# THE UNIVERSITY OF READING

Department of Meteorology
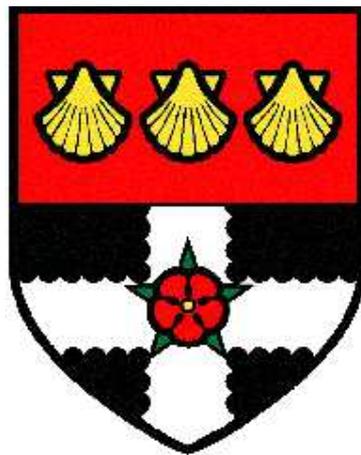
The Impact of Targeted Observations on the Prediction of Weather Systems

Emma Ann Irvine

A thesis submitted for the degree of Doctor of Philosophy

February 15, 2010

Declaration

I confirm that this is my own work and that the use of all material from other sources has been properly and fully acknowledged.

Emma Ann Irvine

Abstract

Targeted observations are additional observations of the atmospheric state taken in regions where the downstream forecast is highly sensitive to the initial conditions. This thesis analyses the impact of targeted observations on the 1-3 day forecasts for northern Europe, the mechanisms for this impact, and the potential size of the impact. The targeted observations were carried out during the Greenland Flow Distortion Experiment (GFDex) and the Norwegian IPY-THORPEX field campaigns, using research aircraft to release dropsondes into total energy singular vector or ensemble transform Kalman filter sensitive areas. The impact of the targeted sondes was assessed using a 24km gridlength limited-area version of the Met Office Unified Model including their variational data assimilation scheme.

The impact of targeted observations on the forecasts for northern Europe is on average small, with a maximum forecast improvement of 5-17% after 18-24 hours (using a total energy metric) for those cases where new information was added to the background field. Targeted observations from the GFDex impacted the forecast through changes to the upper-level fields (through modification of the position of upper-level potential vorticity anomalies). For an IPY-THORPEX case the impact on the forecast of the polar low was caused by the modification of the initial state.

The potential size of the impact is dependent on a number of factors. Increasing the fit of the analysis to the observations increased the maximum improvement by 2-5%. Experiments with singular vector pseudo-soundings showed that approximately 50 observations covering an area of 1×10^6km^2 are required to sufficiently describe any baroclinic structures that might not be well-represented by the background or routine observations. Targeted observations that are positioned within a correlation length-scale (e.g. 200km) of steep orography may degrade the forecast through the anomalous upslope spreading of analysis increments along terrain-following model levels.

Acknowledgments

I am grateful for the help and support provided by very many people during the course of my thesis. I would like in particular to thank the following people:

My supervisors, Sue Gray, John Methven, Ian Renfrew and Richard Swinbank for their continual advice and encouragement. Sue's help and support was invaluable from the start, both in person and occasionally via email and Skype! I am grateful to John for stepping-in to supervise when Sue went on maternity leave, and remaining involved for the course of the project. Ian's superb organisation led to a successful field campaign, and plenty of data for my thesis. Richard arranged access to the Met Office systems, without which few of the experiments in the thesis could have been performed.

Keir Bovis, Richard Dumelow and David Walters at the Met Office for dispensing large quantities of technical help and advice, and patiently and quickly responding to my pleas for help. Keir also ran the first set of GFDex hindcasts and provided the ETKF sensitive-area predictions; David set-up the crisis area mesoscale model for the Norwegian experiments.

My colleagues that I worked alongside during the Greenland Flow Distortion Experiment and IPY-THORPEX field campaigns, not just for their expertise but also for their sense of humour which made the field campaigns so enjoyable. Particular thanks to the Norwegians (project leader Jon Egill Kristjánsson and secretary Gudmund Dalsbø deserve special mention) for making me part of the team, and introducing me to Norwegian food and culture.

Mesoscale group and my colleagues in 1U07 for their help and many conversations, not always work-related.

Nick and my family, for their continual encouragement, which helped me through the difficult parts of my thesis and for supporting my participation in two field campaigns away from home.

Finally my Dad, from whom I inherited my interest in the natural world, and who I greatly miss.

Contents

CHAPTER 1

Introduction

The atmosphere is a complex, highly nonlinear, chaotic system. In order to make a forecast of the atmospheric state at a particular point at some future time, we need two things. Firstly, we need a numerical model that approximates the behaviour of the atmosphere using sets of equations, which we can use to evolve the present state forward in time. Secondly, we need an estimate of the state of the atmosphere at the present time, called the initial conditions. The specification of the initial conditions is a crucial part of the forecast. The chaotic nature of the atmosphere means that by definition, the behaviour of the atmosphere is highly sensitive to the initial conditions. This is more popularly known as the 'butterfly effect,' explained as the flap of a butterfly's wings in Brazil causing a small change to the atmospheric state that results in a tornado in Texas (from a talk by Edward Lorenz in 1972). To get the initial conditions we take observations of the state of the atmosphere using instruments at the ground and on radiosondes, satellites and aircraft. These observations that are fed into the numerical models every day are referred to as routine observations throughout this thesis. The routine observations do not cover the entire globe - in regions which are remote or inaccessible (e.g. high latitude or over the ocean) or regions of political unrest (e.g. large regions of Africa) there are far fewer observations. Therefore the observations have to be combined with an estimate of the initial state from the numerical model, usually a six hour forecast is used, in a process called data assimilation. This produces the initial state which is evolved forward in time using the numerical weather prediction (NWP) model to produce the weather forecast.

Weather forecasts are by their nature imperfect. They will always have errors because our knowledge of the dynamics of the atmosphere is imperfect, and our ability to put this knowledge into computer programs that represent the state of the atmosphere is also limited. This is model error. Even if our knowledge of the processes which govern our weather was perfect, and we could produce a perfect NWP model, our forecasts would still have errors. This is because we can never have the complete and accurate knowledge of the state of the atmosphere at any one point in time which we need to initialise our weather forecast model.

This is initial condition error. In parts of the globe where fewer observations are made, such as in remote or inaccessible regions, the initial conditions may be less accurate as there are less observations to constrain the model, and therefore the initial condition error may be larger (indeed, the use of the forecast model to produce the initial conditions means that model error may contribute to initial condition error). This may lead to forecast busts, where the initial condition error grows significantly throughout the forecast, deteriorating the forecast downstream of the initial condition error. In the mid-latitudes baroclinic wave packets propagate to the east with a group velocity that is faster than the phase speed of the individual systems, leading to downstream development. Therefore an initial error at one location will affect the forecast downstream of that location. Rabier et al. (1996) suggest that initial condition errors may explain many major forecast errors. For example accurate initial conditions are needed throughout a cyclone's lifetime to accurately model the development of a cyclone, as both the forecast error resulting from inaccurate initial conditions and cyclone itself grow exponentially (Zhu and Thorpe, 2006). Shutts (1990) showed that the forecast of the October 1987 storm, which caused much damage to southern England, was greatly improved by the assimilation of one late aircraft observation. In this thesis a method is explored that has potential to improve forecasts by reducing the initial condition error in data sparse regions. The method, known as targeted observing, is explained in the following section. The thesis aims and structure are set out in Section 1.2.

