The Predictability of Extratropical Cyclones

Lizzie S. R. Froude

Environmental Systems Science Centre (ESSC), University of Reading, Harry Pitt Building, Whiteknights, PO Box 238, Reading, RG6 6AL, United Kingdom
e.s.froude@reading.ac.uk

ABSTRACT

Extratropical cyclones (storms) are fundamental to the weather in the mid-latitudes and it is vital that they are predicted as accurately and as far in advance as possible by numerical weather prediction (NWP). In the past studies of the prediction of extratropical cyclones have mainly focused on individual cyclones or cyclone simulations. There have been some statistical studies, but these have used manual or semi-automated methods to identify and track the cyclones. As a result these studies have been limited due to the large amount of work involved. This paper presents a review of previous cyclone predictability studies and then describes a fully automated storm tracking forecast verification methodology. An overview of some results that have been obtained from its implementation are presented and discussed. Results analysing the prediction of storms by different ensemble prediction systems (EPS) are presented followed by some regional analysis of the ECMWF EPS.

1 Introduction

The day-to-day weather in the midlatitudes is largely dependent on the presence or absence of extratropical cyclones. In the presence of these cyclones, weather conditions are generally unsettled, stormy, wet and windy; in their absence, the weather is more settled and dry. Extratropical cyclones can be beneficial, in that they provide the majority of the precipitation received in the midlatitudes and are therefore important for human activities such as agriculture. They can also be very damaging, since under certain conditions they can intensify more than usual, bringing very heavy rainfall and extremely strong winds. This can result in loss of life and economic damage and it is therefore important that these cyclones are predicted as accurately and far in advance as possible by numerical weather prediction (NWP).

In the past studies of the prediction of extratropical cyclones have mainly focused on individual cyclones or cyclone simulations. There have been some statistical studies, but these have used manual or semi-automated methods to identify and track the cyclones. As a result these studies have been limited due to the large amount of work involved. In recent years a new cyclone identification and tracking approach to forecast verification has been developed (Froude et al., 2007a,b; Froude, 2009, 2010a,b). This approach provides detailed information about the prediction of cyclones. The method involves the identification and tracking (Hodges, 1995, 1999) of extratropical cyclones along forecast trajectories. Statistics can then be produced to determine the rate at which the forecast cyclones diverge from the analysed cyclones with increasing forecast time. Detailed information about the prediction of extratropical cyclones can be determined that it is not possible to obtain from other conventional forecast verification methodologies.

This paper has 2 main aims. The first is to provide an overview of previous cyclone predictability studies and the second is to describe the storm tracking forecast verification methodology and give an overview of some of the results that have been obtained from its implementation. This paper continues with a review of previous cyclone predictability studies in section 2 and the storm tracking methodology is described in section 3. Section 4 presents some recent results from Froude (2010a,b) which analyse different ensemble prediction systems (EPS) archived as part of the THORPEX Interactive Grand Global

Ensemble (TIGGE) project (Bougeault and Coauthors, 2010). Section 5 presents some regional analysis of the ECMWF EPS from Froude (2009) and the paper finishes with some final remarks and discussion of future direction in section 6.

2 Previous Extratropical Cyclone Predictability Studies

There have been numerous studies of the prediction of individual extratropical cyclones. These are generally motivated by the severity of a particular cyclone or by the failure to accurately forecast such an event. Examples include that of Morris and Gadd (1988) for the Great October storm of 1987 and that of Pearce et al. (2001) for the European storms of 1999. Studies of such cyclones are not limited to operational forecasts of the time; current numerical models are also used to study the prediction of severe cyclones that occurred in the past. For example Jung et al. (2004) used a recent version of the European Centre for Medium Range Forecasts (ECMWF) to reanalyse and reforecast three major European storms of the twentieth century, including the October 1987 storm. They found that although the prediction of the track and intensity of the storm was very good with this modern high resolution model, the timing of the storm was difficult to predict. As a continuation of this study Jung et al. (2005) explored the prediction of the storms by the ECMWF ensemble prediction system (EPS). The study showed that the EPS was able to predict the large forecast uncertainty associated with the timing of the October 1987 storm as much as 4 days in advance.

There have also been a number studies, which have investigated the impact that some controllable factor has on the prediction of an individual cyclone or cyclone simulation. A common factor that is often studied is the use of specific types of observations. Examples of such studies include Kuo et al. (1997), Xiao et al. (2002) and Pouponneau et al. (1999). The Kuo et al. (1997) study looked at the impact that Global Positioning System (GPS) refractivity data had on the short range prediction by the Penn State/NCAR mesoscale model (MM5, Grell et al., 1994) of an extreme cyclone, which occurred over the Northwest Atlantic in January 1989. Results of the study showed that assimilation of the refractivity data significantly improved the temperature and moisture fields and led to a considerably more accurate prediction of the cyclone. The Xiao et al. (2002) study investigated the impact that satellite derived winds had on the prediction, also by the MM5 model, of a mid-Pacific cyclone that occurred in February 1998. They found that the satellite wind observations increased the cyclonic zonal wind shear and cross-front temperature gradient associated with the cyclone and consequently improved the predicted position and intensity of the cyclone.

Pouponneau et al. (1999) looked at the impact that upper-level wind aircraft data has on the analyses and forecasts of a well-predicted Atlantic cyclone occurring in February 1994. The study used the operational data assimilation and forecasting system of Meteo-France (Courtier and Geleyn, 1988) and made use of an automated cyclone tracking system (Baehr et al., 1999, and references therein) to track relative vorticity maxima. The study showed that the inclusion of upper-level aircraft data modified the vertical structure of the forecasted storm, which led to significant forecast differences. Pouponneau et al. (1999) also suggest the use of an automated tracking algorithm to provide an alternative measure of forecast skill to those measures currently used.

The impact that targeted observations have on the prediction of individual cyclones has also been explored. For example Leutbecher et al. (2002) evaluate the potential to improve forecasts of one of the French storms of December 1999 and a storm that hit Denmark also in December 1999 by using supplementary observations in regions that lack accurate observations. The study used the ECMWF Integrated Forecast System (IFS) and optimal observing regions (where the use of additional observations will reduce the forecast error the most) were identified with singular vectors. Overall the additional observations were found to improve the forecasts of the cyclones.

Another factor that is often explored in such impact studies is the initial state. For example Zou et al. (1998) investigated the impact of uncertainties in the initial conditions of a 5-day forecast of the cyclogenesis of the Atlantic cyclone studied in Kuo et al. (1997). Using the MM5 model they found that forecasts made up to 4.5 days before the storm reached its peak predicted an intense cyclone, whereas the forecast made 5 days before did not. Using a simplified version of the MM5 adjoint model they determined optimal perturbations by minimizing the errors occurring in the initial 12-hours of the 5 day forecast. These perturbations were then introduced into the original analysis from which the 5-day forecast was integrated and the resulting forecast was substantially improved.

