Storm Prediction Research and its Application to the Oil/Gas Industry

STORM PREDICTION RESEARCH:

*Lizzie S. R. Froude and Robert J. Gurney
Environmental Systems Science Centre (ESSC), University of Reading, UK.

Abstract- The accurate prediction of storms is vital to the oil and gas sector for the management of their operations. An overview of research exploring the prediction of storms by ensemble prediction systems is presented and its application to the oil and gas sector is discussed. The analysis method used requires larger amounts of data storage and computer processing time than other more conventional analysis methods. To overcome these difficulties eScience techniques have been utilised. These techniques potentially have applications to the oil and gas sector to help incorporate environmental data into their information systems.

Keywords: storm tracks, ensemble prediction system, eScience, oil and gas industry.

1. Introduction

Storms are a major natural hazard, causing vast amounts of damage and loss of life around the globe (see below). It is therefore vital that they are predicted accurately by Numerical Weather Prediction (NWP). The quality of the forecasts is improving all the time, because of very large worldwide investments in supercomputing and in satellite technologies for new global observing systems. How can the results of these investments be used profitably by the oil and gas sector?

High quality forecasts of severe weather are essential to the management of oil/gas operations both on and offshore. In 1982 a North Atlantic storm, known as the Ocean Ranger Storm, caused an oil rig located

Book Title Information

near Grand Banks, Newfoundland to capsize and resulted in the tragic death of the entire crew of 84 workers. Hurricane Katrina and other hurricanes in the Gulf of Mexico have led to repeated disruption of the oil and gas industries located there, and similar disruptions are faced regularly by operators elsewhere in the world. Accurate and up to date forecast information delivered in understandable ways to operators is therefore crucial to avoiding disasters and minimising disruption in the future.

Oil and gas consultancy Schlumberger (http://www.slb.com) have recognised the importance of accurate weather forecasts and environmental information in general. Their most recent (2008) annual report reflects the business consequences of these risks and states:

"The prices for oil and natural gas are subject to a variety of additional factors, including: … weather conditions."

"Environmental compliance costs and liabilities could reduce our earnings and cash available for operations."

"Severe weather conditions may affect our operations."

Other oil and gas companies and operators note similar risks to their operations in their financial statement. Schlumberger Information Solutions (SIS, http://www.slb.com/content/services/index_sis.asp) are currently funding storm prediction research of the Environmental Systems Science Centre (ESSC), University of Reading in the UK in order to see if they can develop tools that can tailor the new information obtained from this research to their clients.

The storm prediction research being funded by SIS makes use of the storm identification and tracking software of Hodges (1995, 1999). This software has been used in numerous studies to explore both the current and future climate of extratropical and tropical storms (e.g. Hoskins and Hodges 2002, 2005; Bengtsson et al. 2006, 2007a,b, 2008). More recently the method has been applied to forecast data to explore the prediction of extratropical storms by NWP (Froude et al. 2007a,b; Froude 2009). The approach provides detailed information about the prediction of various properties of storms, such as their position, intensity, growth and propagation speed.

NWP models are integrated from a best estimate of the current atmospheric state, known as an analysis. This is obtained by statistically combining different types of observations (e.g. satellite, weather balloons, aircraft) with a previously obtained numerical forecast by a process known as data assimilation. Since the observations will contain errors and are unevenly distributed around the globe, an analysis will always contain errors. Due to the chaotic nature of the atmosphere (Lorenz 1963) these errors grow rapidly throughout the forecast, making it impossible to accurately predict the weather at higher forecast times.

Recent and current work has focussed on a particular type of forecast system known as an Ensemble Prediction System (EPS), which aims to take this uncertainty into account. A set of multiple forecasts (ensemble members) are run from slightly different initial states. One of the forecasts is known as the control and is started from the analysis and the initial conditions for the other ensemble members are obtained by applying small perturbations to the analysis. Since forecast models themselves are also not perfect sometimes the model is also perturbed during the forecast integration.