1.1 Targeted Observations

Targeted observations are observations of the initial state that are made in regions where there are few routine observations, in order to try and reduce the initial condition error in this region and therefore improve the downstream forecast. The observations are targeted in the sense that they are located in 'sensitive regions' where the downstream forecast is highly sensitive to errors in the initial conditions, and therefore where the targeted observations should have maximum forecast impact. Here the word sensitive is used in the sense that a small perturbation to the initial state has a large impact on the forecast, through rapid error growth. Targeting observations to improve operational weather forecasts relies on the ability to accurately diagnose these sensitive areas with enough notice to mobilise observing platforms, such as aircraft, from which the additional observations are made. Targeted observing has only be-

come possible operationally in the past decade through advances in computing speed and the development of mathematical techniques to identify regions of sensitivity in the forecast. Targeted observations could have potential application operationally as part of an adaptive observing system, where targeted observations complement the routine observation network.

Targeted observing could provide large gains in forecast improvement and therefore forecast lead time, for a single event. However, targeted observing using dropsondes is expensive to implement; the cost of the resources to deploy and make the additional observations is not insignificant, without even considering the computational cost of running sensitive area predictions, and the cost of manpower to analyse the predictions and make the additional observations. The use of targeted observations could therefore only be justified where there is the possibility of a high-impact weather event, such as damaging winds or heavy precipitation. Targeted observing could be performed using additional satellite data (by less thinning of satellite data in sensitive regions); this would greatly reduce the cost of targeting. Evaluating the potential benefit of targeted observations is one of the aims of The Observing System Research and Predictability Experiment (THORPEX), a World Meteorological Organisation program (Shapiro and Thorpe, 2004). The research presented in this thesis addresses some of the aims of the THORPEX program.

In this thesis the impact of targeted observations on the forecasts for northern Europe is examined. Figure 1.1 shows the study area for this thesis. Targeted observations were made during two aircraft-based field campaigns in the high-latitudes in regions where the routine observing network is sparse (i.e. has few components). The Greenland Flow Distortion Experiment (GFDex; Renfrew et al. (2008)) targeted observations in the region around Iceland and southern Greenland (region A; Figure 1.1) in February and March 2007. The Norwegian IPY-THORPEX field campaign (IPY relates to the 2007-2009 International Polar Year) targeted observations in the Norwegian sea (region B; Figure 1.1) in March 2008. Sections 1.1.1 and 1.1.2 motivate targeting in these regions and describe the key atmospheric conditions, relevant to the targeting cases, that occur in these regions.

1.1.1   Motivation for Targeting Around Southern Greenland

There are few observations on or around Greenland, yet Greenland has a substantial modulating effect on the atmospheric flow and these effects are not well-captured by NWP models.
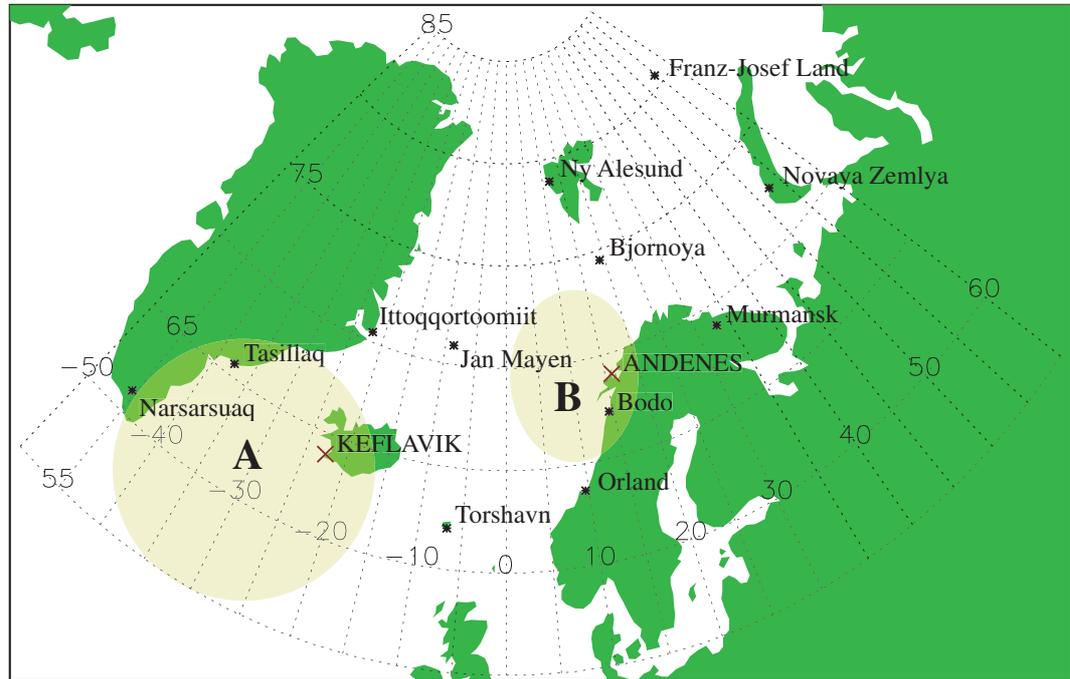
Figure 1.1: *A map showing the study area for this thesis. Relevant radiosonde stations are labelled with stars. The aircraft bases for the GFDex and IPY-THORPEX field campaigns are labelled with red crosses (note that Keflavik is also a radiosonde station). The shading shows the approximate regions where observations were targeted in GFDex (A) and IPY-THORPEX (B).*

This has the potential to lead to initial condition errors which may grow and deteriorate the forecast downstream of Greenland, over Europe. Extending from 60°N to 80°N with an elevated plateau that is up to 3,500m above sea-level, Greenland is a major obstacle to the flow in the high-latitudes of the northern hemisphere. The deformation of the local flow field by Greenland's orography results in three observed phenomena around its coasts: tip jets and reverse tip jets are observed near Cape Farewell (the southern tip of Greenland) in the south, and barrier winds occur along its eastern and south-eastern coasts. On the synoptic-scale the deformation results in lee cyclogenesis and blocking. The field campaign aspect of GFDex aimed to obtain in-situ measurements of high-wind speed events around southern Greenland, to advance our understanding of the flow deformation by Greenland and its effect on downstream predictability (Renfrew et al., 2008). A second aim of GFDex was to investigate the potential of targeted observations in this region. The field campaign goals were achieved during a three-week (19 February to 12 March 2007) field campaign, which used the Facility for Atmospheric Airborne Measurement (FAAM) research aircraft based out of Keflavik, Iceland

(Figure 1.1).

Of the high-wind speed events around the Greenland coasts, only barrier winds occurred during a targeting flight, therefore tip jets and reverse tip jets are not discussed here. Barrier winds are observed along both Greenland's eastern and western coasts, however it is the northerly barrier winds along Greenland's south-east coast which are of interest here. They occur when there is a low pressure system between Greenland and Iceland, which forces cold air against the 'barrier' which is the steep east Greenland coast. Moore and Renfrew (2005) argued that this created a pressure gradient perpendicular to the barrier, resulting in geostrophic flow along the barrier. However, aircraft observations from GFDex have disproved this theory, showing that part of the flow is in fact ageostrophic (Petersen et al., 2009). This study also demonstrated that the main synoptic control on the barrier wind is the position and strength of the lee cyclone. Where there is katabatic flow down the fjords of south-eastern Greenland (these damaging downslope flow events are referred to locally as Piteraqs) the barrier flow may be enhanced, as the katabatic flow turns anti-cyclonically as it exits the fjords (Klein and Heinemann, 2002).

