Mean-sea level pressure from met.no's operational analysis for 12 UTC 12 January 2007. The black dot marks the position of the buoy referred to in section 4.

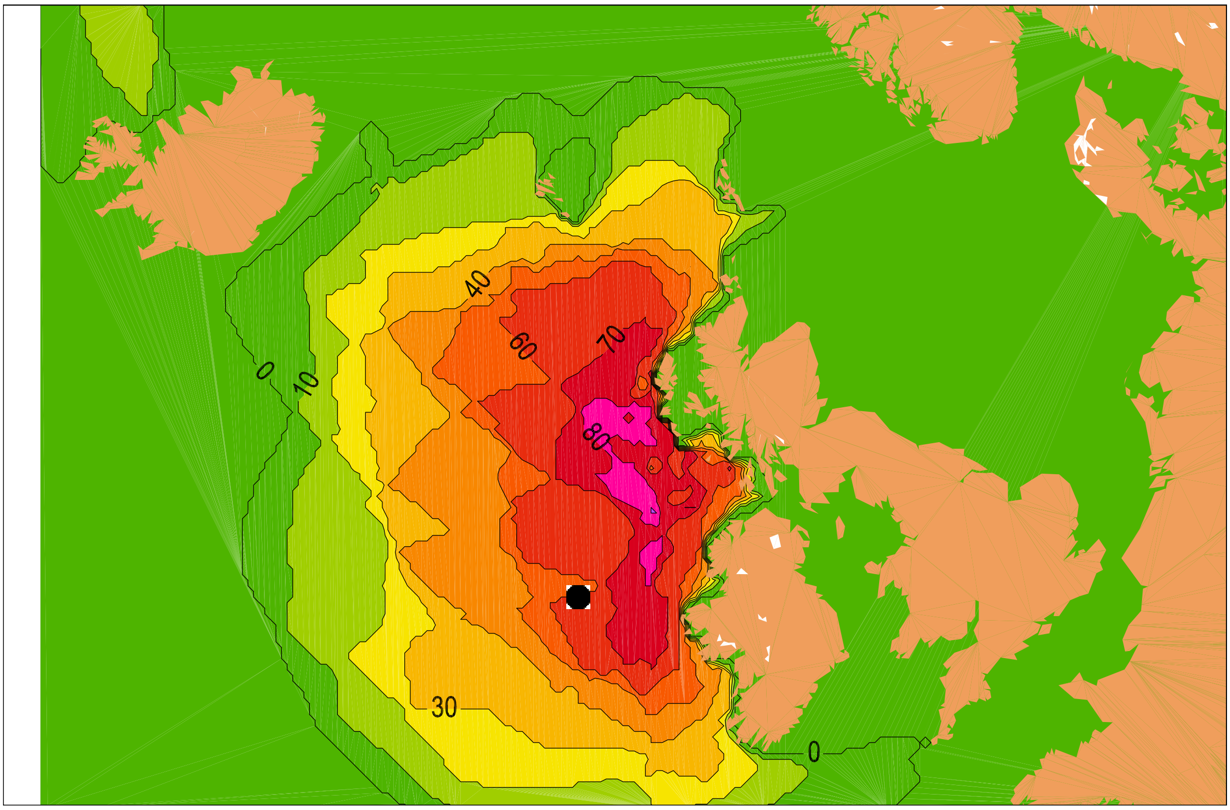

Figure 5: 48-hour forecast of the probability of waves exceeding 10 metre valid on 18 UTC 11 January 2007. The black circles marks the position of the buoy referred to in section 4. Colour scale indicates probability of exceedence. The exceedence probability for the buoy location is about 60% in this case.

