TECHNICAL MEMORANDUM

435

Ensemble Shiprouting

O. Saetra

Research Department

April 6, 2004

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

Abstract

Ensemble ship route forecasts have been tested by forcing an optimal route model with wave and wind data from the ECMWF ensemble prediction system. The ensemble forecasts are compared with the results from the optimal routes based on a high resolution deterministic model. The results have been verified against the optimal routes obtained by forcing the ship route model with analysed weather. A period of one and a half years has been considered, where one set of forecasts and analysis per day have been produced . It has been demonstrated that the spread of the ensemble ship routes is related to the spread of the ensemble forecasts for significant wave height. The spread of the ensemble ship routes is clearly related to the skill of the deterministic forecast. This is the case for both the deterministic optimal routes based on the control forecasts and on the high resolution forecasts. As a result, it is suggested to use the ensemble forecasts to produce error bars on the forecasted optimal route, to give the ship master an idea of the uncertainty of the route suggestions.

One of the main objectives of this study has been to investigate if it is possible to find an optimal route from the ensemble forecasts that on average is less expensive than the deterministic high resolution alternative. A study by Hoffschildt et al. (1999) suggests that the ensemble mean is such an alternative. The problem is however that these results are based on a data set where all situations for which the verification run encounters engine damage are discarded. In this paper, the problem of engine damage has been resolved by allowing the ship to vary speed during the verification run. It was found that the average costs of the ensemble mean route is more or less the same as the costs of the high resolution alternative. Both these alternatives are much less expensive on average than any of the others.

1 Introduction

The benefits of any ship routing system, however good the model is, will always depend on the quality of the forecasted weather parameters that are used to force the system. In 1963, Lorenz suggested that there might be a theoretical limit to the atmospheric predictability (Lorenz 1963). By using a simple three dimensional nonlinear model for atmospheric convection, Lorenz demonstrated that the system was highly sensitive to differences in the initial state. Even small differences in the initial condition could after a while result in huge differences in the final forecasts. If this was true also for the atmosphere, small round off-errors in the computer could lead to different weather patterns, making it virtually impossible to forecast the weather, with any reliability, far into the future. This was described as if the flap of a sea gull's wing on one side of the globe could cause a hurricane at the other side of the globe. Later the effect was nicknamed the ”butterfly effect”. Therefore one has to ask for how long it is possible to make a weather forecast that has any value. Usually, when assessing the performance of weather forecast systems, climatology is used as a reference. The predictability is said to be lost when the scores of the forecast system are no better than the those when climate is used to predict the weather. Due to the fact that forecast models are imperfect, it is impossible to answer this question with absolute certainty. Several studies have been made to estimate the growth of the error kinetic energy spectrum (Charney et al., 1966; Leith 1971). Charney et al. suggest that the limit of the deterministic predictability is about two weeks in the winter and somewhat longer in the summer. The study by Leith indicates the limit to be about half of Charney's estimate. However, the weather predictability can not be taken to be constant. The Lorenz instability is flow dependent, i.e. it depends on the current state of the atmosphere. For example, for Northern Europe a blocking situation with stable weather for long periods is assosiated with higher predictability than normal. One method of estimating the flow dependent predictability of the atmosphere that has received a lot attention the last 15 years, is the use of ensemble predictions. This is basically a Monte-Carlo method where several weather scenarios are obtained with slightly different initial conditions. The results are then used to estimate the probability distribution for different weather parameters. Several operational centres around the world are now producing ensemble forecasts on a daily basis. Currently, the European Centre for Medium-Range Weather Forecasts (ECMWF) runs an Ensemble Prediction System (EPS) with 50 perturbed and one unperturbed members (control forecast).