Another study which used an adjoint model to study the initial condition sensitivity of forecasts of an extratropical cyclone is that of Langland et al. (2002). They use the U.S. Navy global forecast model to study a U.S. east coast cyclone of January 2000. They found that introducing optimal perturbations into the initial state of the 3 day forecast decreased the error in the predicted position of the cyclone from 1860 km to 105 km. Studies such as this and the Zou et al. (1998) study illustrate the sensitivity of the prediction of cyclone development to the initial state.

The sensitivity of the development of individual cyclones to the initial state is also often explored by running an ensemble of simulations. Examples of this include the studies of Sanders et al. (2000) and Hacker et al. (2003). The Sanders et al. (2000) study constructed an ensemble using the National Center for Atmospheric Research (NCAR) CCM2 model (Hack et al., 1993) to explore the sensitivity of explosive cyclogenesis to the initial conditions. Two different cases of cyclogenesis were contrasted: one near Kamchatka and the other in the central Pacific. The cyclogenesis of the Kamchatka case was found to be more predictable than that of the central Pacific case. This was attributed to the weaker upper level predecessor trough of the central Pacific case and to it's smaller horizontal scale.

Hacker et al. (2003) studied a cyclone occurring off the northwest coast of North America in February 1999, which was badly predicted by NWP models. The storm was forecast to hit Vancouver and warnings of severe snow, rain and winds were issued. However, the storm did not hit Vancouver until much later than forecast and in a decayed state. The Hacker et al. (2003) study uses ensembles to study both initial condition and model error, and concluded that model error played a larger role in the poor forecasts.

In a recent study Zhu and Thorpe (2006) investigated forecast error growth, due to errors in the initial conditions and model deficiencies, by following the development of an extratropical cyclone in a simulation obtained by applying upper level potential vorticity (PV) perturbations to an idealized two-dimensional baroclinic jet initial state. Primitive equation models with different vertical discretization and horizontal resolution were used to explore the impacts of model uncertainty. Upper level perturbations of different amplitudes were used to explore the contribution of initial condition uncertainty to forecast error growth. Differences between the forecast error growth arising from inaccurate initial conditions and model deficiencies are discussed.

All of the studies discussed so far have focused on individual cyclones or cyclone simulations. Although such studies can provide a lot of information about the prediction of extratropical cyclones, a statistical analysis of a large number of cyclones is required to obtain an objective assessment of cyclone prediction and predictability. Such statistical studies are considerably less numerous than those of individual cyclones. This has perhaps been mainly due to the large computational requirements involved in such an analysis.

There have in the past, however, been a number of studies which have aimed to provide a statistical evaluation of the prediction of extratropical cyclones by operational models over North America and the adjacent oceans. The first such study is that of Leary (1971), which analysed the prediction of a sample of 417 storms from the November 1969 - February 1970 winter by the then operational model of the National Meteorological Center (NMC, now the National Centers for Environmental Prediction). In this study the cyclones were manually identified and tracked, from analysis and 36-hour forecast surface pressure maps, as those systems that had one or more closed isobar. The forecast cyclones were then compared with the analysed cyclones and statistics for the errors in the predicted pressure, thickness and position were generated. The results showed forecasted cyclones over the ocean did not deepen enough, those to the lee of the Rockies were too deep and too warm and that for strongly deepening cyclones the forecast tracks generally lie to the right of the analysed tracks.

Other studies using this method of manually identifying and tracking cyclones to generate forecast statistics followed. Silberberg and Bosart (1982) analysed forecasts of cyclones by the NMC Limited Area Fine MeSH Model (LFM) and obtained very similar results to Leary (1971). Grumm and Siebers (1989) and Grumm et al. (1992) examined forecasts of cyclones by the NMC Nested Grid Model (NGM) and found that the forecasted cyclones deepened too much over continental areas and did not deepen enough over the ocean. They also found that some of the forecasted cyclones had a tendency to move too slowly and that there was a cold bias during the winter months. These studies used a semiautomated method to identify and track the sea level pressure features, which meant that unlike the earlier studies of Leary (1971) and Silberberg and Bosart (1982) the data from surface pressure maps did not have to be manually entered into a computer.

Other such studies include those of Grumm and Siebers (1990), Grumm (1993) and Smith and Mullen (1993) which compared the prediction of cyclones by the NMC's NGM and the aviation run of the global spectral model (AVN) and found that the AVN had higher levels of predictive skill in both cyclone intensity and position. Sanders (1986, 1987) and Sanders and Auciello (1989) investigated the prediction of explosive cyclogenesis by models from the NMC. The studies showed that the NMC models were to slow to develop systems that were rapidly intensifying. Sanders (1992) evaluated the prediction of cyclones in the central and western North Atlantic region by the operational models of NMC, ECMWF and the UK Meteorological Office. The NMC model was found to have slightly higher levels of skill than the other models, but the models were verified against NMC analyses and the results may therefore contain some bias. Differences in the observations available in the central and western North Atlantic region to each of the forecasting systems may also have played a role.

The statistical studies discussed so far have mainly been concerned with the short range prediction (2 days or less) of cyclones and have been focused on the North America region. In the past there have also been a few studies that have considered the medium range prediction of cyclones over Europe. The studies of Girard and Jarraud (1982) and Akyildiz (1985) investigated the differences between the grid point model then operational at ECMWF and a spectral model. They found that the propagation speed of cyclones was consistently too slow for the grid point model and was also too slow for fast moving cyclones in the spectral model. The cyclone deepening and filling rates were found to be less than that observed for the grid point model, but were found to be more realistic for the spectral model.

All of these statistical studies of cyclone prediction have involved manually identifying and tracking surface features. Although some useful information about the prediction of cyclones has been obtained, these studies have been limited by the time consuming task of manually identifying and tracking the features. It should also be noted that all of the statistical studies discussed are some 15 or more years old. A statistical analysis of the prediction of extratropical cyclones by current NWP is therefore required. The work highlighted in this paper performs a statistical analysis of the prediction of extratropical cyclones by NWP using a fully automated method of cyclone identification and tracking.

3 Storm Tracking Methodology

In this section the storm tracking methodology is described before the presentation of some results obtained with the method in sections 4 and 5. The extratropical cyclones are identified and tracked along the 6-hourly forecast trajectories 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 scales. 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 $1.0 \times 10^{-5} s^{-1}$ 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 are identified the tracking is performed, which involves the minimization of a cost function to obtain smooth trajectories (storm tracks). Only those storm tracks that lasted at least two days, traveled further than 1000 km and had a majority of their lifecycle in 20^{o}N - 90^{o}N or 20^{o}S - 90^{o}S are retained for the statistical analysis. The tracking is also performed with an analysis dataset so that the forecast tracks can be verified.

In order to validate the forecast storm tracks against the analysis storm tracks, it is necessary to have a systematic method of determining which forecast storm tracks correspond to which analysis storm tracks. The matching methodology of Froude et al. (2007b) is used, in which a forecast storm track is considered to be the same system as an analysis storm track (i.e. matched) if the two tracks meet certain predefined spatial and temporal criteria. A forecast track is said to match an analysis track if:

1. At least T% of their points coincide in time, i.e. $T = 100 \times \left(\frac{2n_m}{n_a+n_f}\right)$ 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 points in the analysis track that coincide in time with the forecast track.