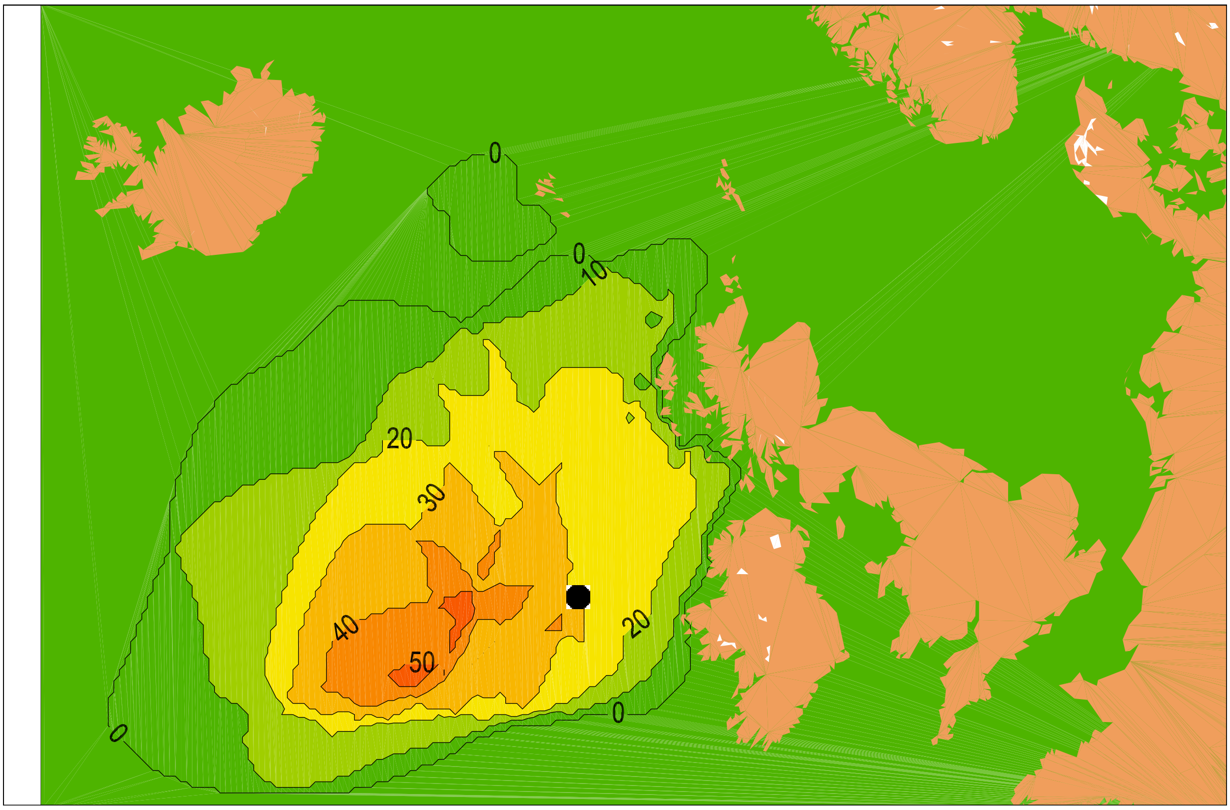

Figure 6: 48-hour forecast of the probability of waves exceeding 10 metre valid on 18 UTC 12 January 2007. The black circles marks the position of the buoy referred to in section 4. Colour scale indicates probability of exceedence. The exceedence probability for the buoy location is about 30% in this case.

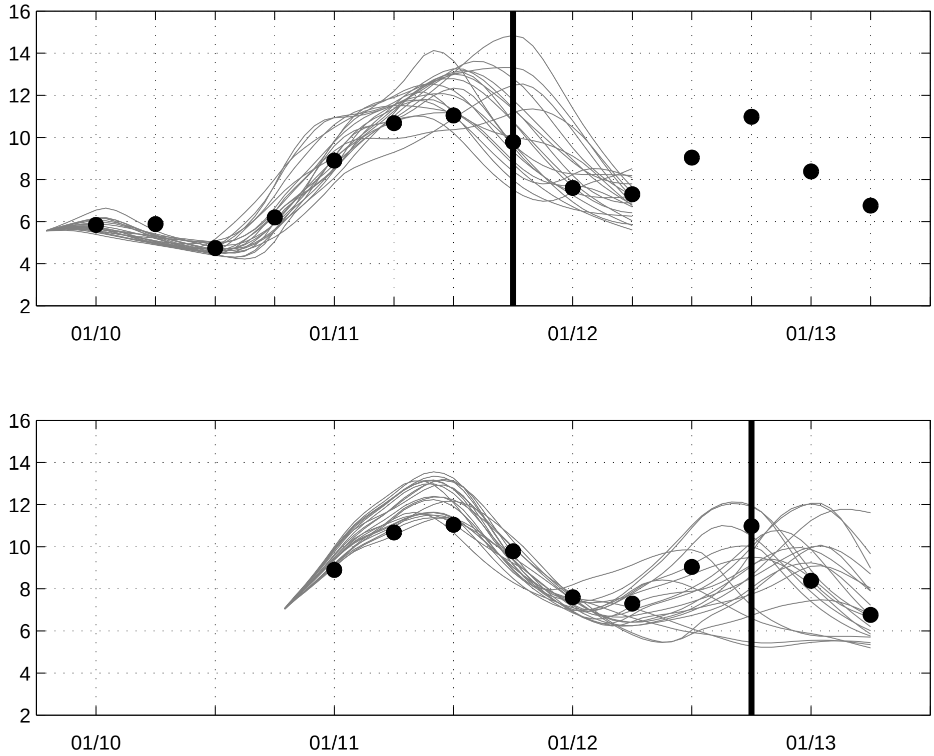

Figure 7: Ensemble forecasts for the buoy location discussed in section 4 for two subsequent days. The thin grey lines are the individual ensemble members and the black dots are the observations from the buoy.