Hoffschildt et al. (1999), later referred to as HO, investigated the potential benefits of using ensemble weather forecasts as forcing for ensemble ship routes. This investigation is a continuation of the work by HO. One of the major findings was that the best performing route, both in terms of fuel consumption and safety, was the ensemble mean route. In more than 50% of the cases, it outperformed both the control and the route based on the high resolution forecast. In this study, the ensemble mean route is the average route over all ensemble members, the control is the EPS forecast initiated from the unperturbed analysis. The high resolution forecast is essentially also based on the unperturbed analysis of the initial atmospheric state, however it uses a higher spatial resolution than the EPS runs. When verifying the ship route forecasts, the cost of all forecast routes is re-calculated using the weather based on the analysis instead of the corresponding forecast weather. The resulting costs are assumed to be the best estimate on how expensive these routes would have been if they had been followed in reality. This approach has one serious problem, which is likely to have large impact on the overall statistics. When the ship is forced along a pre-described route at a pre-described speed, the ship will from time to time run into the centre of a storm without slowing down. In the ship route model, this is heavily penalised through engine damage. However, this is certainly not how a ship captain would operate in a realistic situation. Given a bad ship route forecast, the captain will take preventive action to avoid damaging the ship or jeopardising the safety of the crew. The captain might for instance slow down the ship or change the course of the ship. HO realized this problem, and removed all these cases when computing the statistics. As a consequence, the statistical result will favour the ensemble mean forecast because most of these cases are situations with a storm centred between the departure and the destination. Often in these cases, bifurcations in the ensemble forecasts will occur, with some members to the north and some to the south of the centre of the storm. Hence, the ensemble mean may go through the centre of the storm, usually causing engine damage during the verification. Discarding all these cases results in an unrealistically good performance of the ensemble mean forecasts. One way to overcome this problem is to allow the ship to adjust the speed during the verification run. Technically, this is done by optimising in the time domain only. Namely, the ship is forced to follow pre-described geographical coordinates, but will adjust the speed according to the weather conditions. When the ship approaches the centre of a storm, the ship will slow down to avoid damages. To minimise eventual extra harbour costs due to late arrival, it may then increase the speed once out of the area of bad weather. This is of course not the only option available to a ship captain. He or she may for instance decide to change the course of the ship instead and avoid the storm area all together. This is however difficult to simulate in a realistic way, and a reasonable compromise in a numerical model may be to decide that the ship is allowed to adjust the speed only. At least, it avoids the problems involved in removing all cases when the verifications results in engine damage.

The main questions to be investigated here is whether the ensemble spread of the ship routes could be used to estimate the expected errors in the optimal route forecasts, and which route scenario would be most beneficial for potential end users in terms of the total costs. With an ensemble prediction system in addition to a high resolution deterministic forecast model, the following route options will be compared; the shortest in space (the first guess route), the optimal route from the high resolution model, the ensemble mean route, the maximum density route, the optimal route from the control forecast or a randomly chosen ensemble member. The maximum density route is found by approximating each ensemble member by a Gaussian probability distribution. The full probability distribution is then obtained by adding the distribution from each ensemble member.

2 Ship Route model

The ship route model finds the optimal route between two geographical positions by minimising a cost function in time and space. The cost function calculates costs due to fuel and grease consumption, harbour costs and

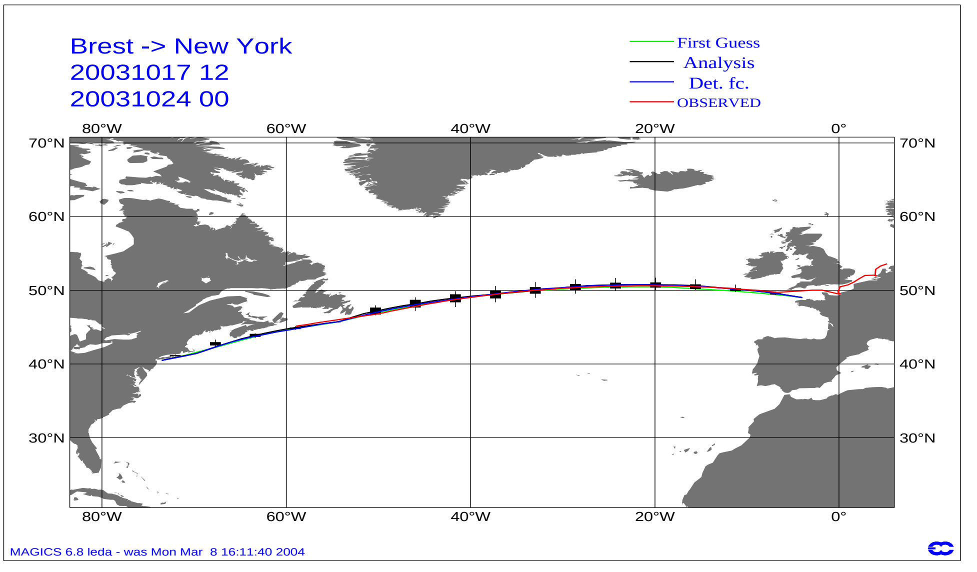

Figure 1: Ship routes for the crossing between Brest and New-York. The plot shows the deterministic forecast (blue line), The first guess (green line), the analysis (black line), the ensemble forecast (error bars). The boxes show the 10 and 90 percentiles for the spread of the ship routes, and the whiskers depict the positions of the outliers. The red line is the route used by the container vessel Leverkusen Express. The forecasts starts at 12Z 14th October 2003.