2. The geodetic separation distance d between the first k points of the forecast track, which coincide in time with the analysis track, and the corresponding points in the analysis track is less than S^o, i.e..$d \leq S^o$.

The forecast tracks that matched analysis tracks are used to generate diagnostics concerning the position, intensity and other properties of the cyclones. In Froude et al. (2007a,b) the sensitivity of the diagnostics to the choice of parameters k, T and S was explored in detail. They found that, although the number of forecast cyclone tracks that matched analysis tracks varied with different choices of the parameters, the diagnostics produced from the matched tracks were basically unaffected. For this paper all the results were obtained using the parameters $k = 4$, $T = 60$% and $S = 4^o$.

As an additional constraint, only those cyclones whose genesis occurs within the first 3 days of the forecast or that already existed at time 0 are considered. Results from the study of Bengtsson et al. (2005) indicated that the skill in predicting cyclone tracks after 3 days is relatively low. If a cyclone was generated in a forecast at a lead time (the time since the start of the forecast) greater than 3 days, and matches a cyclone in the analysis, then it was probably more due to chance than an accurate prediction. Although this may not be the case for the more recent forecast and analysis systems, this constraint is kept so that the methodology is consistent with Froude et al. (2007a). For further details of the methodology please see Froude et al. (2007a,b); Froude (2009); Hodges (1995, 1999); Hoskins and Hodges (2002, 2005).

4 TIGGE: Comparison of the Prediction of Extratropical Cyclones by different Ensemble Prediction Systems

In this section some results are presented which apply the storm tracking methodology to different ensemble prediction systems (EPS) using data from the TIGGE project. Since 1 February 2008 the TIGGE archive has included EPS data from 10 different operational weather centres, namely the Australian Bureau of Meteorology (BoM), the Chinese Meteorological Administration (CMA), the Canadian Meteorological Centre (CMC), the European Centre for Medium Range Weather Forecasts (ECMWF), the

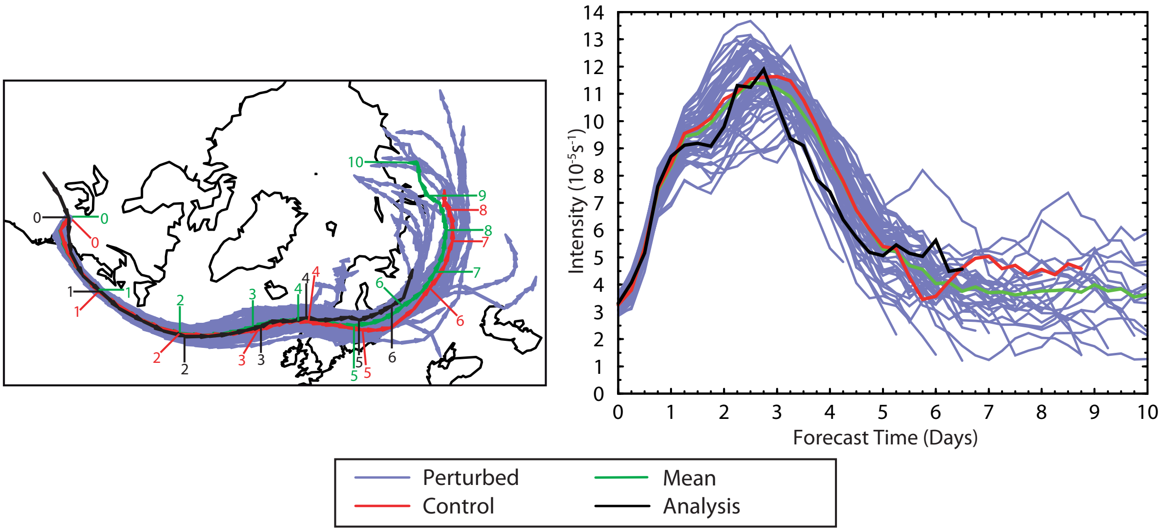

Figure 1: Tracks (a) and intensities (b) of an Atlantic cyclone predicted by the ECMWF EPS. The ECMWF analysis is also shown. Units of intensity are $1.0 \times 10^{-5}s^{-1}$ *(relative to background field removal) and the numbers along the tracks correspond to the forecast lead time in days. The forecast start time (day 0) is 12 UTC on 22 February 2008.*

Japanese Meteorological Administration (JMA), the Korean Meteorological Administration (KMA), the National Centers for Environmental Research (NCEP), the UK Met Office (UKMO), the Brazilian Centre for Weather Prediction and Climate Studies (Centro de Previsao de Tempo e Estudos Climaticos, CPTEC) and Meteo France. EPS data for all of these centres, except Meteo France, has been analysed for the 6 month time period of 1 February 2008 - 31 July 2008. Meteo France was excluded because their forecasts are only integrated out to 3 days, which is not long enough to include the full life cycle of a large number of extratropical cyclones. The storm identification and tracking was performed along each ensemble member and control forecast for each EPS for the 6 month time period. It was also performed with the ECMWF analysis for verification. Since the cyclones are verified against the ECMWF analysis, there may be some positive bias towards ECMWF in the results. However this will probably only be significant in the earlier part of the forecast (Bengtsson et al., 2005).

Figure 1a shows an example of the tracks and intensities of an Atlantic cyclone predicted by ECMWF. The analyzed Atlantic cyclone (shown in black) formed over North America at 00 UTC on 22 February 2008. It then travelled across the Atlantic, intensifying rapidly over the next 3 days before reaching its maximum relative vorticity amplitude of $11.9 \times 10^{-5}s^{-1}$ at 06 UTC on 25 February. The cyclone then moved north of the British Isles, over Scandinavia, and just into Russia while decaying over the next 3.5 days.

The ensemble member tracks are tightly spaced around the analysis track indicating that this particular cyclone is highly predictable. The mean track (calculated by averaging all the ensemble member tracks) and the control track lie virtually on top of each other until day 4 of the forecast. From this point the control track is slightly too far to the south and the mean is closer to the analysis. The spread in the intensity for this cyclone is also small, particularly during the initial growth phase in the first day of the forecast. From this point the ensemble members are more dispersed. Both the ECMWF control and ensemble mean exhibit high levels of predictive skill for this cyclone.