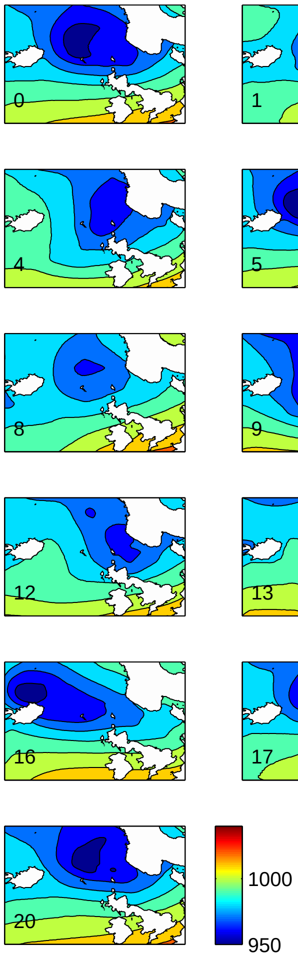

Figure 8: 48-hour forecasts of mean-sea level pressure from all ensemble members valid for 18 UTC 11 January 2007. The ensemble members are numbered from 0 to 20 where 0 refers to the control forecast. The equidistance is 10 hPa.

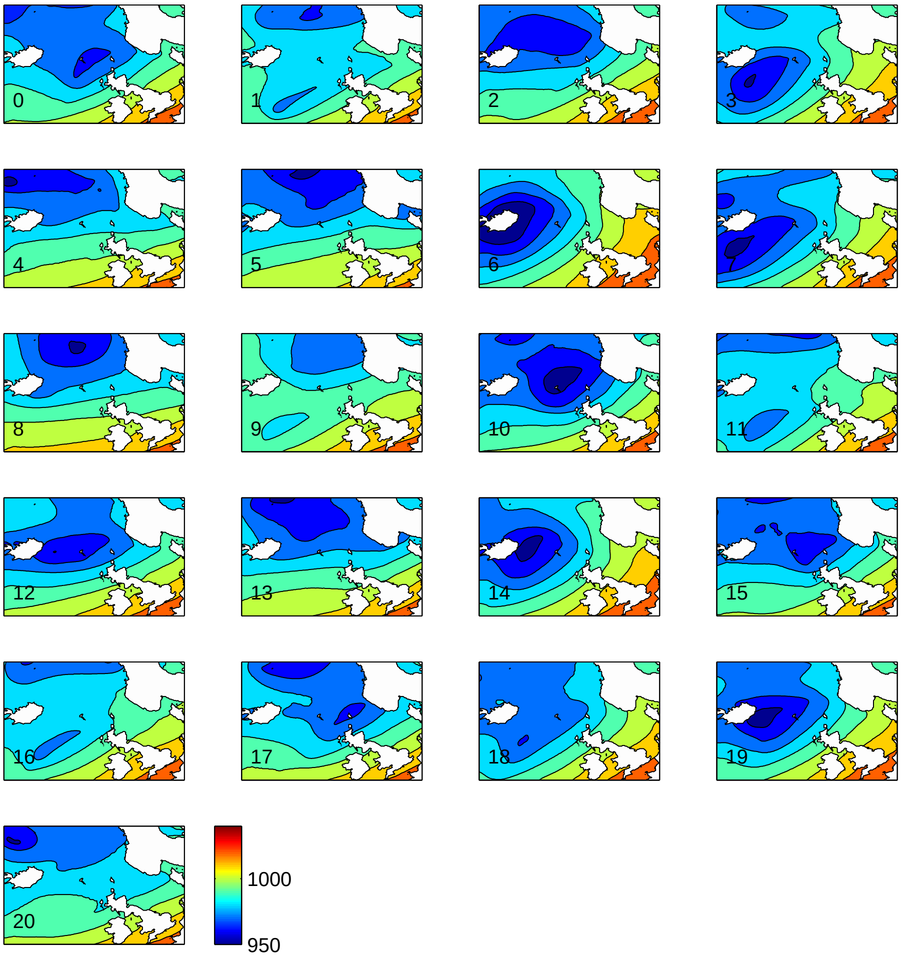

Figure 9: 48-hour forecasts of mean-sea level pressure from all ensemble members valid for 18 UTC 12 January 2007. The ensemble members are numbered from 0 to 20 where 0 refers to the control forecast. The equidistance is 10 hPa.