possible engine damage costs. The present configuration penalises harbour costs relatively hard, such that punctual arrival will almost always be preferred to lower fuel and grease consumption. The fuel and grease cost are calculated as functions of wind speed, wind sea and swell. The present routines are modelling the features of the container vessel Bonn Express, owned by Hapag Lloyd. The route considered is between Brest and New York. To find the optimal route on a given weather scenario, the model starts from a first guess route. This first guess is the optimal route with no wind and waves. In the absence of land, this route would be the great circle, but because for the Brest-New York route this would cross Newfoundland, the first guess is a slightly modified great circle route that follows the coast when approaching the American continent. From the first guess route, the model builds a grid with points along a great circle line normal to the route. To minimise the cost function, the Bellman method is used (Bellman 1957). The advantage of this method is that it is independent of the choice of first guess route, a good first guess will only speed up the convergence of the minimisation. For a more detailed description of the ship route model and the two-dimensional version of the Bellman method, the reader is referred to HO.

The resolution of the atmospheric component of the EPS used to force the ship route model is TL255L40 (spectral triangular truncation with 40 levels in the vertical), which is about 80 km horizontal resolution in grid point space on the Gaussian grid used at ECMWF. A detailed description of the atmospheric component of the ECMWF EPS is given by Buizza et al. (2000). The atmospheric component of the high resolution deterministic model is on TL511L60, which is about 40 km horizontal resolution with 60 layers in the vertical.

The wave model component in the EPS is the ECMWF version of WAM (the WAve Modelling group) cycle 4. WAM was developed during the 1980's by an international group of scientists (Komen et al. 1994). The present version of the EPS wave model runs on a 110 km grid resolution with shallow water physics, 12 directional and 25 frequency bins. The wave component of the deterministic high resolution model is about 55 km with 24

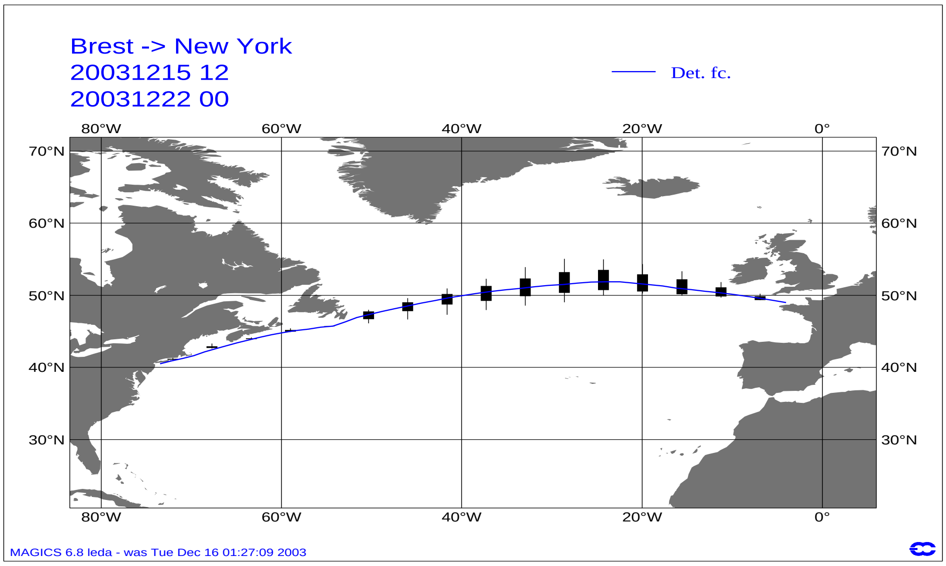

Figure 2: Example of ship route forecast for the route between Brest and New-York in a case where the ensemble spread is relatively large. The plot shows the deterministic forecast (blue line) and the ensemble forecast (error bars). The boxes show the 10 and 90 percentiles for the spread of the ship routes, and the whiskers depict the positions of the outliers. The forecasts starts at 12Z 15th December 2003.

directional and 30 frequency bins.