Figures 2a and b show the mean tracks and mean intensities respectively for each of the nine EPSs. The track of the cyclone is predicted very well by all the centres, although some of the forecast cyclones travel considerably farther into Russia than the analyzed cyclone. There is a larger difference in performance between the centres for the cyclone's intensity than track. Overall ECMWF and KMA have the highest

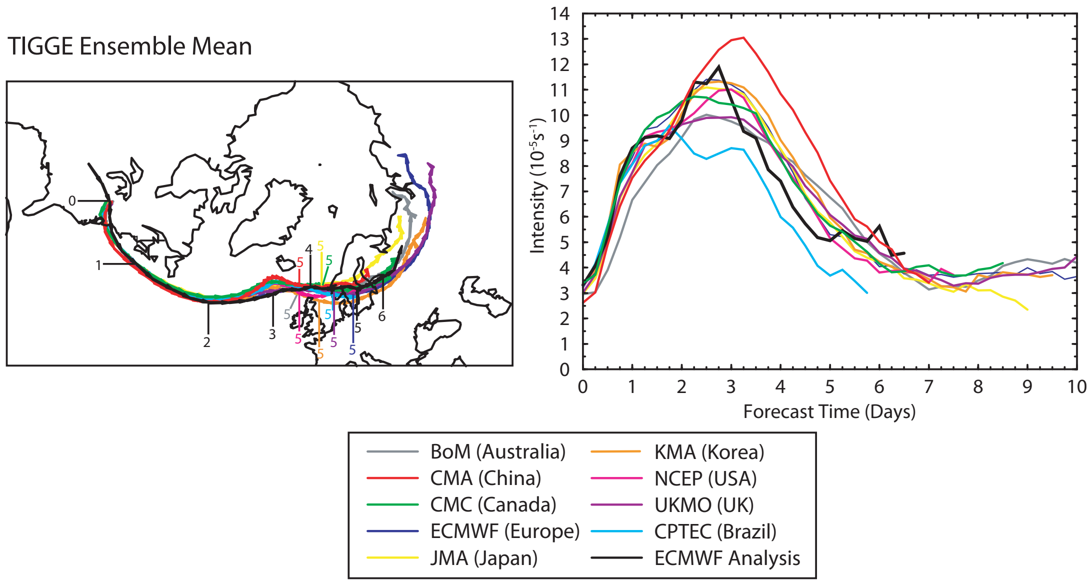

Figure 2: Tracks (a) and intensities (b) of an Atlantic cyclone predicted by the ensemble mean (calculated by averaging all the ensemble member tracks/intensities) of each EPS. The ECMWF analysis is also shown. Units of intensity are $1.0 \times 10^{-5}s^{-1}$ *(relative to background field removal) and the numbers along the tracks correspond to the forecast lead time in days. The forecast start time (day 0) is 12 UTC on 22 February 2008. (Figure from Froude, 2010a)*

level of performance. The CMA mean and control overpredict the maximum intensity of the cyclone and the other centres have an underprediction.

For this particular cyclone there is only a small difference in skill between the control and ensemble mean. However, for other cyclones, there can be a larger difference. The relative performance of the different EPSs can also vary considerably for different cyclones (see Froude, 2010a). This highlights the importance of performing a statistical analysis of a large number of cyclones to assess the skill and determine the strengths and weaknesses of the different EPSs.

Figure 3 shows the ensemble mean error in cyclone position, intensity and propagation speed for each EPS in the NH for the 6-month time period. To calculate the ensemble mean error, the mean track, mean intensity, and mean propagation speed of the matching ensemble member tracks (including the control) are computed for each cyclone in each ensemble forecast at each forecast lead time. The mean error in position is calculated as the mean geodetic separation distance between the mean tracks and the corresponding ECMWF analysis tracks. Also the mean intensity error was calculated similarly, from the filtered vorticity value at the cyclone centres, using the absolute intensity difference as the measure of error. The propagation speeds of the analysis and ensemble member cyclones were calculated at each point on their tracks by comparing the position of consecutive points on the tracks. Since the points on the tracks are 6 hours apart, the speed calculated at each point corresponds to the average propagation speed of the cyclone in the next 6 hours.

Firstly considering the position of the cyclones, there is a large difference in the predictive skill of the different EPS. ECMWF shows the highest level of skill, although there may be some bias because all the EPS were verified against the ECMWF analysis. JMA, NCEP, UKMO and CMC have approximately 1 day less skill than ECMWF throughout the forecast. It is worth commenting that while CPTEC has the least skill this is perhaps to be expected since the NH extratropical region is not the focus in the construction of their ensemble (e.g. they only apply perturbations in the region of 45S-30N).

For the intensity of the cyclones the skill of the different EPS in relation to each other in general remains

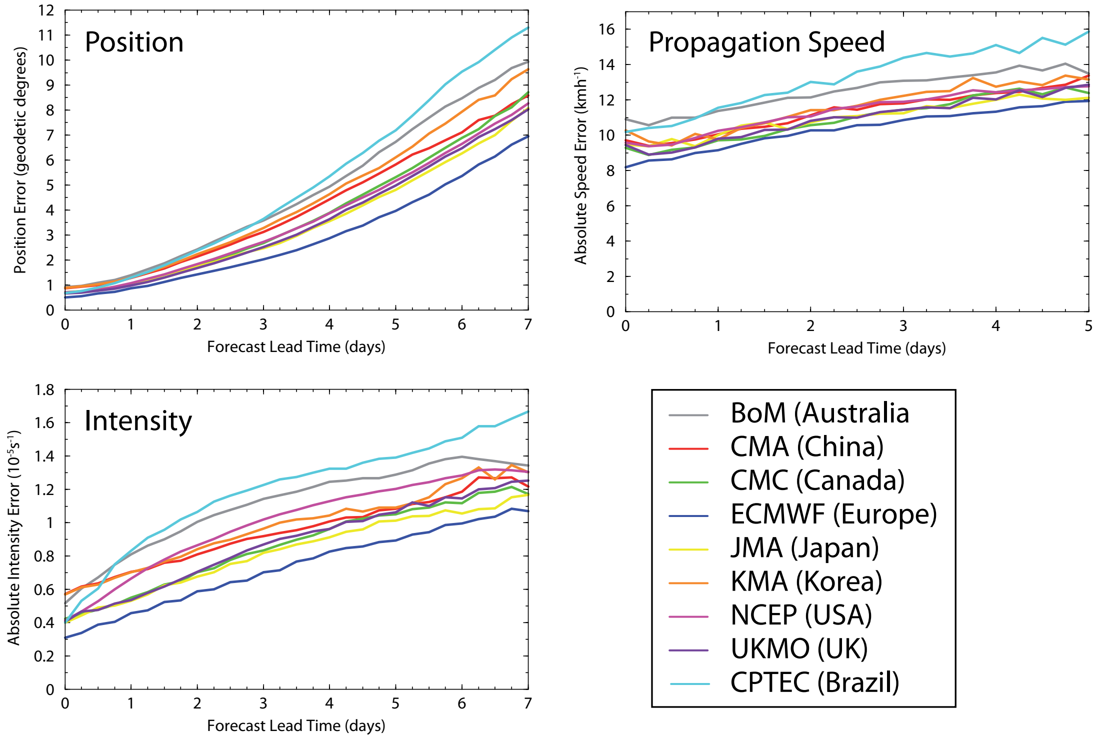

Figure 3: EPS mean error in (a) position, (b) intensity and (c) propagation speed. Units of position, intensity and propagation speed error are geodetic degrees, $1.0 \times 10^{-5}s^{-1}$ *(relative to background field removal) and* kmh^{-1} *respectively. (Figure from Froude, 2010a)*

the same. That is EPS with smaller/larger errors in position generally 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. For position NCEP has errors comparable to the CMC, UKMO and JMA, but for intensity it has larger errors comparable with CMA and KMA. The error growth is faster initially for NCEP, CPTEC and BoM. This is perhaps because these EPS are integrated at low resolutions and are not able to accurately capture the cyclones' growth and decay (see for example Jung et al., 2006). However the CMC EPS is also integrated at a low resolution and does not have this rapid error growth. Perhaps the use of 4DVar is compensating for this by providing a better initial state (e.g. Johnson et al., 2006).