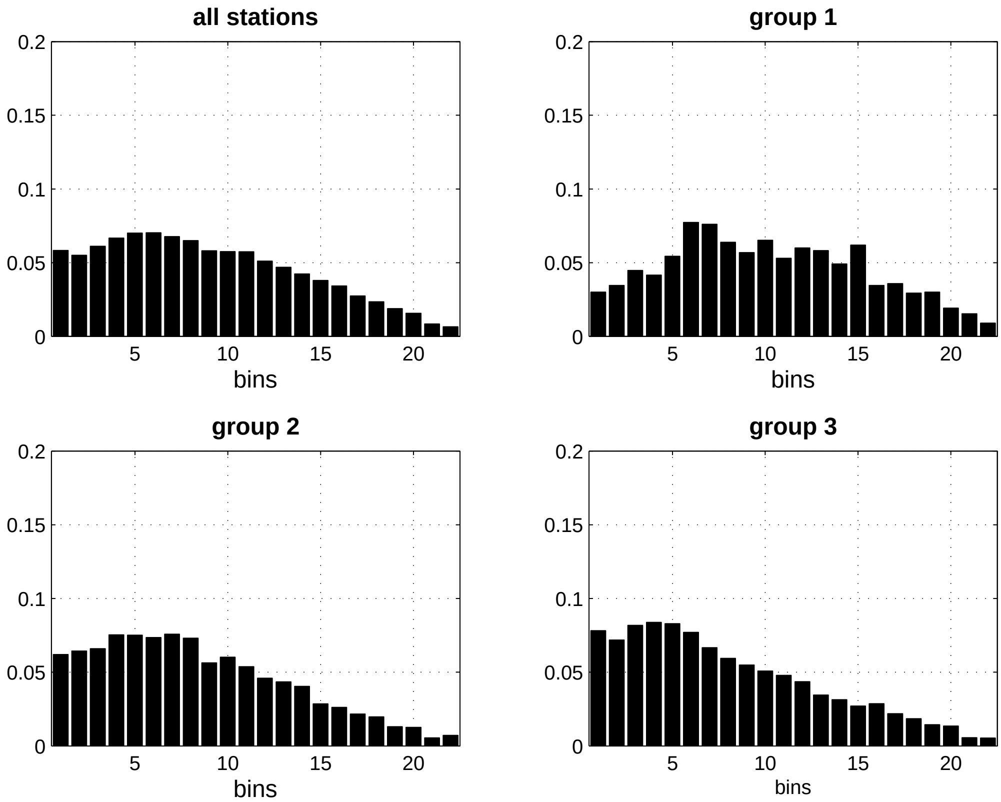

Figure 10: Rank histograms for 48-hour. The upper left plot is when all available stations are used. For the three other plots, the data have been divided into sub-regions with different variability. Group 1 represents the Norwegian Sea north of 65°N, the second group is for buoys located in the North Sea and group 3 are observations from buoys around the British coast.