The study of Froude et al. (2007b) explored the prediction of storms by the European Centre for Medium Range Weather Forecasts (ECMWF) and National Centers for Environmental Prediction (NCEP) EPS (Buizza and Palmer 1995; Molteni et al. 1996; Buizza et al. 2007; Toth and Kalnay 1993, 1997). It showed that the ECMWF EPS had a slightly higher level of performance than the NCEP EPS and highlighted a number of benefits ensemble prediction offers over single deterministic forecasts in the prediction of storms (such as providing early warnings). As a continuation of this study Froude (2009) explored the regional differences in the prediction of storms by the ECMWF EPS. The results of this study are discussed in section 2 of this paper.

Current work is using the storm tracking approach to analyse and compare 9 EPS from different operational centres. The data for this is being obtained from the THORPEX Interactive Grand Global Ensemble (TIGGE, http://tigge.ecmwf.int/) archive. The first objective of this paper is to present a selection of results from the above discussed work and to describe how this type of information could be useful to the oil and gas sector.

Analysing EPS data using the storm tracking software requires large amounts of data processing and storage. This motivated the use of eScience methodologies, which use distributed computing with distributed data resources to help reduce both the computation time and the storage required to perform the analysis (Froude 2008). The second objective of this paper is to discuss how eScience techniques are being used to help with the research and how they could potentially be used to help integrate environmental information into Schlumberger's information systems.

This paper continues by describing the storm prediction research in section 2 and is followed by a description of the eScience techniques used to perform the research in section 3. The paper finishes with a final summary in section 4.

2. Storm Tracks and Ensemble Prediction

In this section a brief description of how the storm tracking techniques is applied to ensemble forecast data to investigate the prediction of extratropical cyclones by EPS is given. For further details the reader is referred to Froude et al. (2007b). Some results will then be presented to illustrate the type of information that can be obtained from the method and their application to the oil and gas sector are discussed.

2.1. STORM TRACKING

For a given EPS, the extratropical cyclones are identified and tracked along the 6-hourly forecast trajectories of each of the perturbed ensemble members and the control forecasts in both hemispheres using the automated tracking scheme of Hodges (1995, 1999). Before the cyclones are identified the planetary scales with total wavenumber less than or equal to five are removed (Hoskins and Hodges 2002, 2005) so that the cyclones can be identified as extrema without being masked by the larger scale flow. The data are also reduced to a resolution of T42, to ensure that only the synoptic scale features are identified. Vorticity features, at the 850-hPa level, exceeding a magnitude of 1.0×10^{-5} s^{-1} are identified, as positive extrema in the northern hemisphere (NH) and negative extrema in the southern hemisphere (SH), and considered as cyclones. Once the cyclones have been identified the tracking is performed, which involves the minimization of a cost function (Hodges 1999) to obtain smooth trajectories (storm tracks). Only those storm tracks that last at least two days, traveled further than 1000 km and had a majority of their lifecycle in 20°N - 90°N or 20°S - 90°S are retained for the statistical analysis. The tracking is also performed on the analysis data of the same time period to generate analysis storm tracks to use for the verification of the forecast storm tracks.

2.2. EXAMPLE STORM

Fig. 1 shows the tracks and intensities of hurricane Dennis predicted by the European Centre for Medium Range Weather Forecasts (ECMWF) EPS published in Froude (2009). We note that although this storm clearly originates as a tropical cyclone, it spends a majority of its lifetime in the extratropics and is therefore included in the extratropical analysis. There is a large level of uncertainty in the forecast of this storm as can be seen from