The mean error in propagation speed is large throughout the forecast for all the EPS ranging from around 8-16 kmh^{-1}. It should be noted that the speed error is different in nature to the position or intensity error in that it would not necessarily be expected to grow with lead time. However, there will be a cumulative effect of a consistent error in speed on the position of the cyclone with increasing lead time. The relative skill of the different EPS is similar to the position error. It was only possible to plot the propagation speed error to day 5 as there was insufficient data beyond this point for this particular diagnostic.

Figure 4a shows the bias in the intensity error given in Figure 3b. CMC has the smallest bias (not exceeding $0.5 \times 10^{-5}s^{-1}$) and ECMWF, CMA, JMA and KMA also all have small biases. ECMWF is the only system which consistently overpredicts the intensity of cyclones. JMA and KMA underpredict, and CMA and CMC vary, but the biases of all these systems are very small. On the other hand, BoM, NCEP, CPTEC and UKMO all significantly underpredict cyclone intensity. BoM, NCEP and CPTEC in particular show a dramatic increase in negative bias in the earlier part of the forecast (shorter lead time). This corresponds to the rapid error growth exhibited by these systems in the earlier part of the forecast (figure 3b).

BoM, NCEP, CPTEC and UKMO all significantly underpredict cyclone intensity. BoM, NCEP, and

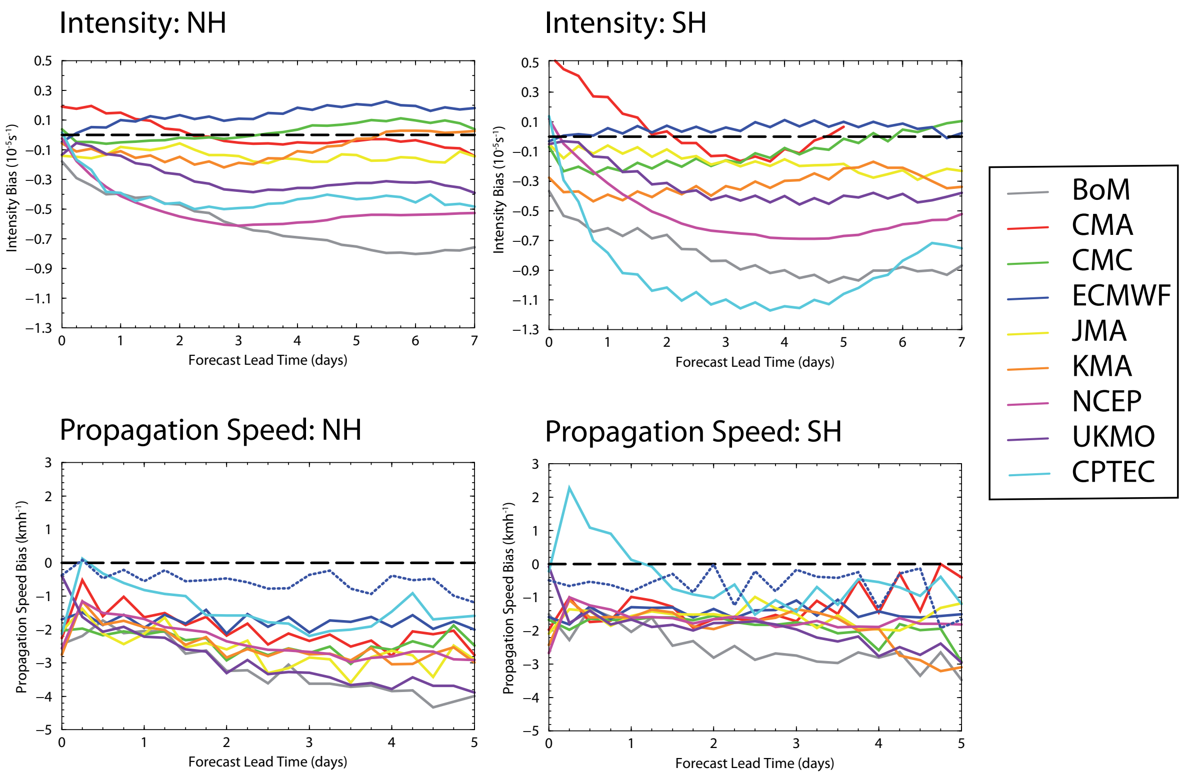

Figure 4: EPS mean bias in (a) intensity and (b) propagation speed. Units of intensity and propagation speed bias are $1.0 \times 10^{-5}s^{-1}$ *(relative to background field removal) and kmh*$^{-1}$ *respectively. (Figure from Froude, 2010a)*

CPTEC in particular show a dramatic increase in negative bias in the earlier part of the forecast. This corresponds to the rapid error growth exhibited by these systems in the initial period (Figure 3b).

Figure 4b shows the bias in the propagation speed error given in Figure 3c. It is interesting that all of the EPS underpredict the propagation speed of the cyclones. Hence cyclones will in general arrive earlier than they are forecast to. The magnitude of the bias varies between centres with BoM and UKMO having the largest and CPTEC, CMA and ECMWF having the smallest. A similar negative bias was found for the control forecasts of each EPS (not shown). This shows that the bias must be due to a deficiency in the models rather than the perturbation methodologies. The magnitude of this bias is small, but the cumulative effect will result in the 5-day forecast being approximately 200-400 km behind the analysed cyclone, which would be of importance to many forecast users.

For further details of this work please see Froude (2010a,b). The key results of this study were as follows:

- There are large differences between the different EPS in the skill of predicting extratropical cyclones.
- The ECMWF ensemble mean and control forecast has the highest skill for all cyclone properties.
- The ensemble mean provides little advantage over the control forecast for cyclone position, but it provides a significant advantage for cyclone intensity.
- The ECMWF and JMA EPS have excellent spread-skill relationships for cyclone position.
- The EPS are much more underdispersive for intensity and propagation speed than for position.
- The cyclones propagate too slowly in all the EPS.
- The UKMO, NCEP, BoM and CPTEC EPS underpredict storm intensity and the other EPS have smaller bias.