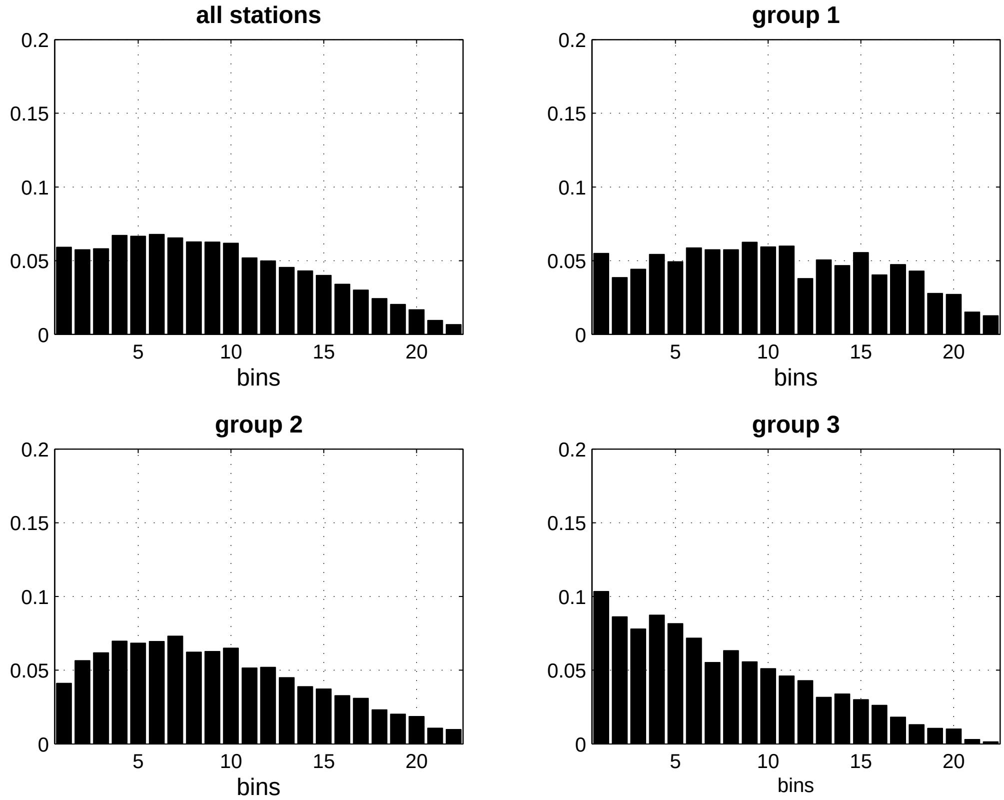

Figure 11: Rank histograms for 60-hour. The upper left plot is when all available stations are used. For the three other plots, the data have been divided into sub-regions with different variability. Group 1 represents the Norwegian Sea north of 65°N, the second group is for buoys located in the North Sea and group 3 are observations from buoys around the British coast.

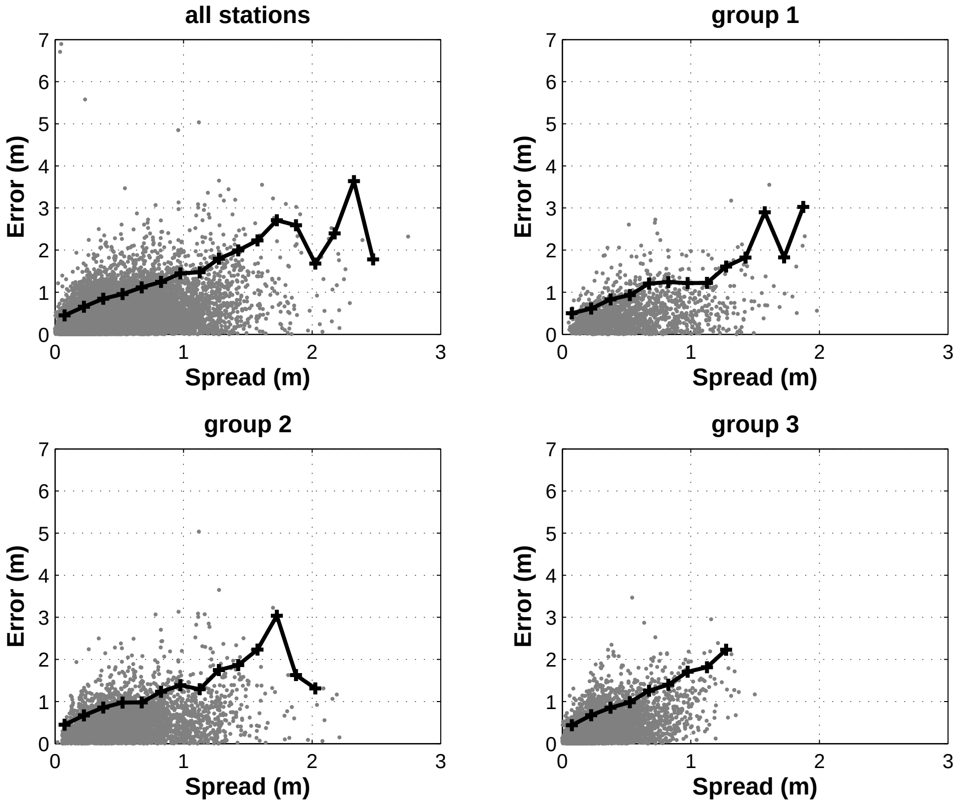

Figure 12: Spread-skill relation for the 48 hour forecasts. The solid line is the 90 percentile of the observations errors for the ensemble mean. The error distribution is depicted as the grey points. The groups are statistics for the three areas explained in Fig. 10.