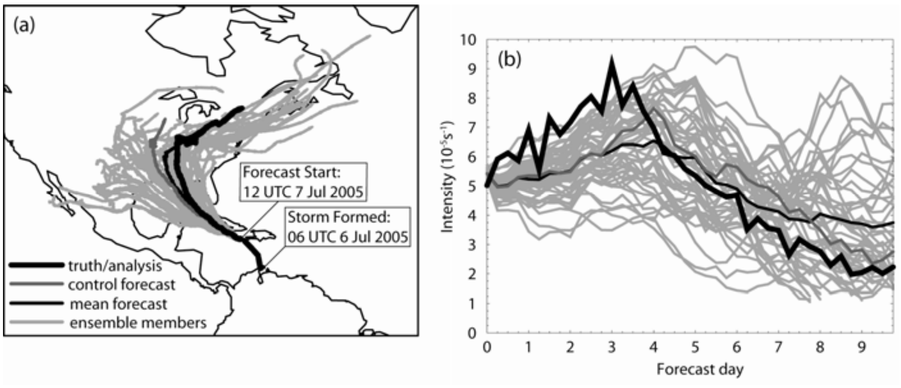

Figure 1. Tracks (a) and intensities (b) of hurricane Dennis predicted by the ECMWF EPS forecast started 1200 UTC 7 July 2005, from Froude (2009). The ECMWF analysed track and intensities are also shown (in black).

the large spread of the ensemble forecast. For the storm's track, some of the ensemble members that lie to the left of the storms analysed track (which is shown in black in fig. 1) travel as far as the west coast of Mexico, whereas ensemble members that lie to the right travel into eastern Canada, with one almost reaching Greenland. The mean forecast storm track from the ensemble provides a reasonably accurate prediction of the storm's track, lying just a little to the left of the truth. It is also more accurate than the control forecast, which lies further to the left. One advantage of an EPS is that in general the mean forecast will provide a better forecast than the single control forecast (Leith 1974; Toth and Kalnay 1993, 1997).

There is also a large spread (and uncertainty) in the predicted intensity of the storm. The growth of the storm is underpredicted by the ensemble forecast, as can be seen from the slope of the mean curve compared with that of the analysis curve between day 0 and 4 in the figure. The control forecast also underpredicts the growth of the storm.

The study of Froude (2009) showed that in general storms that originate in the tropics, such as hurricane Dennis above, are more difficult to predict accurately than those that originate in the extratropics. Dynamically tropical storms are very different to extratropical storms; although they can reach far higher intensities and cause far more damage, they are much smaller in spatial scale. This makes them more difficult to predict without higher resolution models. Information concerning the difficulties and uncertainty in predicting tropical storms could be very important for managing operations in the Gulf of Mexico and similar regions severely affected by tropical storms.

2.3. STORM PREDICTION STATISTICS

In order to generate statistics comparing forecast storm tracks with analysed storm tracks for a large number of storms it was necessary to have an automated objective method of determining which forecasted cyclones correspond to which analysed cyclones. This was achieved using a matching methodology (Froude et al. 2007a,b). For the results presented in this paper, a forecast track was said to match an analysis track (i.e. considered to be the same storm) if the following criteria were satisfied:

(i) At least 60% of their points overlapped in time, i.e.$100 \times [2n_m / (n_A + n_F)] \geq 60\%$ where n_A and n_F denote the total number of points in the analysis and forecast tracks respectively and n_m denotes the number of point in time that occur in both the analysis and forecast tracks.

(ii) The *geodesic* separation distance between the first 4 points in the forecast track, which coincide in time with the analysis track, and the corresponding points in the analysis track must be less than 4°.

The matched forecast tracks can then be used to generate statistics. In the studies of Froude et al. (2007a,b) a variety of different matching criteria were considered. Although the choice of matching criteria had a considerable impact on the number of forecast tracks that match analysis tracks, the statistics generated from the matched tracks were unaffected. In this chapter we therefore only present the results obtained with the above criteria. For further details of the matching method please see Froude et al. (2007b).

2.4. INTENSITY AND PROPAGATION SPEED BIAS

Fig. 2 shows the bias in intensity and propagation speed of storms predicted by the ECMWF ensemble members, in different regions of the extratropics, for the 1 year time period of 6 Jan 2005 – 5 Jan 2006. For the intensity, a positive value corresponds to the storm intensity being overpredicted by the forecasts and vice versa. For the speed a positive value corresponds to the storms moving too quickly in the forecasts and vice versa.