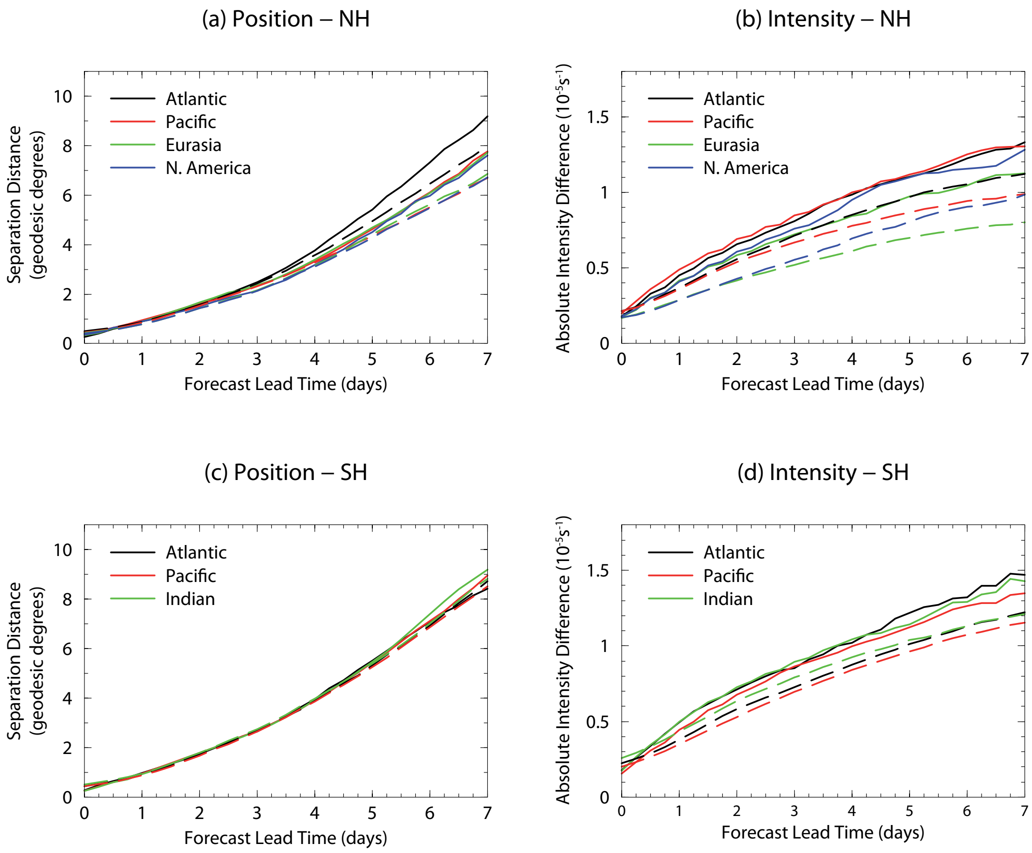

Figure 5: ECMWF ensemble mean error (solid lines) and ensemble spread (dashed lines) in cyclone position, in (a) NH and (c) the SH, and cyclone intensity, in (b) NH and (d) SH, for the different regions as a function of forecast time. Units of separation distance and intensity difference are geodesic degrees and $10^{-5}s^{-1}$ (relative to background field removal) respectively. (Figure from Froude, 2009)

5 Regional analysis of the Prediction of Extratropical Cyclones by the ECMWF EPS

The storm track analysis of the TIGGE data sets was performed on entire hemispheres. A regional analysis of the ECMWF EPS has also been performed (Froude, 2009). Some of the results from this study are presented in this section. In order to explore smaller geographic regions it was necessary to have a larger data sample. For this study the 1 year period of 6th January 2005 - 5th January 2006 was analysed. Four different regions were considered in the NH, in the extratropical latitude band of $20^o - 90^oN$ and within the following longitude bands: 1) Atlantic $= 280^oE - 0^o$, 2) Pacific $= 220^o - 240^oE$, 3) Eurasia $= 0^o - 120^oE$ and 4) North America $= 240^o - 280^oE$. In the SH three different regions were considered in the latitude band of $90^o - 20^oS$ and within the following longitude bands: 1) Atlantic $= 300^o - 20^oE$, 2) Pacific $= 150^o - 290^oE$ and 3) Indian $= 20^o - 120^oE$.

Figure 5 shows the ensemble mean error and the ensemble spread for the position and intensity of the cyclones for all the regions in the northern and southern hemispheres. Ensemble spread was calculated as the mean geodesic separation distance (absolute intensity difference) of the ensemble member tracks from the ensemble mean track. For an EPS to be statistically reliable the average distance of the ensemble mean from the analysis should be equal to the average distance of the ensemble members from the ensemble mean (i.e. mean error = spread).

Firstly considering the position of the cyclones, in the NH the Atlantic region has larger ensemble mean error than the other regions from day 3 of the forecast (approximately 1 day less skill from day 4). The other regions in the NH have comparable error. In the SH there is no real difference in the error between the regions, but the errors are slightly larger than in the NH. The spread and skill curves are almost identical in the SH, but in the NH the ensemble is slightly under-dispersive, from day 4 for the Atlantic regions and from day 5 for the other regions.

For the intensity, the ensemble mean error is larger over the ocean regions (Atlantic and Pacific) than over the land regions (Eurasia and North America). This is probably to be expected with this absolute measure of intensity error, since storms would generally be expected to more intense over the ocean than over the land. Observational coverage probably also plays a role, since the higher frequency of surface and upper air observations over the land will also improve the prediction of the intensity of the storms. In the SH, as with the position error, the mean error is fairly comparable between regions. Since the SH regions are all oceanic, there will be less variation in the intensity and other properties of the storms between regions, which is reflected in the statistics. The ensemble spread is significantly under-dispersive from very early on in the forecast in all regions, but particularly in the NH (also found previously by Froude et al., 2007b). In the NH, there is a smaller difference between the mean error and spread curves for the Atlantic region ($2\times10^{-5}s^{-1}$ at day 7) than the other regions ($3\times10^{-5}s^{-1}$ at day 7). In the SH, the difference between the mean error and spread curves is similar in all regions, but is less than in the NH.

Figure 6a shows the intensity bias separately for the perturbed ensemble members and control forecast for the NH and SH regions respectively. Rather interestingly the diagnostics show that the EPS over-predicts cyclone intensity over the ocean regions (Atlantic and Pacific in NH and all regions in SH) and underpredicts the intensity over the land. The small magnitude of the bias should be noted, but there is a clear systematic pattern between the ocean and land based regions. It should also be noted that the biases were computed using the filtered values of vorticity (see section 3). If the biases were computed from the original values of vorticity of the full resolution data, then higher values may be obtained (this will be investigated as future work). In the oceanic regions of the NH the bias grows in magnitude until day 2 (presumably corresponding to the optimisation time of the singular vector perturbations), it then decreases and becomes slightly negatively biased from day 3. In the SH, the bias also increases until day 2, but it then levels off rather than decreasing as in the NH. This difference was also found by Froude et al. (2007b).

Figure 6b shows the propagation speed bias. It was not possible to show the results for North America as the data sample was insufficient for this particular diagnostic. This region is smaller than the other regions and so less data is available to produce stable statistics. There is a negative bias, corresponding to the forecast storms propagating too slowly, for all of the regions. The NH Atlantic region has a bias of twice the magnitude of the other regions, which corresponds with the larger position error in this region (figure 5a).