In modelling studies of the effect of Greenland's orography on the general circulation, Petersen et al. (2003) found upstream blocking was an important characteristic of the flow field around Greenland. Evidence for this is the dipole in surface pressure often observed on synoptic charts with high pressure over southern Greenland and correspondingly low pressure immediately to the east of Greenland. Blocking of cold dense air on the upstream side of a mountain leads to an increase in drag, thus decelerating the flow (Ólafsson and Bougeault, 1997).

Mesoscale cyclones are observed year-round in the Arctic and high-latitude oceans, with an observed frequency maxima in cyclones off the south-east Greenland coast in autumn and winter (Harold et al., 1999a). These cyclones are typically 200-600km diameter, and generally have a surface wind speed less than $15\mathrm{ms}^{-1}$, and so are not of sufficient strength to be considered a polar low (see Section 1.1.2). This maxima is absent in the spring and summer when there is less mesoscale cyclone activity over the whole region. This is linked to a seasonal reduction in the extent of the pack ice off the south-east Greenland coast and around Cape Farewell. A maximum in cyclone activity in this region is likely to be associated with cyclogenesis caused by one of two mechanisms: lee cyclogenesis or baroclinic instability

along the ice edge (Harold et al., 1999b). Lee cyclogenesis occurs as a result of flow over the topographic barrier undergoing vortex stretching as it descends. Baroclinic instability is the conversion of potential energy (associated with large horizontal temperature gradients) into kinetic energy. There is a region of baroclinic instability in the lee of Greenland, where there is the cold land and sea ice to the west, and warmer Atlantic ocean to the far east. The temperature gradients are strongest in winter, and this is when the generation of cyclones through baroclinic instability is prevalent. The importance of these two mechanisms is supported by the results of Klein and Heinemann (2002) who showed that katabatic winds contribute to the formation of mesocyclones through the generation of cyclonic vorticity as the downslope flow is stretched, and also the enhancement of the low-level baroclinic environment. Another possible mechanism for cyclogenesis is the generation of low-level potential vorticity (PV) streamers generated through flow-splitting or gravity wave breaking, which has been observed to occur in the lee of the Alps (Aebischer and Schär, 1998).

1.1.2   Motivation for Targeting Polar Lows

Polar lows were specifically targeted for the first time during the IPY-THORPEX field campaign. The campaign aimed to get in-situ measurements of weather events in the Arctic, such as polar lows, Arctic fronts, and flows influenced by orography, in order to better understand these phenomena and improve weather forecasts of them. This was achieved using the German Aerospace Centre (Deutsches Zentrum für Luft-und Raumfahrt; DLR) aircraft, which was deployed from Andenes in Norway (Figure 1.1) during February and March 2008. Polar lows are short-lived mesoscale weather systems that occur in the polar regions and are associated with severe weather such as strong surface winds and heavy snowfall. Following Rasmussen and Turner (2003) we consider a mesoscale weather system to be a polar low if it forms over the sea to the north of the polar front, has a horizontal scale of between 200-1000km and has surface winds of at least gale force (above 15ms^{-1}). The severe weather associated with polar lows, and the difficulty in producing accurate forecasts of their occurrence makes them a hazard to ships, oil rigs and coastal communities in the Arctic regions. With the observed decrease in Arctic sea-ice predicted to continue (see for example Arzel et al. (2006)), the Arctic is becoming more easily accessible. This is leading to increased use by humans, whether for drilling of oil and gas, new shipping routes along the North-east passage or for tourism, thus potentially placing more people in the path of polar lows.

In a changing climate, the frequency and location of polar low occurrence itself may change. The current generation of climate models have horizontal grid lengths on the order of 100km, an insufficient horizontal resolution to be able to resolve polar lows. To resolve the smaller polar lows, which have a horizontal scale of approximately 200km, it would be necessary to have a horizontal grid length of 20km or smaller. Despite this shortcoming, it is possible to use climate models to study whether future climatic conditions will be more or less favourable for the development of polar lows than at present. Kolstad and Bracegirdle (2008) looked at the occurrence of marine cold-air outbreaks in the Norwegian sea using climate model output that was used to form the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report in 2007. They found that as the limit of the Arctic sea-ice retreated northwards, so the area where conditions were favourable for polar low formation moved with it. The effect of this retreat on polar low formation is not clear. Polar lows may form further north and in regions where at the moment the sea-ice extends to the coast, prohibiting the formation of polar lows due to the lack of latent heat flux from the ice-covered ocean. The authors suggest that the northward retreat of the sea-ice could lead to stronger polar lows in areas like the Norwegian seas, where polar lows will have a longer fetch of warm water over which to travel before they make landfall (although the Arctic air may also be warmer which would reduce the likelihood of polar lows). With increased human use of Arctic regions, and indications that polar lows will still be a real hazard to the users of these regions, it is important to be able to produce more accurate forecasts of polar lows. Targeting observations in the data-sparse regions where polar lows form may represent a way of achieving this, while at the same time expanding the dataset of in-situ measurements of these phenomena.

1.2 Thesis Aims

This thesis contributes to the on-going assessment of the impact of targeted observations on a forecast and the potential benefit of targeted observations. It also covers the more general areas of numerical weather prediction, data assimilation and mid- to high-latitude dynamics. This thesis has three main parts. The first part of the thesis assesses the impact of targeted observations made during the GFDex, including how the targeted observations had an impact on forecasts. The targeting method follows previous targeting experiments (see Langland (2005) for an overview) but the analysis is novel in the use of a limited-area rather than

global model, and 24km grid which is higher resolution than previous studies. It also uses a four-dimensional variational (4D-Var) data assimilation scheme to assimilate the targeted observations. Like previous targeting studies, this study is limited by the small number of cases; however, this study attempts to overcome this weakness by using a small ensemble of forecasts to assess the robustness of the results from the individual case studies.

The middle section of the thesis also deals with the targeted observations from the GFDex field campaign, investigating the potential benefit of the observations using one case study. This is investigated in two ways. Firstly, the errors attributed to the dropsonde observations during the assimilation are modified, having first calculated the dropsonde observation errors using GFDex dropsonde data. The set of observations is also varied, changing the spacing and the proximity of the targeted observations to the steep orography of Greenland. Secondly, it is investigated whether targeted observations show a small forecast impact (on average) because of the failure of the data assimilation scheme to extract important dynamical structures. This is investigated by the generation of pseudo targeted observations which contain a singular vector structure that is not contained in the background or routine observations, i.e. assuming that the background has an initial condition error with tilted growing structure.

The last section of the thesis analyses the impact of directly targeting a small dynamical feature, namely a polar low, in a manner that differs from conventional targeting. This novel approach to targeting identifies the sensitive region and then targets only the main dynamical feature within the verification region, that will have the greatest impact on the forecast of the verification region some time later. This is the first time that observations have been targeted in this region, and the first targeted observations of a polar low. The analysis of these targeted observations follows the GFDex analysis in using a limited-area model at high resolution.

The thesis aims can be condensed to form three succinct questions:

- What is the impact of targeted observations on the forecast for Northern Europe?

- How do targeted observations impact the forecast?

- Do targeted observations have the potential to have a large impact on the forecast?