Fig. 2 shows that storm intensity is generally overpredicted by the ensemble forecasts over the oceanic regions (Atlantic, Pacific and Indian) and underpredicted over the land based regions (Eurasia and North America). The predicted speed of the storms is consistently too slow over all regions, but the bias is larger over the Atlantic in the NH. This will affect the predicted time in which a storm will strike in a certain place. In

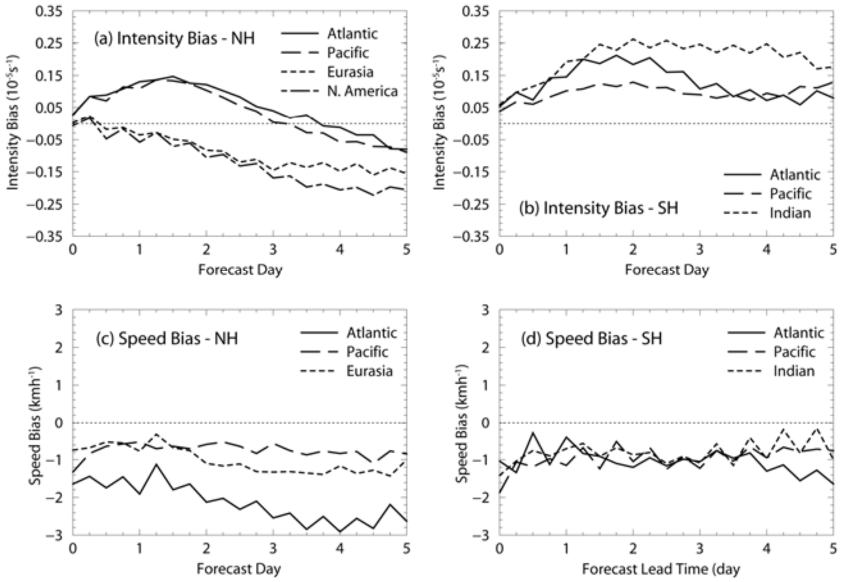

Figure 2. Bias in predicted intensity in (a) NH and (b) SH, and in propagation speed in (c) NH and (d) SH of the ECMWF ensemble members for the period of 6 Jan 2005 – 5 Jan 2006, from Froude (2009), for different regions.

general a storm will strike earlier than predicted by the forecast. It is important that this and other biases of forecast systems are taken into account when deciding what action needs to be taken in response to the prediction of a storm. For example, suppose a very severe storm is heading directly for an oil platform/rig and the decision must be made to stop operations in preparation for when the storm strikes. If storms generally strike earlier than predicted by the forecasts, operations should be stopped some time before the storm is forecast to strike. The amount of time before should depend on the size of the bias. Operations located in regions with larger biases should be stopped earlier than those in regions with smaller biases. Stopping operations is clearly very costly and so this needs to be balanced against the risk of potential damage that could be caused by an oncoming storm.

2.5. DIFFERENT ENSEMBLE PREDICTION SYSTEMS

Ensemble Prediction first became operational in 1992 at both ECMWF and NCEP (Buizza and Palmer 1995; Molteni et al. 1996; Toth and Kalnay 1993, 1997). Nowadays a large number of the major operational weather centres run an EPS. As mentioned in the introduction, data from 9 different

EPSs, obtained from the TIGGE archive, are being analysed with the storm tracking approach. Fig. 3 shows the ensemble mean error in storm position and intensity for each of the different EPS for the 4 month period of 1 Feb 2008 – 31 May 2008. The ensemble mean error is calculated by determining the mean track/intensity from the ensemble member storm tracks and then calculating the error between this and the corresponding analysis storm track/intensity (see Froude et al. 2007b for further details). All the EPSs were verified against the ECMWF analyses so the results for ECMWF may be subject to some positive bias in the earlier part of the forecast. In the future we hope to be able to perform the verification against analyses from the other centres, but we currently only have access to the ECMWF analyses at the 6 hourly frequency required for the storm tracking.

Fig. 3 shows that there are large differences between the different EPS. In general EPS with smaller/larger errors in position have smaller/larger errors in intensity. However, NCEP has a larger error in intensity in relation to the other EPS than it does for position. Further analysis of the TIGGE data (not shown) indicates a lot of differences in the ability of the different EPS to predict storms. For example the ECMWF EPS has a small bias to overpredict the intensity of the cyclones, whereas the UKMO has a bias to underpredict them. Like the ECMWF EPS, all of the EPSs underpredict cyclone propagation speed, although some centres have larger biases than others.