For further details of this work please see Froude (2009). The key results of this study were as follows:

- The error in cyclone position is larger over the Atlantic in the NH. It is larger in the SH (than the NH) but comparable between the regions.
- The error in cyclone intensity is larger over the ocean than over the land.
- The spread in position is slightly less than mean error in position from day 3 for all regions in NH, but in the SH they are comparable.
- The spread in intensity is less than mean error in intensity for all regions, with larger differences for NH Pacific and Eurasia.
- In general storms are overpredicted over the ocean and underpredicted over the land.
- The forecast storms move too slowly in all regions, but the bias is larger over the Atlantic in the NH.

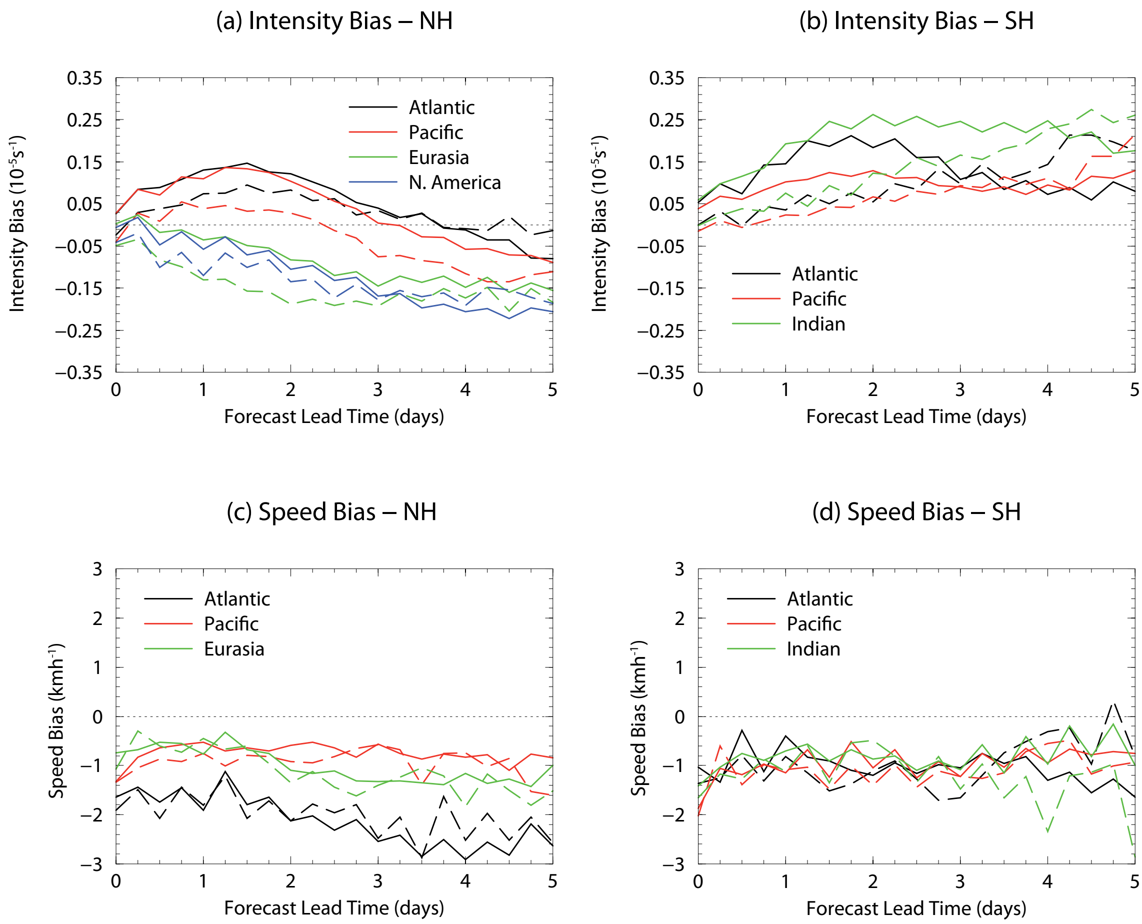

Figure 6: Bias in intensity in (a) NH and (b) SH and propagation speed in (c) NH and (d) SH of the ECMWF perturbed member tracks (solid lines) and control forecast tracks (dashed lines) for the different regions as a function of forecast time. Units of intensity and speed error are $10^{-5}s^{-1}$ *(relative to background field removal) and kmh*$^{-1}$*. (Figure from Froude, 2009)*

6 Final Remarks and Future Directions

Extratropical cyclones are fundamental to the everyday weather of the midlatitudes. They provide essential rainfall for human activities such as agriculture, but can also cause large amounts of damage by their strong winds and heavy precipitation. It is therefore very important that these cyclones are predicted as accurately and as far in advance as possible by NWP. In the past studies of the prediction of extratropical cyclones have mainly focused on individual cyclones or cyclone simulations. There have been some statistical studies, but these have used manual or semi-automated methods to identify and track the cyclones. As a result these studies have been limited due to the large amount of work involved. This paper has described a storm tracking forecast verification methodology and has given a overview of some of the results that have been obtained from its implementation. Detailed information about the prediction of current NWP, particularly EPS, has been obtained.

In future work we plan to explore the causes of error in storm prediction. More sophisticated diagnostics will be applied to forecast models to explore the vertical structure, tilt and lifecycle of extratropical cyclones. This will provide further understanding of the causes of error in storm prediction in relation to the cyclone dynamics. Forecast experiments will be performed to assess the impact of different factors such as resolution, perturbation methods and ensemble size. We also plan to work more directly with forecast users who could benefit from storm prediction information. We have just begun a new project with marine informatics company British Marine Technology Group Ltd ARGOSS (BMT ARGOSS. http://www.argoss.nl). The project will involve the development of forecast tools for providing storm prediction/uncertainty information from NCEP EPS data for decision making at sea. Accurate information about storm prediction is vital for many marine activities such as ship routing and oil and gas operations.

The storm tracking methodology presented in this paper provides detailed statistical information about the prediction of extratropical cyclones by NWP. It provides an alternative method of forecast verification to more conventional approaches such as root mean square error or anomaly correlation coefficient. Since storms are so fundamental to the day to day weather in the midlatitudes the storm tracking approach provides useful information about the ability of NWP to predict the weather. The disadvantage of the method is that it is more time consuming and requires larger data samples than the conventional forecast verification approaches. The approach is potentially useful to both forecast users and model developers.

Acknowledgements

The author would like to thank Kevin Hodges, Robert Gurney and Lennart Bengtsson for their help and advice on this work and the TIGGE contribution centres and data centres are acknowledged for providing the EPS data.

References

Akyildiz, V., 1985: Systematic errors in the behaviour of cyclones in the ECMWF operational models. *Tellus*, **37A**, 297–308.

Baehr, C., B. Pouponneau, F. Ayrault, and A. Joly, 1999: Dynamical characterization of the FASTEX cyclogenesis cases. *Q. J. R. Meteorol. Soc.*, **125**, 3469–3494.

Bengtsson, L., K. I. Hodges, and L. S. R. Froude, 2005: Global observations and forecast skill. *Tellus*, **57A**, 515–527.