This thesis attempts to answer these questions, and has the following structure. In Chapters 2 and 3 the background material needed to understand and interpret the work presented

in subsequent chapters is explained. In Chapter 2 an overview of the Met Office forecast system is given, with particular emphasis on the data assimilation scheme. In Chapter 3 an overview of previous targeting experiments is given, including idealised experiments, and the methods used to identify where to target observations are explained. In Chapters 4 to 7 the experiments conducted and their results are presented. In Chapter 4 the set-up of the GFDex targeting experiments is explained and the impact of the targeted observations on the forecasts for Northern Europe is analysed. This section of work has been accepted for publication (Irvine et al., 2009). Chapter 5 takes one case study from the GFDex field campaign and examines the potential to increase the forecast impact through the specific set of observations assimilated and the modification of the errors assigned to the observational data. In Chapter 6 the potential for targeted observations to impact a forecast is also examined, by testing whether the lack of impact is due to the inability of the data assimilation scheme to extract important structure from the targeted observations. Chapter 7 examines the impact of targeted observations of a polar low from the IPY-THORPEX field campaign on the forecast for Scandinavia, with particular focus on the time that the polar low made landfall. The main conclusions from this thesis and directions for future work are discussed in Chapter 8. The acronyms and abbreviations used in this thesis are listed in Appendix A.

CHAPTER 2

The Met Office Forecast System

This section describes the different components of the Met Office numerical weather prediction system, that are important in motivating the use of targeted observations and in understanding how targeted observations impact on a forecast. Figure 2.1 illustrates the different components of the Met Office forecast system in limited-area mode; this chapter describes the observation types (Section 2.1), observation processing (Section 2.2), data assimilation (Section 2.3) and forecast model (Section 2.4) components.
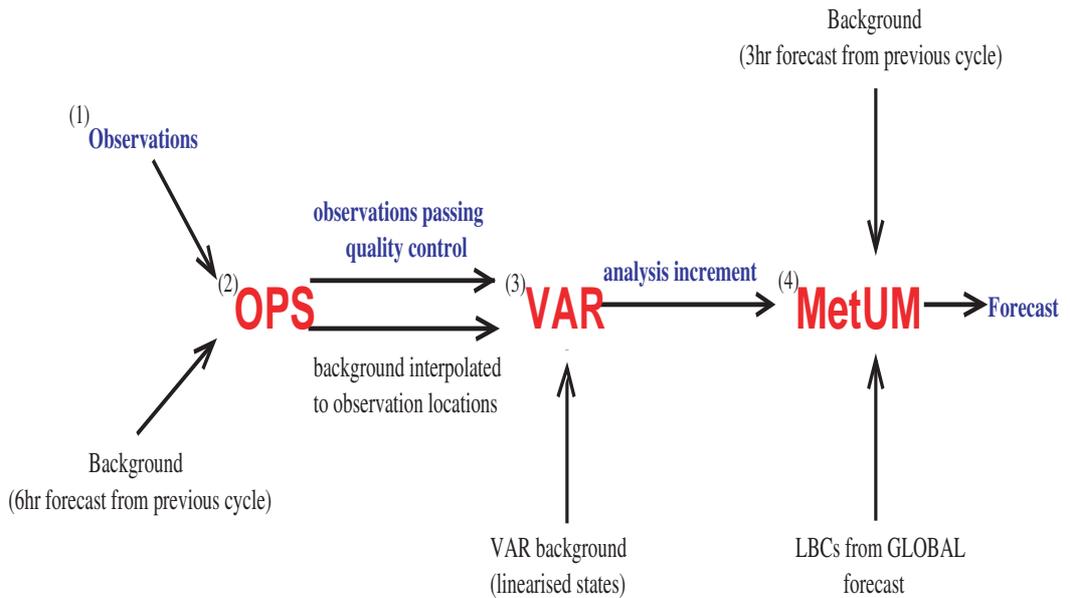
Figure 2.1: *The Met Office forecast cycle for a limited-area model. The main components of the cycle are highlighted in red. Observations are first processed by the observation processing system (OPS), then passed to the four-dimensional variational data assimilation scheme (VAR) before being used (along with lateral boundary conditions (LBCs)) by the Met Office Unified Model (MetUM) to make a forecast. The path that the observational data take is highlighted in blue. The numbers correspond to the four sections of this chapter.*

2.1 Observation Types

In order to ensure coverage of all areas of the globe, many different observation types are used in the forecast system. Together, these observation types make up the global observing system (GOS). The observations can be split into two categories, in-situ and remotely sensed. Observations which are made in-situ are for example sonde, aircraft and station observations. Instruments which make observations which are remotely sensed may be ground-based, for example wind profilers, or space-borne such as sounders producing radiance profiles.

2.1.1 In-situ Observations

Station observations provide information on surface conditions, such as pressure, wind and temperature. Such measurements are usually over populated areas of land, although the advent of automatic weather stations has enabled their deployment in remote uninhabitable regions such as Antarctica. Surface data is also received from ships and buoys, extending the area covered by surface observations.

Some commercial aircraft are fitted with instruments that measure the air temperature, wind direction and speed at the aircraft flight level, and also provide profiles as the aircraft takes off and lands at airports. The accuracy of these measurements has been shown to be as good as radiosonde observations (Moninger et al., 2003). Some aircraft now carry advanced systems which allow them to measure and report humidity. There are two main types of reports from aircraft: automatic reports called aircraft meteorological data relay (AMDAR), and manual reports called Aircraft Reports (AIREPS). AMDAR are continuous reports from the aircraft along their flight track, whereas AIREPS are received at intervals of 10° longitude over the oceans. Note that some American aircraft report via a different system called Aircraft Communications Addressing and Reporting System (ACARS). There is good data coverage of AMDAR reports over the USA and Europe and above the oceans along major intercontinental aircraft routes (Figure 2.2(b)). The temporal coverage however is highly variable due to the flight schedules (i.e. more flights during the day than at night), and much less data is received when inclement weather (e.g. a snowstorm) prevents aircraft from taking-off and landing.

Radiosonde ascents provide profiles of temperature, pressure, wind direction and speed and relative humidity. The balloons are generally released from land-based stations with each station providing between two and four ascents each day. They are most concentrated in

(a)

(b)