This type of information would potentially be useful to the oil and gas industry when deciding which forecast system (or systems) to use for decision making concerning operations. For example if an intense storm is expected to strike in a region where oil/gas operations are being carried out, it may be more important to know the timing of the storm to high accuracy than the exact intensity when deciding when to shut down operations. In this case it would be better to use a forecast system with higher skill in predicting the propagation speed of the storms than the intensity. If several different EPSs are used, then the question of how best to combine the information they provide in order to get the best estimate of position, intensity, timing etc. would need to be addressed.

3. Storm Tracks and eScience

As discussed in the introduction, the storm track analysis of EPS requires large amounts of data processing and storage. In this section some of the eScience techniques used to help overcome these difficulties are discussed.

Froude (2008) developed a web application to allow users to run the storm tracking software directly from a web browser with remote datasets and using distributed computing. Users are currently able to compute storm

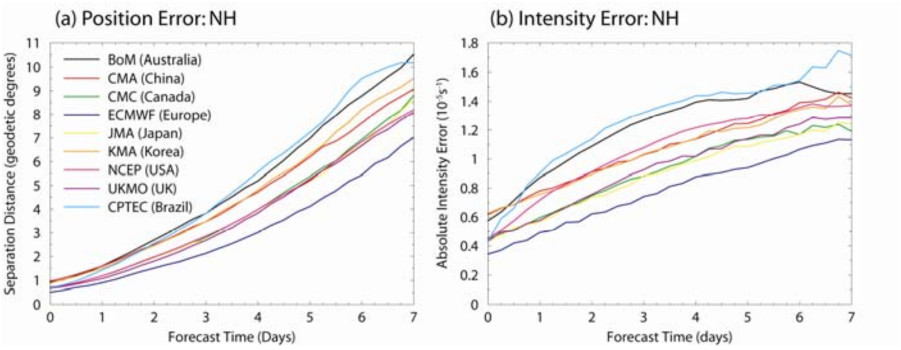

Figure 3. Ensemble mean error in (a) position and (b) intensity of storms in the NH extratropics for the period of 1 Feb 2008 – 31 May 2008 for the different EPS.

tracks from the NCEP re-analysis (Kalnay et al. 1996) and NCEP EPS (Toth and Kalnay 1993, 1997) datasets. A list of jobs can be constructed and executed across multiple computers to reduce computation time. The progress of each job can be monitored and once completed; the computed storm tracks can be downloaded and plotted in a web browser. Froude et al. (2007b) made use of the web application to help reduce the amount of data storage required and to help reduce the computation time involved in the processing of the NCEP EPS data.

The web application was written using Java Servlets/Java Server Pages (Hall 1999). It accesses the remote data using the Open-source Project for a Network Data Access Protocol (OPeNDAP, http://www.opendap.org). OPeNDAP allows data to be accessed over the Internet. OPeNDAP also has a sub-sampling facility, so that a specific part of a data file can be requested. This allows the user to download just the parts of the data they require, rather that downloading the entire data file. The storm tracking program was modified to work with OPeNDAP. It can now be used to compute storm tracks from remote datasets. The OPeNDAP sub-sampling facility is used to request specific meteorological fields and time periods requested by the user in their job list. This use of sub-setting dramatically reduced the amount of data that are needed to be stored locally. For example, the NCEP EPS data files include a large number of meteorological fields at a large number of different pressure levels. For the storm-tracking analysis only mean sea level pressure or vorticity at the 850 mb level were required. These fields are selected with the sub-sampling facility rather than downloading the entire file.

The web application allows users to submit a list of jobs to the Condor (Thain et al. 2005) pool in ESSC. Condor is a software system that manages a collection of jobs by making use of the computational power of machines over a network. Users can submit a list of multiple jobs to Condor, which chooses where and when to run them. Each job in the user's job list is submitted as a separate job to the Condor pool and is run on a different machine. This allows a much faster throughput than using just a single machine.