Bougeault, P. and Coauthors, 2010: The THORPEX interactive grand global ensemble (TIGGE). *Bull. Am. Meteorol. Soc.*, **91**, 10591072.

Courtier, P. and J. F. Geleyn, 1988: A global spectral model with variable resolution. Applications to the shallow water equations. *Q. J. R. Meteorol. Soc.*, **114**, 1321–1326.

Froude, L. S. R., 2009: Regional differences in the prediction of extratropical cyclones by the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **137**, 893–911.

— 2010a: TIGGE: Comparison of the prediction of northern hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **25**, 819–836.

— 2010b: TIGGE: Comparison of the prediction of southern hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **in review**.

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

— 2007b: The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. *Mon. Wea. Rev.*, **135**, 2545–2567.

Girard, C. and M. Jarraud, 1982: Short- and medium- range forecast differences between a spectral and grid-point model. ECMWF Technical report no. 32, 176pp, ECMWF, ECMWF, Shinfield Park, Reading, UK.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A desscription of the fifth generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note 398+STR, NCAR, Boulder, CO, 138pp.

Grumm, R. H., 1993: Characteristics of surface cyclone forecasts in the aviation run of the global spectral model. *Wea. Forecasting*, **8**, 87–112.

Grumm, R. H., R. J. Oravec, and A. L. Siebers, 1992: Systematic model forecast errors of surface cyclones in the NMC's nested-grid model, December 1988 through November 1990. *Wea. Forecasting*, **7**, 65–87.

Grumm, R. H. and A. L. Siebers, 1989: Systematic surface cyclone errors in the NMC's nested grid model November 1988-January 1989. *Wea. Forecasting*, **4**, 246–252.

— 1990: Systematic model forecast errors of surface cyclones in the NGM and AVN, January 1990. *Wea. Forecasting*, **5**, 672–682.

Hack, J. J., B. A. Boville, B. P. Briegleb, J. T. Kiehl, P. J. Rasch, and D. L. Williamson, 1993: A description of the NCAR Communuty Climate Model (CCM2). NCAR Tech. Note 382+STR, NCAR, Boulder, CO, 108pp.

Hacker, J. P., E. S. Krayenhoff, and R. B. Stull, 2003: Ensemble experiments on numerical weather prediction error and uncertainty for a North Pacific forecast failure. *Wea. Forecasting*, **18**, 12–31.

Hodges, K. I., 1995: Feature tracking on the unit sphere. *Mon. Wea. Rev.*, **123**, 3458–3465.

— 1999: Adaptive constraints for feature tracking. *Mon. Wea. Rev.*, **127**, 1362–1373.

Hoskins, B. J. and K. I. Hodges, 2002: New perspectives on the northern hemisphere winter storm tracks. *J. Atmos. Sci.*, **59**, 1041–1061.

— 2005: A new perspective on southern hemisphere storm tracks. *J. Climate*, **18**, 4108–4129.

Johnson, C., B. J. Hoskins, N. K. Nichols, and S. P. Ballard, 2006: A singular vector perspective of 4DVAR: The spatial structure and evolution of baroclinic weather systems. *Mon. Wea. Rev.*, **134**, 3436–3455.

Jung, T., S. K. Gulev, I. Rudeva, and V. Soloviov, 2006: Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Q. J. R. Meteorol. Soc.*, **132**, 1839–1857.

Jung, T., E. Klinker, and S. Uppala, 2004: Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part I: Analyses and deterministic forecasts. *Meteorol. Appl.*, **11**, 343–361.

— 2005: Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part II: Ensemble forecasts. *Meteorol. Appl.*, **12**, 111–122.

Kuo, Y. H., X. Zou, and W. Huang, 1997: The impact of global positioning system data on the prediction of and extratropical cyclone: An observing system simulation experiment. *Dynamics of Atmospheres and Oceans*, **27**, 439–470.

Langland, R. H., M. A. Shapiro, and R. Gelaro, 2002: Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm. *Mon. Wea. Rev.*, **130**, 957–974.

Leary, C., 1971: Systematic errors in operational National Meteorological Center primitive-equation surface prognoses. *Mon. Wea. Rev.*, **99**, 409–413.

Leutbecher, M., J. Barkmeijer, T. N. Palmer, and A. J. Thorpe, 2002: Potential improvement to forecasts of two severe storms using targeted observations. *Q. J. R. Meteorol. Soc.*, **128**, 1641–1670.

Morris, R. M. and A. J. Gadd, 1988: Forecasting the storm of 15-16 October 1987. *Weather*, 70–90.

Pearce, R., D. Lloyd, and D. McConnell, 2001: The post-Christmas ‘French’ storms of 1999. *Weather*, **56**, 81–91.

Pouponneau, B., T. B. Franck Ayrault, and A. Joly, 1999: The impact of aircraft data on an Atlantic cyclone analyzed in terms of sensitivities and trajectories. *Wea. Forecasting*, **14**, 67–83.

Sanders, F., 1986: Explosive cyclogenesis over the West-Central North Atlantic ocean, 1981-84. Part II: Evaluation of LFM model performance. *Mon. Wea. Rev.*, **114**, 2207–2218.

— 1987: Skill of NMC operational dynamical models in prediction of explosive cyclogenesis. *Wea. Forecasting*, **2**, 322–336.

— 1992: Skill of operational dynamical models in cyclone prediction out to five-days range during ERICA. *Wea. Forecasting*, **7**, 3–25.

Sanders, F. and E. P. Auciello, 1989: Skill in prediction of explosive cyclogenesis over the western north atlantic ocean, 1987/88: A forecast checklist and NMC dynamical models. *Wea. Forecasting*, **4**, 157–172.

Sanders, F., S. L. Mullen, and D. P. Baumhefner, 2000: Ensemble simulations of explosive cyclogenesis at ranges of 2-5 days. *Mon. Wea. Rev.*, **128**, 2920–2934.

Silberberg, S. R. and L. F. Bosart, 1982: An analysis of systematic cyclone errors in the NMC LFM-II model during the 1978-79 cool season. *Mon. Wea. Rev.*, **110**, 254–270.

Smith, B. B. and S. L. Mullen, 1993: An evaluation of sea level cyclone forecasts produced by NMC’s Nested-Grid Model and Global Spectral Model. *Wea. Forecasting*, **8**, 37–56.

Xiao, Q., X. Zou, M. Pondeca, M. A. Shapiro, and C. Veldon, 2002: Impact of GMS-5 and GOES-9 satellite-derived winds on the prediction of a NORPEX extratropical cyclone. *Mon. Wea. Rev.*, **130**, 507–528.

Zhu, H. and A. Thorpe, 2006: Predictability of extratropical cyclones: The influence of initial conditions and model uncertainties. *J. Atmos. Sci.*, **63**, 1483–1497.

Zou, X., Y.-H. Kuo, and S. Low-Nam, 1998: Medium range prediction of an extratropical oceanic cyclone: Impact of initial state. *Mon. Wea. Rev.*, **126**, 2737–2763.