(c)
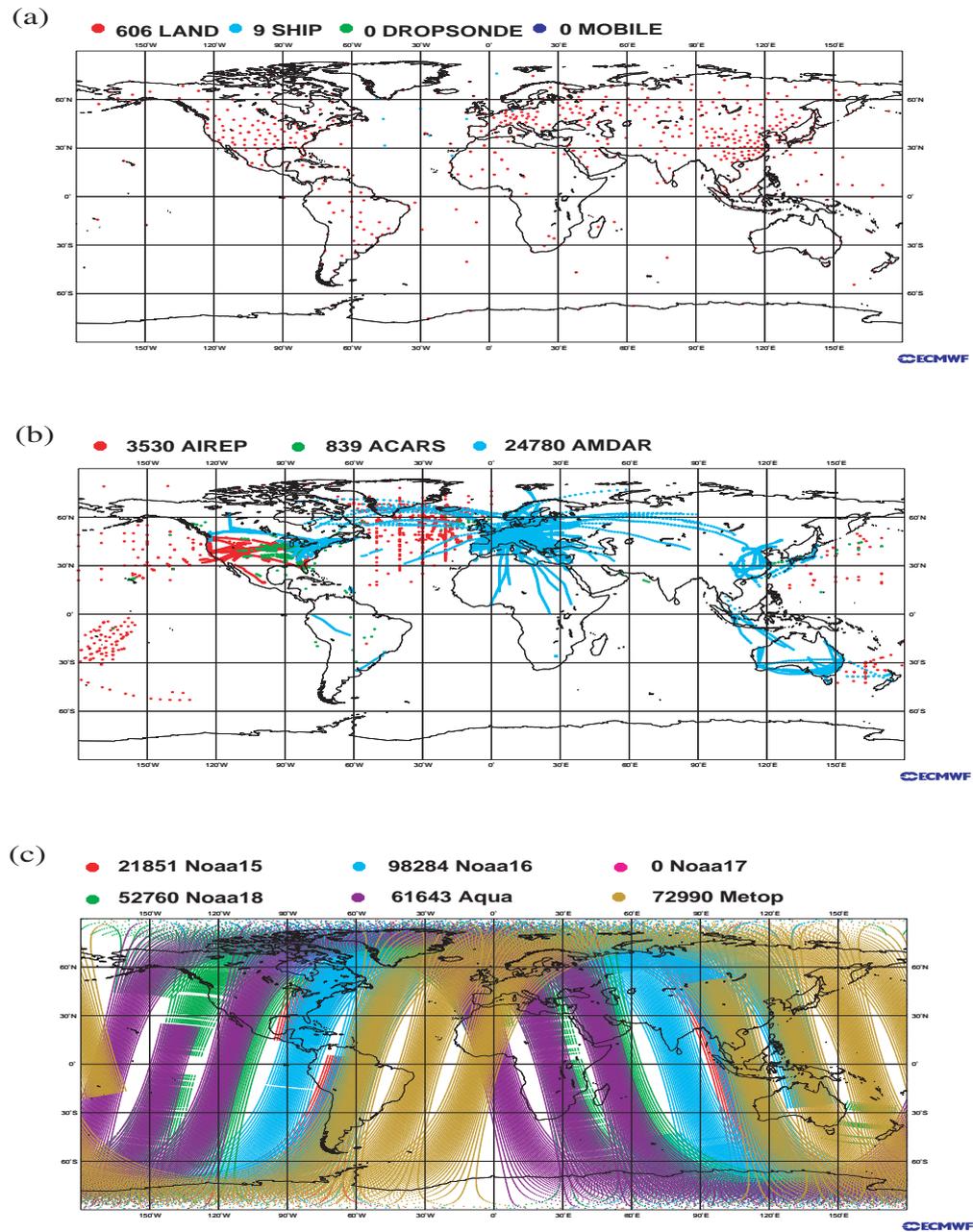
Figure 2.2: *Data coverage plots from ECMWF showing observations that were available for assimilation into the 1200 UTC forecast on a day in September 2009 for (a) Radiosonde observations released from land, ships or mobile units, (b) aircraft observations from AIREP, AMDAR and ACARS and (c) Advanced Microwave Sounding Unit-A (AMSU-A) satellite observations from various satellites. Above each panel the number of available observations is given.*

densely populated areas of the world, with few to no ascents over sparsely populated areas and areas of political conflict such as large regions of Africa (Figure 2.2(a)). There are a few ships which release radiosondes, but these are not stationary so the location of the ascent varies from day to day as the ship travels along its shipping route.

Focussing on the North Atlantic region (the region that this thesis is concerned with), around Iceland and Greenland, and up into the Arctic, the data coverage of in-situ observations is poor. There are a few land-based radiosondes scattered on Greenland, Iceland, Norway and islands in the seas surrounding these land-masses (Figure 2.2(a)). There are some AMDAR measurements around Greenland and Iceland, which lie under the Europe-North America flight track taken by commercial aircraft. In the Norwegian and Barents seas the sole AMDAR measurement comes from flights between the Norwegian mainland and Longyearbyen on Spitsbergen (Figure 2.2(b)).

Observing-system experiments (OSEs) compare forecasts containing different components of the GOS with analyses containing all components of the GOS, to deduce the relative importance of each type of observation. Graham et al. (2000) analysed the contribution of different types of conventional observations to 20 cases where forecasts of a particular synoptic event were enhanced by observations made prior to the initialisation of a 60-hour forecast of that event. This found the radiosonde network to be most important overall, although for weather systems developing over the North Atlantic aircraft winds were more important than radiosonde observations. This is in contrast to the findings of Bouttier and Kelly (2001) and Dumelow (2008) who found that radiosonde observations were most important to the forecast skill over Europe. Although in places sparse, radiosonde data is important because it provides a profile of data, which has high vertical resolution and small errors which are well-characterised. The relative importance of aircraft data is smaller because it provides data at a single level only (except near major airports). The same is true for station observations.

2.1.2 Remotely-sensed Observations

Wind profilers use radar to provide vertical profiles of wind, over populated areas of land (data coverage not shown). They can degrade the forecast in conditions of strong convection in the boundary layer (Cardinali, 2009), however the profiles are important to the forecast skill, particularly above 400hPa (Graham et al., 2000).

Putting meteorological instruments on satellites that can remotely-sense the Earth's atmosphere has helped to greatly increase the data coverage of the GOS. Figure 2.2(c) shows data coverage of AMSU-A observations for one forecast. A data coverage plot of ATOVS (Advanced TIROS (Television and Infrared Observational Satellite) Operational Vertical Sounder) would look similar. The satellites that the instruments are mounted on are polar orbiting; one orbit of the Earth takes approximately 90 minutes and provides a wide-swathe of observations. The ATOVS instruments measure the radiation emitted by the atmosphere, from which a temperature profile can be derived. An obvious problem with this method is that clouds emit infra-red radiation as black-bodies with an energy proportional to the temperature of the atmosphere at the top of the cloud. The instruments cannot see through the cloud, therefore profiles are obtained only as far down as the cloud top height. The error in the temperature profiles is larger than for radiosonde data, as the temperature is not measured directly but derived from radiance profiles, making assumptions about the profile in order to do this.

Space-borne instruments are an important component to the GOS. Over sparsely-populated or inaccessible areas of land or over the oceans they are often the only source of observational data with which to constrain the analysis. Their inclusion in the GOS has resulted in large increases in forecast skill, particularly for the Southern Hemisphere (Simmons and Hollingsworth, 2002). Bouttier and Kelly (2001) found that over data-sparse regions satellite data was the main contributor to forecast skill, with an equivalent contribution to that made by aircraft and radiosonde data over data-rich areas.

In the region around Greenland and Norway, there is good data coverage from satellite observations. However this region has large temperature gradients at low levels, associated with the transition from ice/snow-covered land or sea-ice to open ocean which is much warmer. Therefore these regions are often baroclinically unstable and so are cloudy. This means that satellite observations are useful only to constrain the atmospheric profile above the cloud top, and often provide little to no information on the profile in the lower atmosphere.

2.2 Observation Processing

Observational data are sent onto the Global Telecommunications System (GTS), and received at forecast centres around the world. In order for data to be eligible for use in the 1200

UTC forecast, the measurement must be taken between 0900 UTC and 1500 UTC. In practice the cut-off time for data to be received by the Met Office for inclusion in the 1200 UTC global forecast is 1345 UTC to allow time for the forecasts to be run and forecast products to be disseminated to the end user. In order to ensure that the best forecast is available to use as the background for the subsequent forecast cycle, an update cycle is run after the operational forecast, which includes all observations that were taken within the data assimilation window but not necessarily received in time to be used in the operational forecast.