Between the 14th and the 21th October 2003, a sea trial was carried out with a similar ship, the Leverkusen Express, during its voyage from Hamburg to New York. As part of the SEAROUTES project, a number of sensors were installed on board to measure the ship responses, such as heave, roll and pith motion, during the crossing. The Leverkusen Express left the English Channel and was situated outside Lands End in Cornwall around noon on the 17th October. The positions of the ship is depicted as a red curve in Fig1 together with model runs. The result based on the high resolution wind and wave forecasts is shown as a blue line. This is from a 10-day forecast starting at 12Z 17th October. The green line is the first guess route, i.e. the shortest route in space. The black curve is the model run based on analysed weather. The forecast has been overlaid with error bars based on the ensemble forecast. The solid boxes are the 10 and 90 percentiles for spread of the ship routes, meaning that 80% of the ensemble routes are confined inside the boxes. The whiskers on the error bars show the positions of the outliers. During this crossing, the weather was unusually calm. This is reflected by all model runs being very close to the first guess route, which is also the case for the actual route chosen by the ship master.

A more typical example is shown in Fig.2, where the ensemble spread is much larger. Given that the hypothesis of relationship between ensemble spread and skill of deterministic forecast holds, the error bars would give the ship master an idea of uncertainty of the forecast ship route. If this is true, information on ensemble spread could be a useful tool in ship routing.

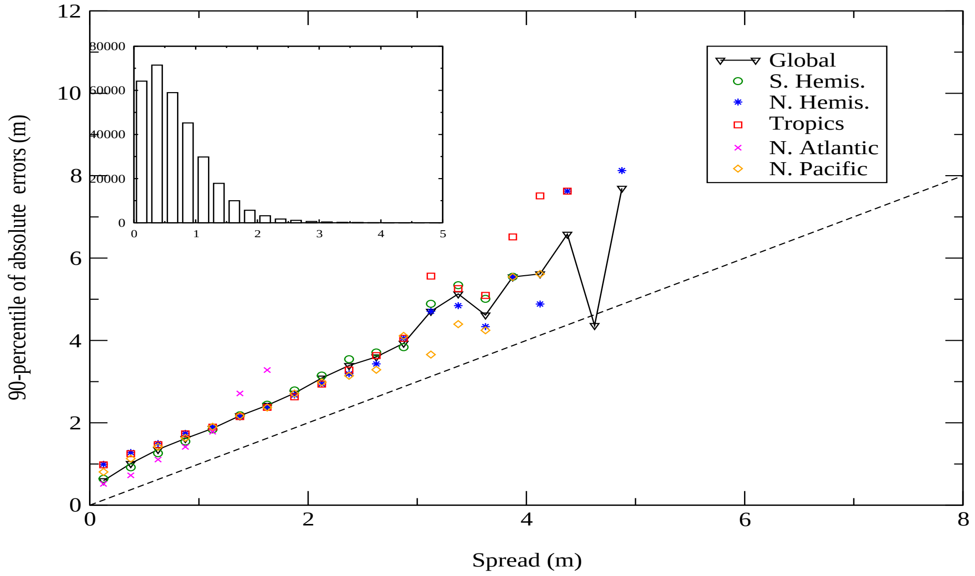

Figure 3: The spread-skill relationship for waves. The plot is based on global altimeter observations. The black solid line is the result when all data are used. The symbols show the result when the data are divided into different areas. For reference, the black dashed line shows the 45 degree diagonal. The bar diagram in the upper left corner gives the number of observations in each bin.

3 Ensemble Spread

Looking at the ensemble spread to estimate the expected model error skill is a convenient way of using the ensemble forecasts. However, it needs to be established that there really is a relationship between the forecast skill and the ensemble spread. Ideally, both ensemble and high resolution forecast and observations are random draws from the same probability distribution. Therefore, the ensemble spread will be a realistic estimate of the errors. As the ensemble forecasts are attempts to estimate the real probability distribution, we need to find a method to establish the relationship between spread and skill. The most important environment parameter for the ship route model is the wave height. Saetra and Bidlot (2004) investigated the performance of the ensemble wave forecasts by verifying against buoy observations and satellite data. In their paper, a method for testing the spread skill correlation was suggested. The ensemble spread was divided into bins, and the individual errors of the wave forecasts were stored within each spread bin. By taking the 90-percentile as an upper bound for the errors in each bin, a curve that is increasing with increasing spread was obtained. Fig3 shows the results for the case when the day 5 forecasts are compared with global altimeter data from satellite. The inlaid histogram shows the number of observations within each bin. The errors are the absolute value of the difference between observations and forecasts. A clear relationship between ensemble spread and errors in the deterministic wave forecasts is demonstrated.