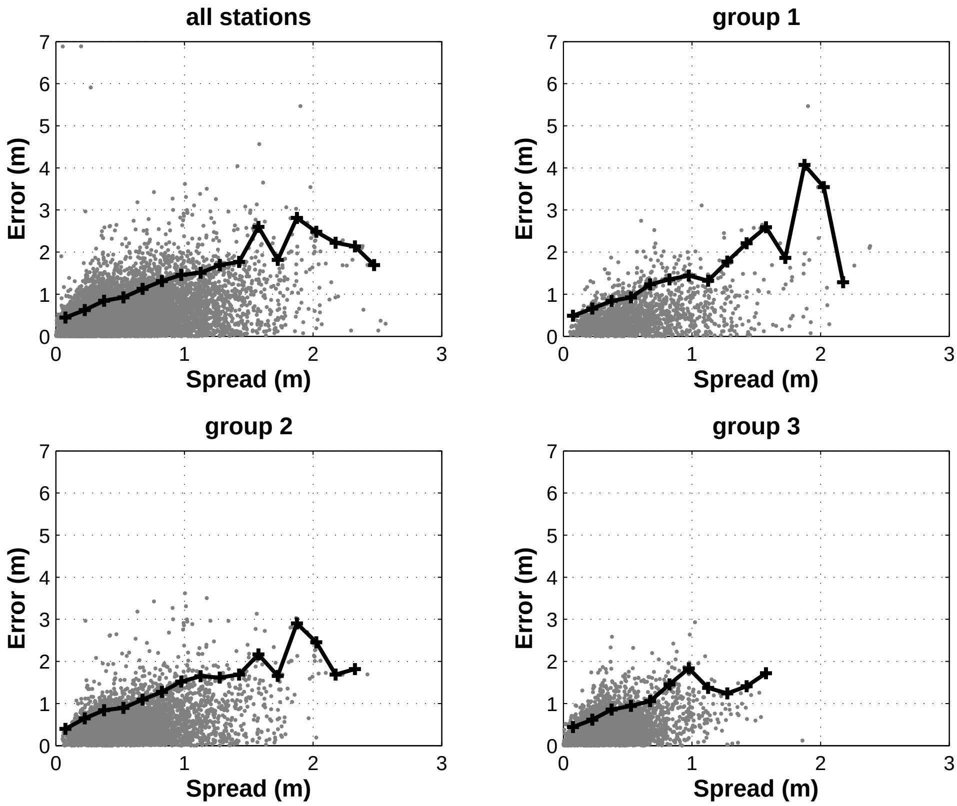

Figure 13: Spread-skill relation for the 60 hour forecasts. The solid line is the 90 percentile of the observations errors for the ensemble mean. The error distribution is depicted as the grey points. The groups are statistics for the three areas explained in Fig. 10.

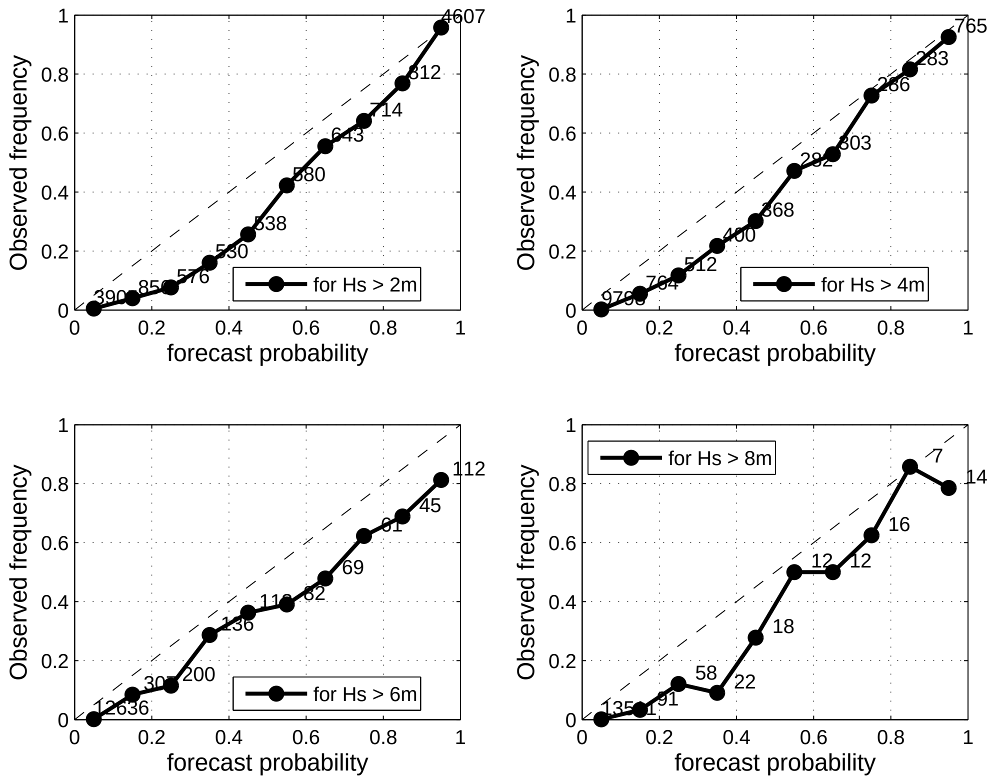

Figure 14: Reliability diagrams for the 48 hour forecasts. The threshold values are 2, 4, 6, and 8 metres.