Before the observations can be used by the data assimilation scheme, the data must first be quality-controlled to remove any observations that have a high probability of being in error. The error could be caused by an instrument malfunction or a human error in reading or transcribing the observation. The data assimilation scheme uses Gaussian error statistics, thus assuming that observations whose errors do not fit this assumption have been removed. At the start of the processing, observations are assigned an initial probability of being in gross error (PGE). For radiosonde observations, the initial PGEs (Table 2.1) are 0.015 for pressure, temperature and relative humidity observations and 0.005 for wind observations (Ingleby and Lorenc, 1998). A series of checks are then made, during which the PGE is incremented if the observation fails a check. At the end of the quality control process the observation is then rejected if the PGE >0.5. The variables which are measured are not necessarily those needed for the assimilation, nor are they necessarily on the same levels that the model uses. The second purpose of the observation processing scheme is then to transform the observations into the correct format for the assimilation. The dropsondes used to make targeted observations are treated as radiosonde observations during the observation processing stage of the forecast. This section will therefore focus on the processing of a radiosonde observation in the Met Office forecast suite.

Here an overview of the probability theory used in the quality control is given, following Lorenc and Hammon (1988) and Ingleby and Lorenc (1998). The Met Office observation processing scheme is based on Bayesian theory. Bayes' theorem gives the conditional probability that an event A will occur, given that event B has occurred:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \qquad\qquad (2.1)$$

where P(A) is the probability of event A occurring, and P(B) is the probability of event B occurring, and P(A|B)=P(A∩B)/P(B) (i.e. the probability of A occurring given that B has occurred is the probability of A *and* B occurring divided by the probability of B occurring).

In order to use this theory to process observations, we have to make some assumptions. Firstly, we assume that the forecast background has Gaussian errors with variance V_b. Good observations are assumed to also have Gaussian errors with variance V_o. The probability density function N for the Gaussian distribution with mean y and variance V is then:

$$N(y, V) = (2\pi V)^{-0.5} e^{\frac{-y^2}{2V}}, \tag{2.2}$$

Observations in gross error have uniform density k. The probability that an observation is in the range y_o to $y_o + dy_o$ given the background value of y_b, $P(O)$, is then:

$$P(O) = P(O|\overline{G})P(\overline{G}) + P(O|G)P(G) = N(y_o|y_b, V_o + V_b)P(\overline{G}) + kP(G)dy_o, \tag{2.3}$$

where the subscripts o and b refer to observation and background respectively, and y_0 is normally distributed with mean y_b. $P(G)$ is the initial probability of gross error (initial PGE) in an observation and $P(\overline{G})$ is the probability that the observation is not in gross error. The variance of the observation errors V_0 can be estimated by calculating the root mean square of the observation minus background differences. $P(G)$ is estimated by using previously quality controlled data to find the occurrence of bad data. The density of the observations in gross error can be calculated the same way, or empirically. The k used in the current observation processing system are listed in Table 2.1, and are taken from Ingleby and Lorenc (1998). The new PGE after the background check is $P(G|O)$, i.e. the probability of gross error given the given the background value, and is calculated as:

$$P(G|O) = \frac{P(O|G)P(G)}{P(O)} = \frac{kP(G)}{v(y_o|y_b, V)(1 - P(G)) + kP(G)} \tag{2.4}$$

where V is the variance of the observation minus background errors.

The processing applied to radiosonde data is now described, following Ingleby (1998). The radiosonde (or dropsonde) data that is sent to the GTS consists of pressure, drybulb and dewpoint temperature and wind direction and speed on standard and significant levels. Note

Table 2.1: *Probabilities Applied to Radiosonde Data*

Variable	Initial PGE	k
Pressure	0.015	0.043 hPa^{-1}
Temperature	0.015	0.01 K^{-1}
Wind	0.005	$0.00005 \text{ m}^{-1}\text{s}$
Relative Humidity	0.015	$0.01 \text{ \%}^{-1}$
Geopotential Height	0.015	0.05 m^{-1}

that radiosondes measure relative humidity, but report dewpoint temperature. The standard levels used for reporting radiosonde or dropsonde information are: 1000, 925, 850, 700, 500, 400 and 300hPa (note that we did not release dropsondes from above 300hPa). Data is also reported at the surface and tropopause levels (if relevant) and at significant wind and temperature levels (defined as turning points) including the level of maximum wind speed. Reporting data this way minimises the amount of data to be encoded and sent to the GTS, whilst ensuring that there is redundancy in the information reported, i.e. the complete profile can be reconstructed from the significant levels without requiring the standard level data.

The first step in the processing is to convert the reported variables into the variables used by the model. Thus the dewpoint temperature is converted to relative humidity, and the wind speed and direction into wind components. The dependence of the calculated relative humidity on the temperature measurement means that if a temperature measurement at a particular level fails the observation processing checks then the relative humidity measurement at that level is automatically failed.

Next follows a series of checks to verify that the observations within a profile are consistent with each other. For the wind data this means checking that different wind speeds are not reported at the same pressure (which would imply a malfunction in the pressure sensor), and also re-computing the standard-level data from significant-level data. If the re-computed standard level data differs from the reported standard level data by $>8\text{ms}^{-1}$ then the standard pressure level and significant levels immediately above and below the standard pressure level have their PGE incremented. The same two checks are applied to temperature data. A common error in reporting temperature is attributing the wrong sign to the measurement (i.e.

making a positive temperature negative). This is checked by differencing the temperature from the background; if the difference is >20°C and changing the sign of the observation reduces the difference to 5°C or less then this is done. A check is also performed for super-adiabatic lapse rates more than 50hPa above the surface. Radiosonde temperature measurements are biased due to heating by solar radiation during the day, and cooling at night. No radiation corrections are performed for the Vaisala sondes, as the biases are removed before the data is sent to the GTS.

The data must then be converted to the same levels as used by the model. This allows for the data to be checked against the background fields, and against each other in the final phase of the observation processing, but is obviously required to be done before the data is made available to the assimilation scheme. The radiosonde data is reported on many more levels than are used in the model (the number of significant levels is large), which means that the radiosonde profile contains structure that the model is not able to represent, introducing representivity errors. The observational data is therefore averaged over the model layers, which results in smaller representivity errors than for the reported data, as there is some cancellation of errors. The only variable which is not averaged is relative humidity, which is instead optimally interpolated to model layers as averaging such information reduces saturation and amount of cloud predicted (Lorenc, 2007). Humidity observations near 100% are also increased by a factor to correct for a bias in the radiosonde instruments although it has been shown that this correction factor is now not needed and that the corrections could be stopped (Lorenc, 2007).

After the observations have been converted onto model levels, the background and buddy checks are performed. The averaged values are compared to the background values, and the PGE re-computed using (2.4). For the temperature check, the background error is multiplied by 1.5 at 300hPa and above, and by 1.2 below, to ensure that the check is not too strict around the tropopause level and does not erroneously reject good temperature observations. The accuracy of the background is dependent on synoptic features for example greater errors occur in the region of a rapidly or explosively developing cyclone. In cases like these, the check against the background would reveal a large departure from the background and the PGE would be set higher which could result in the observation being rejected if there were no further checks (Lorenc and Hammon, 1988). To prevent this, a final check is carried out,

which is a procedure called 'buddy checking.' This compares the departure of an observation from the background with that of up to 12 closest observations of similar type. Therefore dropsonde observations could be buddy-checked with nearby radiosonde observations, but would not generally be buddy-checked with each other as they have the same call-sign. The point is not to see how close a match the two sets of observations are, but how close a match the departure from the background field is. The buddy check is useful to show where it is the background not the observation which is incorrect.