Whether this result is valid also for the ship routing model remains to be demonstrated. If this is the case, there must also be a relationship between the spread of the ensemble ship routes and the spread of the ensemble waves. To produce statistics to validate the ensemble ship routing, the model has been run for a period of one and a half year, from 1 November to 31 March 2002. An ensemble forecast together with a deterministic high resolution forecast was produced for each day during in this period. For verification, an additional route was

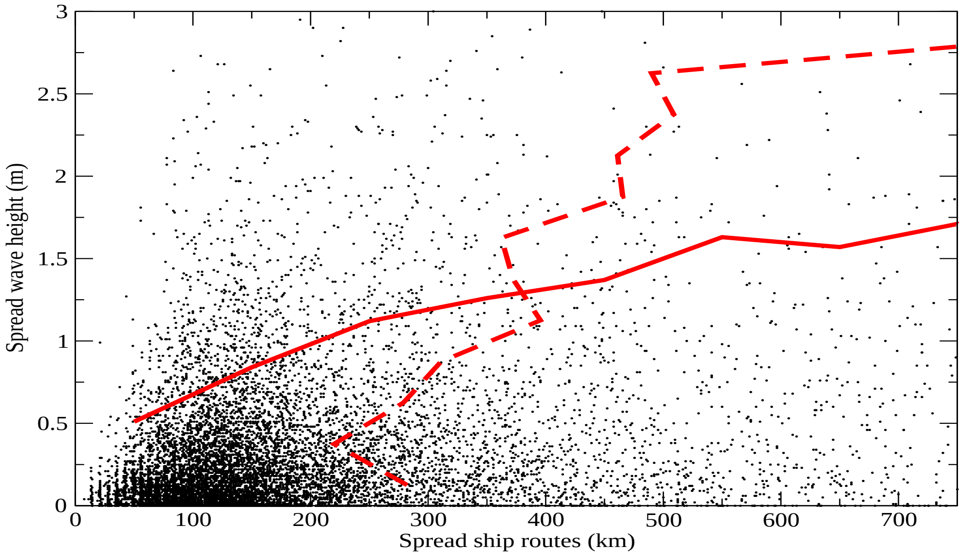

Figure 4: Relationship between the spread of the ensemble waves and the ensemble ship routes. The black dots represent the data pairs. The red solid line is the 90 percentile of the wave spread in the ship spread bins and the red dashed line is the 90 percentile of the ship spread in wave spread bins.

calculated based on the analysed weather for the same period. This route was then taken as the "observation".

In this study, the ship spread is defined as the average distance between the 10 and the 90-percentile over the whole route. The percentiles are found by ranking the ship routes from the most southern to the most northern route. To relate the ship spread to the spread in the ensemble waves, a decision must be made to where the wave spread is calculated. The problem is that each ensemble member for the ship routes is located at different positions, where the ensemble waves might have different spread. To resolve this, it was decided to calculate the ensemble spread for the waves at the position of the ensemble mean route. As was the case for the ship spread, the ensemble spread for the waves is defined as the difference between the 10 and the 90-percentile in this case.

One problem when applying spread of ship routes is that at both ends, departure and arrival points, the ship spread will always be zero. Accordingly, spread is always decreasing when the ship is approaching the arrival harbour, and vice versa when it has recently left the departure port. This makes the data set in the vicinity of the two ports unsuitable for the spread statistics. These areas will introduce a bias in the ensemble statistics towards low spread in the data for the ship positions, whatever the case is for the ensemble spread of the waves. For the statistics regarding ship spread, only "open ocean" areas have therefore been considered. On the eastern side of the ocean basin, all data east of 17.5°W have been discarded. On the western side, one additional problem is caused by the fact that almost all optimal routes tends to be close to Newfoundland, regardless of the weather conditions. The optimal route will only be forced off-shore towards open ocean areas under rather unusual situations. So for most weather conditions, the ship spread will be low in the area along the Newfoundland coast. This will also introduce a bias in the spread statistics, and therfore data west of 50°W have also been discarded.