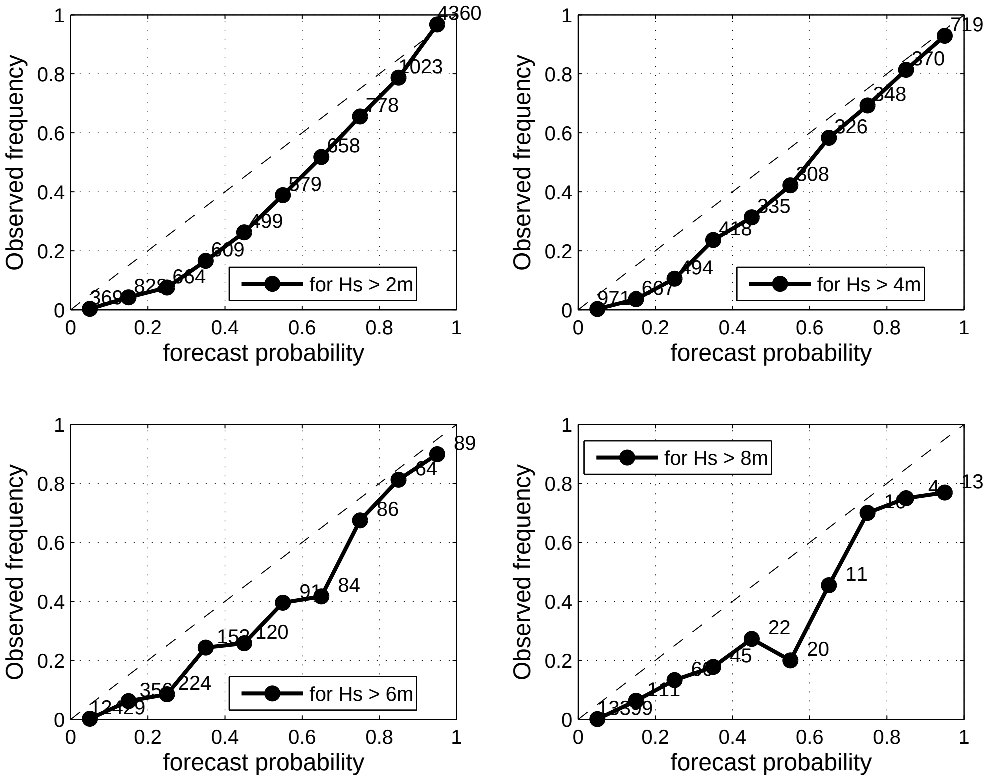

Figure 15: Reliability diagrams for the 60 hour forecasts. The threshold values are 2, 4, 6, and 8 metres.