The NCEP re-analysis data at the Climatic Data Center (CDC) is stored in yearly files (January–December). The OPeNDAP sub-sampling facility allows users to select a time period within a given year (i.e. the same file). It is not, however, possible to select a period that begins in 1 year and ends in another (e.g. a December–February season) because the data for such a period is split across two file. To overcome this problem, the OPeNDAP aggregation server (http://www.opendap.org/server/agg-html/agg.html) was used. This is a piece of software that can be used to create aggregated datasets by effectively merging individual files so they appear as one large file. The NCEP re-analysis dataset was aggregated so that it could be treated as one large 50 year file rather than 50 smaller 1 year files. The aggregation of the data means that the user is able to run the storm tracking software with NCEP re-analysis data from any time period between 1943 and the present.

Figure 4 shows a flow chart, from Froude (2008), illustrating how the different components of the web application fit together. The user constructs a list of jobs in their web browser (labelled [1]), which is then sent to the server (labelled [2]). The server then submits this list of jobs to the Condor pool in ESSC (labelled [3]). Condor puts the jobs into a queue and then sends the jobs to different computers (labelled [4]) as and when they become available. The program accesses the data using the OPeNDAP protocol. The NCEP EPS data are accessed directly from the OPeNDAP server (labelled [6]), whereas the re-analysis data are accessed via the aggregation server at ESSC (labelled [5]). Once all the jobs have finished running the output from the storm tracking program (labelled [7]) is put onto the server for the user to download or plot. While a set of jobs are running, the user is able to check the progress of each individual job from their web browser. For further details of the web application, please see Froude (2008).

The work discussed in section 2.4, analysing the TIGGE data sets requires even larger amounts of data storage and processing than the work of Froude et al. (2007b), which made use of the web application. These data are not available via OPeNDAP and so it is not possible to access the data remotely. However, in order to reduce the computation time of this data processing the University of Reading Campus Grid was used. This consists of a Condor pool of approximately 150 Linux machines and therefore dramatically speeds up the data processing. The importance of Condor to

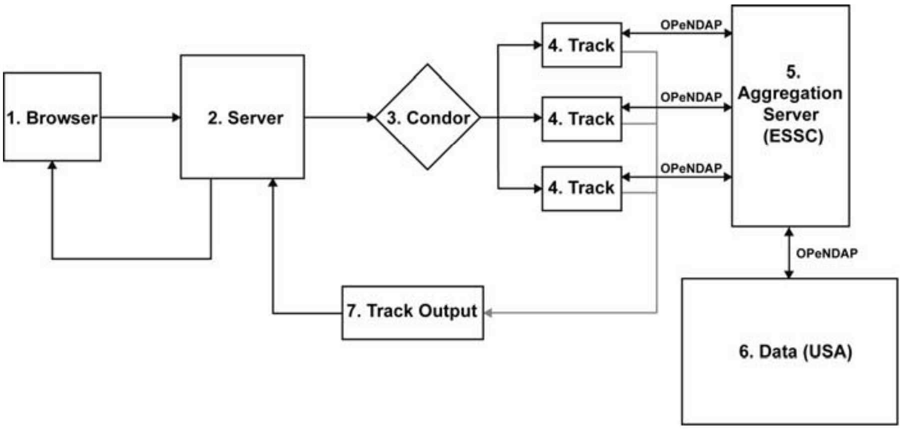

Figure 4. Flow chart to illustrate how storm tracking web application works, adapted from Froude (2008).

this work cannot be overemphasized. Without it, the data processing would have been extremely difficult with the facilities available. Condor was particularly well suited to the processing of the EPS data, since the storm tracking for each ensemble member could be performed on a different computer.

In summary, the storm tracking analysis of EPS data sets requires very large samples of data. Without the use of eScience methodologies, it would not have been possible to store and analyse such large amounts of data. We are currently working with Schlumberger to explore how storm prediction information (and other environmental information in general) can be incorporated into their information systems. This information is potentially very valuable to the management of operations both on and offshore. It is anticipated that the use of eScience will help with this task considerably.

The eScience technologies discussed in this section could also potentially be useful in other areas of scientific research. The TRACK web application has been used by scientists from the US Navy to study past storms using the NCEP re-analysis data. eScience methodologies, such as those used by the web application, could be useful for operational NWP. For example, distributed computing techniques such as Condor would be ideal for ensemble prediction.