In Fig.4, the ensemble ship spread and the ensemble wave spread data pairs are displayed as a cloud of points together with the spread-skill relation calculated with the method described above. The solid red line is the

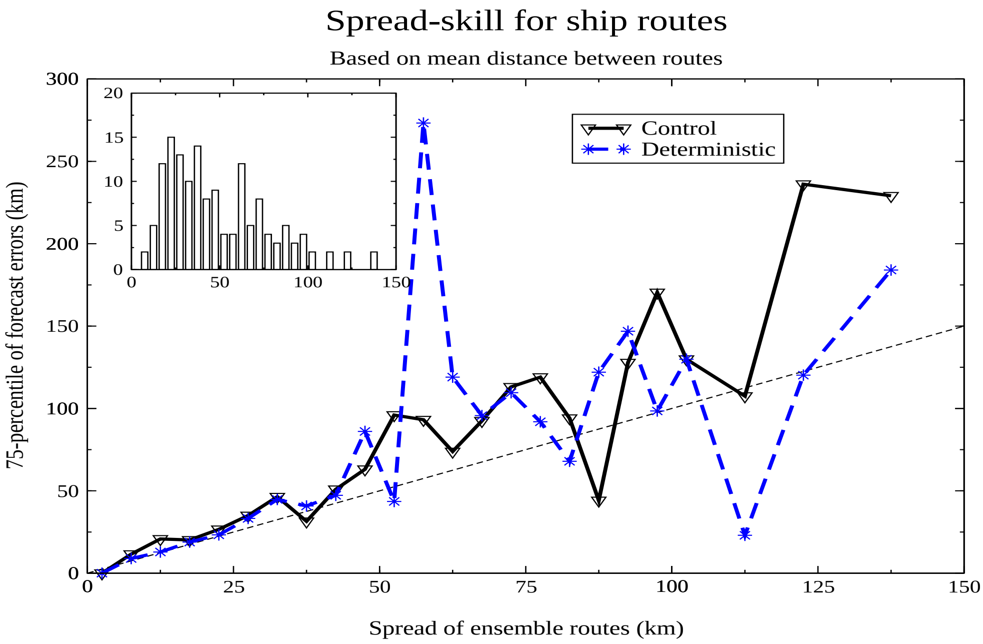

Figure 5: The spread-skill relationship for ship routes. The black solid line is the skill of the control forecast as a function of the ensemble spread of the ship routes. The result for the forecasts based on the high resolution model is shown as a dashed line. For reference, the thin dashed black line gives the 45 degree diagonal line. The histogram gives the number of observations in each bin.

result when the ship spread is divided into bins and the line shows the 90-percentile for the wave height spread. The dashed red line is the opposite case, when the wave spread is divided into bins and the line displays the 90-percentile for the ship spread. Since both lines are increasing with larger spread, there seems to be a relationship between the ensemble spread of significant wave height and the ensemble spread of the ship routes.

In Fig.5, the spread-skill relation for the ship route model is shown. Here, the ensemble spread has been divided into bins. The error distribution for the deterministic forecasts are then used in each bin to determine an upper bound to the errors. The blue dashed line is the result for the high resolution model. For comparison, the thin dashed black line shows the 45 degree diagonal line. The black solid line is the result for the control forecast routes. In this case, the ensemble spread is taken to be the inter-quartile range, i.e. the distance between the 25 and the 75-percentile. The forecast skill values are the 75-percentiles of the error distributions. The bar histogram shows the number of cases in each bin. For both forecasts, high resolution and control, a spread skill relation is present. The curves start zigzagging for ensemble ship spread above 50km, but this is most probably because of too few cases of large ensemble spread.

A relationship between both wave ensemble spread and ship ensemble spread and the relationship between the skill of a deterministic ship routing system and ensemble spread have been showed. Therefore, using the ensemble forecasting system as an estimate of the probable forecast errors seems justified. Although the value of this type of information in economical terms may be difficult to estimate, it seems reasonable to assume that this additional piece of information would be helpful for a ship master. One rather simple way of providing this information, which is easy to explain to a person with no specific statistical training, is to overlay the ship route forecast with error bars as shown in Fig. 1 and 2.