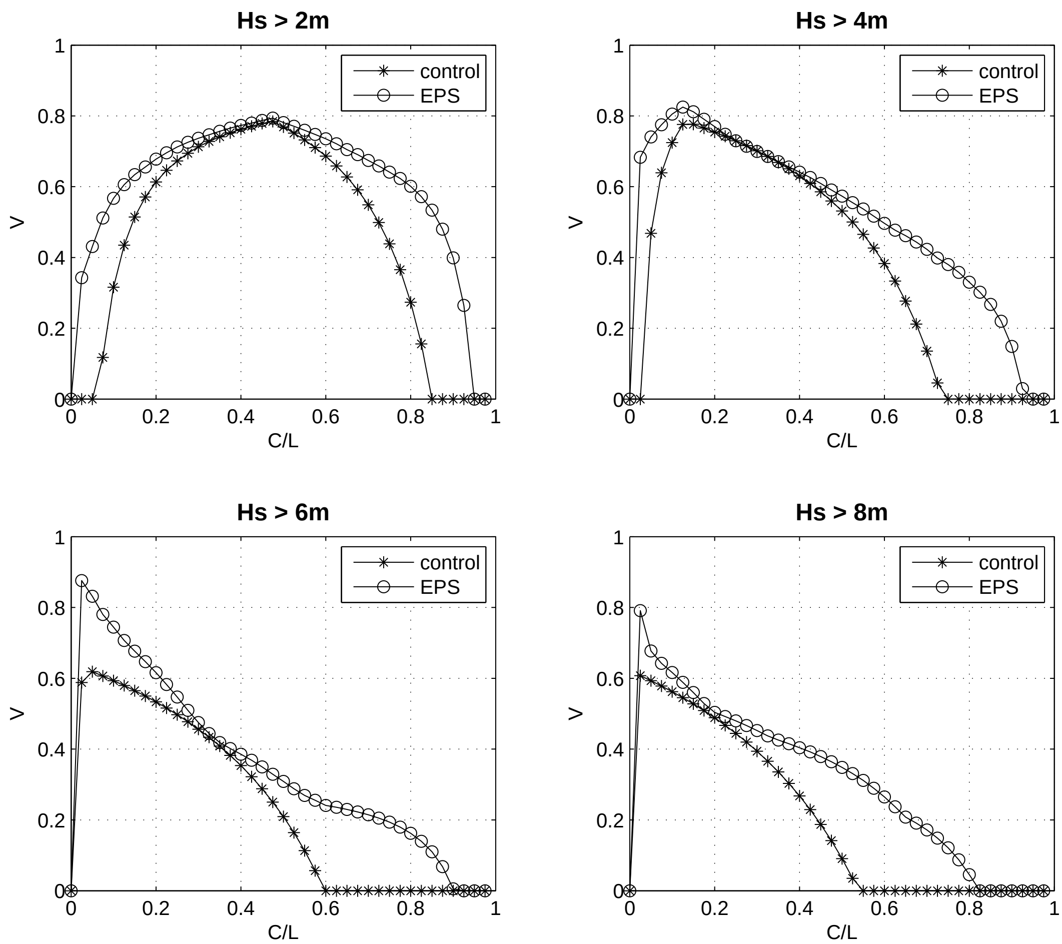

Figure 16: Economic value for the 48 hour forecasts. The line with the open circles are the results for WAMEPS while the results based on the deterministic forecast alone is displayed by the stars. The threshold values are 2, 4, 6, and 8 metres.

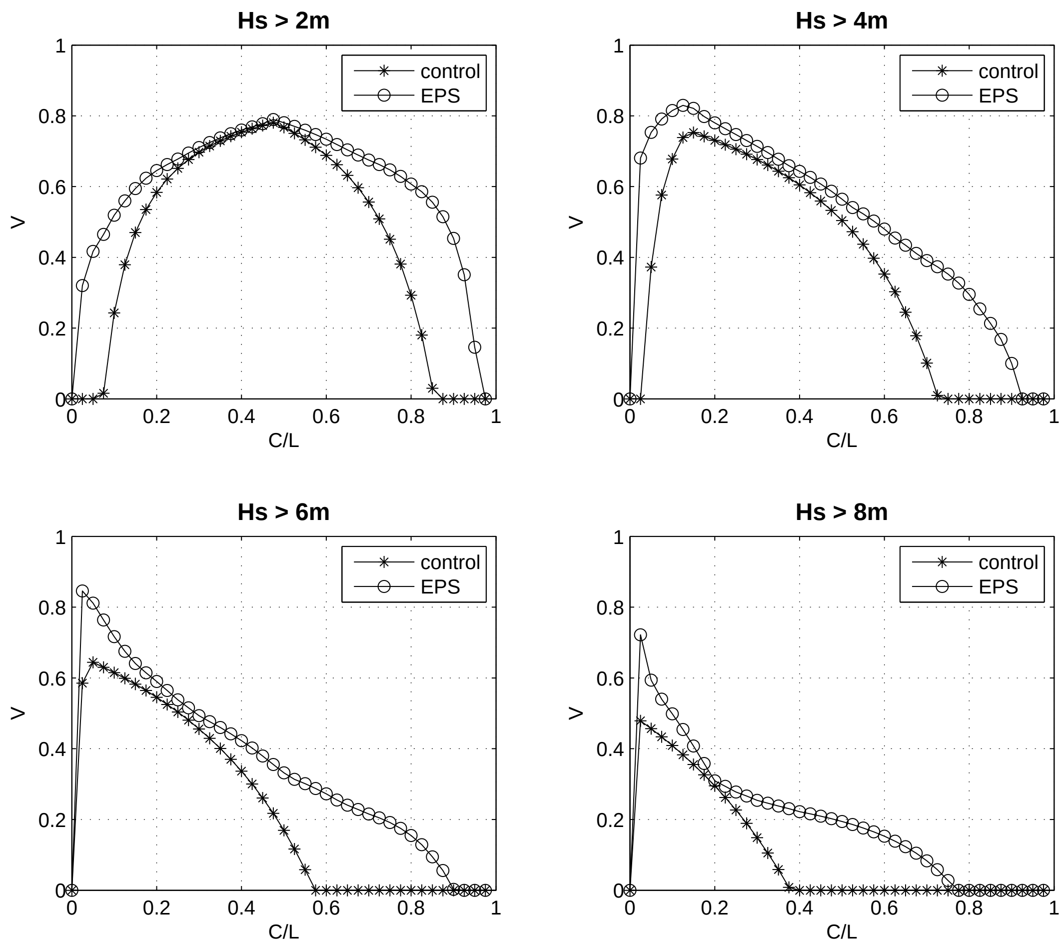

Figure 17: Economic value for the 60 hour forecasts. The line with the open circles are the results for WAMEPS while the results based on the deterministic forecast alone is displayed by the stars. The threshold values are 2, 4, 6, and 8 metres.

Table 1: Table of ROC areas for 48 hour forecasts.

Threshold	8m	6m	4m	2m
Area	0.96	0.97	0.97	0.96
Observed cases	66	382	1791	6519
Total number	13761	13761	13761	13761

Table 2: Table of ROC areas for 60 hour forecasts.

Threshold	8m	6m	4m	2m
Area	0.96	0.97	0.96	0.93
Observed cases	72	413	1863	6530
Total number	13695	13695	13695	13695