4. Summary

The oil and gas sector, particularly the offshore production platforms, need accurate information on extreme events that may disrupt operations. EPS allow better estimation of extreme events, but analysis shows that there can still be biases in intensities and timing predicted by these systems. Handling the large amount of data in such EPS is also a challenge. While weather forecasts are already essential to safe offshore operations, there is a need for still better analysis of predictions and display of results if the oil and gas sector is to make optimal use of the new environmental information now available.

5. Acknowledgements

The authors would like to thank Schlumberger and NERC for funding the research discussed in this article. They also wish to thank ECMWF and NCEP for providing the data.

References

Bengtsson, L., Hodges, K. I. and Roeckner, E.(2006) Storm tracks and climate change. Climate Dynamics. 19, 3518-3543.

Bengtsson, L., Hodges, K. I. and Esch, M. (2007a) Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analyses. Tellus. 59A, 396-416.

Bengtsson, L., Hodges, K. I., Esch, M., Keenlyside, N., Kornblueh, L., Luo, J.-J. and Yamagata, T. (2007b) How may tropical cyclones change in a warmer climate? Tellus 59A, 539-561.

Bengtsson, L., Hodges, K. I., and Keenlyside, N. (2008) Will extratropical storms intensity in a warmer climate? J. Climate, accepted.

Buizza, R. and Palmer, T. N. (1995) The singular-vector structure of the atmospheric global circulation. J. Atmos. Sci. 52, 1434-1456.

Buizza, R., Bidlot, J.-R., Wedi, N., Fuentes, M., Hamrud, M., Holt, G. and Vitart, F. (2007) The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). Q. J. Roy. Meteorol. Soc., 133, 681-695.

Froude, L. S. R., Bengtsson, L. and Hodges, K. I. (2007a) The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. Mon. Wea. Rev. 135, 315-333.

Froude, L. S. R., Bengtsson, L. and Hodges, K. I. (2007b) The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. Mon. Wea. Rev. 135, 2545-2567.

Froude, L. S. R., (2008) Storm tracking with remote data and distributed computing. Computers and Geosciences. 34, 1621-1630.

Froude, L. S. R., (2009) Regional differences in the prediction of extratropical cyclones by the ECMWF ensemble prediction system. Mon. Wea. Rev. accepted.

Hall, M. (1999) Servlets and Java Server Pages (JSP) 1.0: A tutorial. See http://www.apl.jhu.edu/~hall/java/Servlet-Tutorial/.

Hodges, K. I. (1995) Feature tracking on the unit sphere. Mon. Wea. Rev. 123, 3458-3465.

Hodges, K. I. (1999) Adaptive constraints for feature tracking. Mon. Wea. Rev. 127, 1362-1373.

Hoskins, B. J. and Hodges, K. I. (2002) New perspectives on northern hemisphere winter storm tracks. J. Atmos. Sci. 59, 1041-1061.

Hoskins, B. J. and Hodges, K. I. (2005) A new perspective on southern hemisphere storm tracks. J. Climate. 18, 4108-4129.

Kalnay, E., Kanamitsu, R. and Kistler, R. (1996) The NCEP/NCAR 40-year re-analysis project. Bull. Am. Meteorol. Soc. 74, 2317-2330.

Leith, C. E. (1974) Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, 102, 409-418.

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

Molteni, F., Buizza, R., Palmer, T. N. and Petroliagis, T. (1996) The ECMWF ensemble prediction system: Methodology and validation. Quart. J. Roy. Meteor. Soc. 122, 73-119.

Thain, D., Tannenbaum, T. and Livny, M. (2005) Distributed computing in practice: The Condor experience. Concurr. Comput. Pract. Exp. 17, 323-356.

Toth, Z. and Kalnay, E. (1993) Ensemble forecasting at the NMC: the generation of perturbations. Bull. Am. Meteorol. Soc. 74, 2317-2330.

Toth, Z. and Kalnay, E. (1997) Ensemble forecasting at NCEP and the breeding method. Mon. Wea. Rev. 125, 3297-3319.