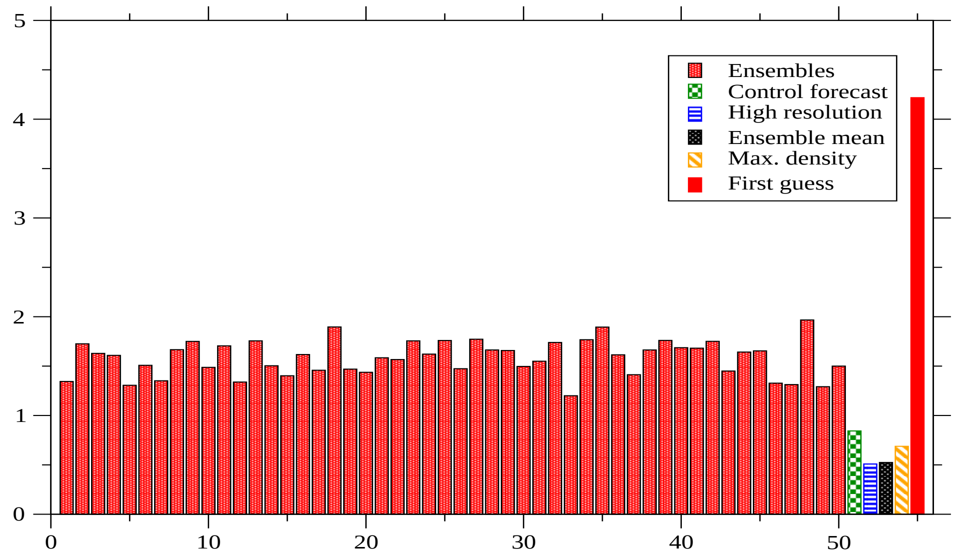

Figure 6: The average costs of different optimal route alternatives relative to the costs of the analysed routes. The vertical axis shows how many percent more expensive a certain alternative is compared to the analysed route. The alternative are explained in the text.

4 Cost Statistics

The question as to which strategy is the most beneficial for a potential user is discussed. Faced with several options, which route should actually be chosen and used as the suggested route for the ship master? To try to answer this question, the cost of different scenarios will be compared over the period studied here. The output of the cost function in the ship route model will be used to accumulate the total cost for each scenario, with one ship starting every day at 12Z over the whole period of the simulation. For comparison, an additional run will be made with the analysed weather for the same starting day. As the outcome of this latter route is by definition the best possible, since it is the minimum solution of the ship route cost function based on analysed weather, it is used to normalise the total costs of all scenarios. As explained in the introduction, all forecast routes are re-run on the analysed weather to estimate the actual cost would have been, assuming that the analysis is the best possible approximation of the actual weather. The geographical positions for the routes were held fixed, while the route was optimised in the time domain. This is to simulate the behaviour of a ship master when the forecasted route may cause damage to the ship or the goods.

In Fig.6, the average cost of choosing the various scenarios relative to the cost of the analysis route are displayed as a bar histogram. The vertical axis shows how many percent more expensive a certain alternative is compared to the analysed route. The most striking feature in this plot is the size of the red bar on the far right side. This shows that the first guess route is on average more than 4% more expensive than the analysis. Any alternative is much better than this, so at least, ship routing pays! The second worst alternative, according to this, is a randomly chosen ensemble member. This is about 1.5% more expensive to run than the analysis route. Note that the control forecast is performing better than any ensemble member. This is a well known feature of the EPS. The control forecast verifies better than randomly chosen perturbed members for most weather parameters (Saetra and Bidlot 2002). The argument is that the analysis is the best approximation to the present weather, and accordingly, it is the initial condition with best chances of correctly predicting the future weather. The control forecast was 0.84% more expensive than the analysis in this study. The alternative with the best performance in this study, is the optimal routes based on the weather forecasts from the high resolution deterministic model. This option is only 0.51% more expensive than the analysis. The ensemble mean route however, is only marginally more expensive with 0.52% obtained here. The final option is to produce a probability distribution from the EPS forecast. This is achieved by assigning a Gausian probability distribution to each ensemble member. Adding the subsequent probability fields for each ensemble member yields the total probability distribution from the EPS. The optimisation program for the ship routes can then be used to obtain the maximum density route. In this study, however the maximum density route was 0.69% more expensive than the analysis.

5 Conclusions

A ship route model for the route between Brest and New York was run on daily ensemble forecasts over a period of one and a half years to produce statistics for spread-skill and total costs. From a previous study by Saetra and Bidlot (2004) it has been demonstrated that the skill of the deterministic forecasts for waves is strongly related to the ensemble spread of the waves, i.e. the ensemble spread can, for this parameter, be used to draw error bars on the time series of the forecasts. In this study, a similar relationship between the ensemble spread of the ship routes and the skill of the optimal routes based on the deterministic forecasts was established. This was achieved by first demonstrating the relation between spread of ensemble waves and the spread of the ensemble ship routes. Secondly, the spread-skill for the ship routes was demonstrated by dividing the spread into bins, and then finding a statistical bound on the error of the deterministic route within each bin. Therefore the spread of the ship route can be used to estimate the probable errors in the forecasts. Consequently, useful information can easily be added to forecast plots by drawing error bars on the predicted optimal routes as in Fig.1 and 2.

The next question to be answered is whether it is possible to obtain routes based on the ensemble forecast that on averaged are cheaper to run, and save the operator more money than a deterministic system does? The results by Hoffschildt et al. (1999) indicates that the ensemble mean route is such an option. However, this result is obtained after all cases are omitted for which the verification runs into engine damage. As argued in the introduction, this is an action that strongly favours the statistics of the ensemble mean because the rejected cases are generally situations with bad weather and possible branching of the ensemble routes. Often, the ensemble mean will go right through bad weather areas in such cases. This assumption has been confirmed here. However, by allowing the ship route model to adjust the ship speed during the verification run, ship engine damage can by avoided for all cases. Consequently, no data had to be discarded. The result is that the cost of the optimal routes from the deterministic forecasts are on average marginally better than the route based on the ensemble mean. The difference in expenses is only 0.01%, so the cost efficiency of these two options must be regarded as equal. All other options considered were more expensive than these two routing alternatives. From this point of view, the answer to the question above must be that the EPS cannot be used to produce a more cost effective alternative to the deterministic route. On the other hand, one might argue that even though the EPS is run on a coarser resolution, it is still able to compete with the high resolution model. The only reason why the resolution is coarser is because of limited computer resources. Ideally, the EPS should be on the same resolution as the deterministic model, i.e. the deterministic model and the control forecast would be one and the same. At the present stage, this can only be taken as a speculation but, if the EPS was run on the resolution of the deterministic model, would the average cost difference between the ensemble mean and the deterministic model be the same as the difference between the ensemble mean and the control is as in Fig6? If this is the case, then the ensemble mean has the potential of becoming the best alternative as judged on cost consideration.

Acknowledgements

This research has been funded by the European Commission through the SEAROUTES project, Contract G3RD-CT-2000-00309. I would also thank Peter Janssen, Jean-Raymond Bidlot, Hans Hersbach and Saleh Abdalla for support and valuable discussions.

References

Bellman, R., 1957: Dynamic programming. Princton Univ. Pr., 240 pp.

Buizza, R., D. S. Richardson and T. N. Palmer, 2002: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Q. J. R. Meteorl. Soc*. **129**, 1269-1288.

Charney, J. G. et al., 1966: The feasibility of a global observation and analysis system. *Bull. Amer. Meteor. Soc.* **47**. 200-220.

Hoffschildt, M., J-R. Bidlot, B. Hansen and P.A.E.M. Janssen, 1999: Potential benefits of ensemble forecasts of ship routing. *ECMWF Research Department Tech. Memo.* No. **287**. 25pp.

Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, Eds., 1994: Dynamics and Modelling of Ocean Waves. *Cambridge University Press,* 533 pp.

Leith, C. H., 1971: Atmospheric Predictability and Two-Dimensional Turbulence. *J. Atmos. Sci.* **28**. 145-161.

Lorenz, E. N., 163: Deterministic Nonperiodic Flow. *J. Atmos. Sci.* **20**, 130-141.

Saetra, O., J-R. Bildot, H. Hersbach and D. Richardson, 2004: Effects of Observation Errors on the Statistics for Ensemble Spread and Reliability. *Mon. Wea. Rev.*, **132**. *in print*.

Saetra, O. and J-R. Bidlot, 2002: Assessment of the ECMWF ensemble prediction system for waves and marine winds. *ECMWF Research Department Tech. Memo.* No. **388**. 29pp.

Saetra, O. and J-R. Bidlot, 2004: On the Potential of Using Probabilistic Forecasts for Waves and Marine Winds Based on the ECMWF Ensemble Prediction System. Accepted for publication in *Weather and Forecasting.*