2.3   Data Assimilation

We seek via data assimilation to find the best guess of the current state of the atmosphere through optimal blending of a prior forecast, known as the background state, and current observations. This best guess of the current state is the analysis. The ability of the data assimilation scheme to produce an accurate analysis is an important factor to consider when analysing the impact of targeted observations. This chapter describes the set-up of two configurations of the variational data assimilation scheme currently used by the Met Office with a particular focus on the representation of observation and background error. The three dimensional variational (3D-Var) scheme (Lorenc et al., 2000) is used by the Crisis Area Mesoscale Model (CAMM), and was used in the hindcast runs for the Norwegian targeting experiment described in Chapter 7. The global and limited-area north-Atlantic European (NAE) models now use a 4D-Var scheme. The 4D-Var scheme (Rawlins et al., 2007) is an extension of the 3D-Var scheme, in which both the time and location of the observations are taken into account during the assimilation process. This scheme, as it is described, was operational for the NAE model in Spring 2007, and used in the hindcast experiments for the GFDex targeting experiments, described in Chapters 4-6.

Figure 2.3 is a schematic illustration of how the 4D-Var scheme works. The aim is to minimise a cost function with background and observation penalty terms to produce an analysis increment $\delta\mathbf{x}$ that is added to the background state when the forecast is initialised. When the forecast is integrated forward, it produces a trajectory which best fits the observations and background through the data assimilation window. The 3D-Var scheme differs because it considers all observations at a single time only and does not consider the evolution of the background state. Table 2.2 summarises some of the key differences between the 3D-Var and

4D-Var (where times are given they relate to a 1200 UTC forecast). The data assimilation window is 6 hours, and the background state is a 6 hour forecast from the previous model run for 4D-Var and 9 hour forecast for 3D-Var (Table 2.2). There are two main components to the 4D-Var system: an outer loop where the full non-linear MetUM is run to provide the background fields needed for the inner loop, and an inner loop where the linearised model is run at lower resolution and the cost function minimised via a descent algorithm to give the analysis increments needed to initialise the forecasts.
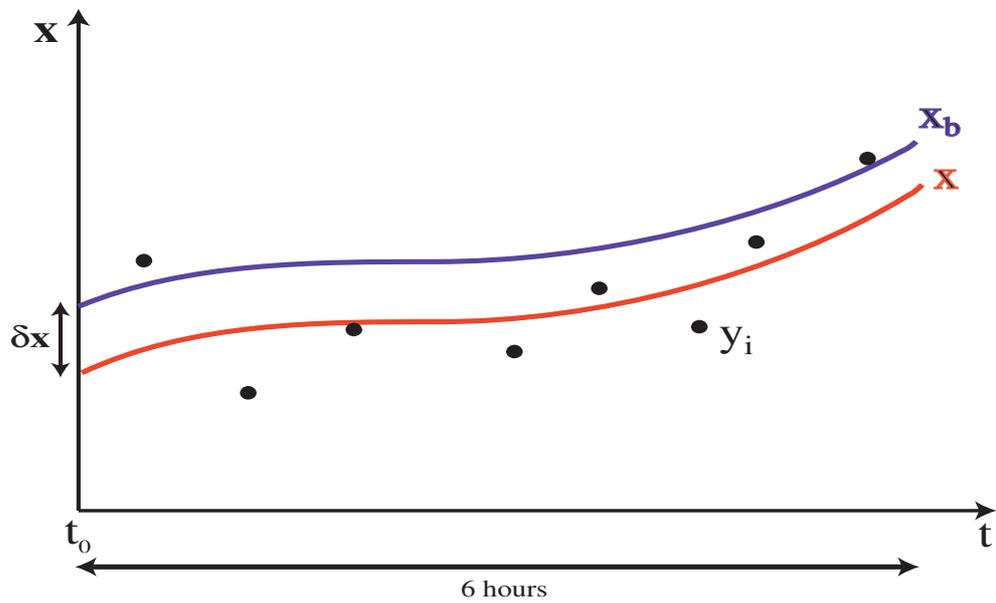
Figure 2.3: *Schematic illustrating 4D-Var, showing the background state x_b evolution (blue line), the guess state x (red line) and observations y (dots). If x minimises the distance to x_b and y (weighted by the error covariances over the assimilation time window) then δx is the analysis increment.*

Variational data assimilation schemes use a least-squares approach to find a state which minimises the distance to the model background and to the observations. This leads to a cost function, J, which has the following form:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathbf{h(x)})^T (\mathbf{y} - \mathbf{h(x)}), \qquad (2.5)$$

where $\mathbf{x}$ is the model state for which J is minimised, $\mathbf{x}^b$ is the background state, $\mathbf{y}$ is the vector of observations and the nonlinear operator $\mathbf{h}$ maps from model to observation space. In matrix notation, $\mathbf{AA}^T$ is equivalent to $\mathbf{A}^2$; therefore this is a least-squares approximation.

Table 2.2: *Attributes of 3D-Var and 4D-Var data assimilation schemes for a forecast from 1200 UTC*

Attribute	3D-Var	4D-Var
Dimensions	x, y, z	x, y, z, t
Analysis Time	1200 UTC	0900 UTC
Observation-background departures	at 1200 UTC	model timestep nearest observation time
Number of linearisation states	1	1 per timestep
Perturbation forecast model?	NO	YES
Forecast error covariance	Static	Allowed to vary

This form of *J* gives equal weighting to the observations and the background state, assuming that both states are equally likely. However we know that this is not a good assumption; both the model background and observations have errors. Incorporating the background and observation errors, we obtain the cost function for 3D-Var:

$$J_{3DVAR}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathbf{h}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}(\mathbf{x})), \qquad (2.6)$$

where **B** is the background error covariance matrix and **R** the observation error covariance matrix. The first term in the above equation calculates the penalty to the forecast background state, weighted by the inverse of the background errors. The second term calculates the penalty to the observations, weighted by the inverse of the observation errors. The inverse of the errors is used as a weighting so that larger weighting is given to an observation or background state that is well-known, i.e. has a smaller error.

4D-Var takes into account not just the position of an observation but the approximate time at which it was made. Thus we add an additional level of complexity to (2.6) (Courtier et al., 1994):

$$J_{4DVAR}(\mathbf{x}(t_0)) = \frac{1}{2}(\mathbf{x}(t_0) - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}(t_0) - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - \mathbf{h}_i(\mathbf{x}(t_i)))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{h}_i(\mathbf{x}(t_i))), \quad (2.7)$$

where the penalty to the observations is evaluated at the timestep *i* closest to the observation, where *N* is the number of timesteps. Here the model estimate of the observation at the

timestep closest to the time of the observation is obtained by propagating the initial state $\mathbf{x}(t_0)$ forward in time:

$$\mathbf{x}(t_i) = \mathbf{M}\mathbf{x}(t_0). \tag{2.8}$$

$\mathbf{M}$ is the perturbation forecast model, which in this set-up is a linearised version (that is not the tangent) of the full non-linear MetUM. During the previous model run the model is forced to write out forecast fields at additional timesteps to provide the background fields, called linearisation states, which are required for the data assimilation for the subsequent forecast. This previous model run is therefore the 'outer loop' in the data assimilation scheme. The 4D-Var uses one linearisation state for each timestep instead of the one used at initial time by the 3D-Var scheme. The background fields are interpolated in space to each observation position, and the field at the timestep closest to the time of the observation is used, producing a vertical profile of model data at each observation point. The 3D-Var and 4D-Var formulations in operation at the Met Office use an incremental approach to facilitate the minimisation. Instead of minimising J with respect to a model state $\mathbf{x}$, they minimise J with respect to an increment to the initial state (Courtier et al., 1994). In this formulation the background field is used as the first guess state, i.e. $\mathbf{x}^g = \mathbf{x}^b$, and J minimised to find an increment $\delta\mathbf{x}$. The next guess state is then the previous guess state plus the increment. The cost function for incremental 4D-Var is then:

$$J(\delta\mathbf{w}) = \frac{1}{2}(\delta\mathbf{w} - \delta\mathbf{w}^b)^T \mathbf{B}^{-1}(\delta\mathbf{w} - \delta\mathbf{w}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - \mathbf{y}_i^o)^T \mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{y}_i^o), \tag{2.9}$$

where $\delta\mathbf{w}$ is the increment to the guess state, and $\delta\mathbf{w}^b$ the background minus guess state, both interpolated to a lower resolution grid than $\mathbf{x}$, $\mathbf{x}^g$ and $\mathbf{x}^b$. $\mathbf{y}^o$ are the model's estimate of the observations, now a function of $\delta\mathbf{w}$ rather than $\mathbf{x}$. For the model runs in this thesis, the grid-spacing was 48km, half the grid spacing for the forecast model (which was 24km).

In order to deal with the above equation, we have to make some assumptions. We make the simplifying assumption that the errors between different observations are uncorrelated, which reduces $\mathbf{R}$ to a diagonal matrix. The structure of the matrix is shown in Figure 2.4(a) for a system with four variables u, v, T and p at two locations. The squares on the diagonal contain variances, the blank squares are zero elements. It can be seen that only the only non-

zero elements are the variances of each observation with itself at the same point. We would ideally like **B** to have the same structure. However the model variables are clearly correlated with each other, e.g. stronger wind where the pressure gradient is greater, therefore the same assumptions cannot be made for **B**. This creates a problem; **B** has dimensions of the state vector (i.e. 10^{14} elements) and is too large to store. The problem is solved by using a control variable transform, transforming from model variables into a set of control variables, chosen so that they are uncorrelated with each other, allowing us to indirectly model a simplified form of **B**.

Figure 2.4: *The structure of the (a) observation and (b) background error covariance matrices for a four-variable system on a 2×2 grid. Matrix (a) is diagonal, where the diagonal elements are the variances σ^2 of each variable. Matrix (b) is block diagonal, where the diagonal elements are the variances σ^2, and the off-diagonal elements σ^2_{12} are autocovariances.*

The model variables are transformed into a vector of control variables **v** by the transformation **U**. **U** contains not only the simple parameter transform, but also vertical and horizontal transforms.

$$\delta\mathbf{w} = \mathbf{Uv}, \tag{2.10}$$

where the transformation **U** is defined as the square root of **B**

$$\mathbf{B} = \mathbf{UU}^T. \tag{2.11}$$

The cost function can then be re-written to get the background term in terms of the control variables by substituting (2.10) and (2.11) into (2.9):

$$J(\mathbf{v}) = \frac{1}{2}(\mathbf{v} - \mathbf{v}^b)^T(\mathbf{v} - \mathbf{v}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - \mathbf{y}_i^o)^T\mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{y}_i^o). \quad (2.12)$$

Note that the first term now contains no references to **B**. **B** is now indirectly modelled via the control variable transform.

The control variable transform is an important part of the 4D-Var system, as it allows us to solve an otherwise intractable problem. The control variables used in the Met Office VAR system are streamfunction ψ, velocity potential (divergent wind) χ, unbalanced pressure (deviation from hydrostatically balanced pressure) p_u and relative humidity μ (Lorenc et al., 2000). Streamfunction is used as the balanced variable. The velocity potential and unbalanced pressure are residual fields, so that they are uncorrelated with each other and with the streamfunction field. The relative humidity is used as the control variable for moisture, because it has a smaller correlation than specific humidity with the temperature field (Lorenc et al., 2003).

The control variable transforms remove multivariate covariances from **B**, through the choice of control variables which are uncorrelated with each other, and also introduces balance. The structure of **B** is shown in Figure 2.4(b). The matrix is now block diagonal; the diagonal elements are the variance of the variable with itself at the same location, and the off-diagonal elements are the autocovariances (i.e. the variance of a given variable at one point with itself at another point). Therefore in this formulation **B** is not diagonal as it still contains the autocovariances of the control variables. These are dealt with by applying spatial transforms in such a way that the autocorrelation of a field can be modelled using statistics.

Accurate representation of the background error covariance matrix is essential because it is this that controls the spread of information from the observations. This can be seen by expressing the analysis increment $\delta\mathbf{x}$ as the best linear unbiased estimate, BLUE (Kalnay, 2003):

$$\delta\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}(\mathbf{y} - \mathbf{h}(\mathbf{x}^b)), \quad (2.13)$$

where $\mathbf{x}^a$ is the analysis and $\mathbf{H}$ is a linearised form of $\mathbf{h}$. It is clear from (2.13) that the size of the analysis increment is proportional to the departure of the observation from the background field. The observation and background errors act as a weighting. As $\mathbf{B}$ is the last operator to act on the analysis increment, it controls the spreading of the increment both in space and between variables. It is therefore important that the error correlations which make up $\mathbf{B}$ are correctly specified, so that they do not lead to anomalous spreading of information. Following Bannister (2008) the spreading of data between gridpoints by $\mathbf{B}$ can be demonstrated by considering the case where a single observation y is assimilated at point k. For a single observation, the analysis increment at point l reduces to:

$$x_l{}^a - x_l{}^b = B_{lk} \frac{y - x_k{}^b}{B_{kk} + \sigma^2} \tag{2.14}$$

where σ^2 is the variance of the observation, B_{kk} is the forecast error covariance at the position of the observation, B_{lk} is the autocovariance of the forecast error between points k and l, and the subscripts k and l denote position.

The $\mathbf{B}$ used in 4D-Var is not properly flow-dependent. If $\mathbf{B}$ were properly flow-dependent then the 4D-Var would be equivalent to the Kalman filter (described in Section 3.1). $\mathbf{B}$ is allowed to evolve with time during the assimilation window, but is reset back to its original climatological values before a subsequent iteration. This quasi flow-dependence is one of the advantages of 4D-Var over 3D-Var, which uses static error covariances as it does not consider the time dimension in the minimisation of the cost function. The error correlations which are used in the implicit representation of $\mathbf{B}$ are calculated using the National Meteorological Center (NMC) method (Parrish and Derber, 1992). The NMC method takes the difference between a 24-hour and 48-hour forecast valid for the same forecast time (i.e. initialised 24 hours apart) for a large set of forecasts, and computes the error covariances:

$$\mathbf{B} \approx \frac{1}{2} \langle (\mathbf{x}^{48} - \mathbf{x}^{24})(\mathbf{x}^{48} - \mathbf{x}^{24})^T \rangle, \tag{2.15}$$

where the superscripts 48 and 24 refer to the forecast times, and the angled brackets denote an average over a month of data (Bannister, 2008). Ingleby (2001) calculated $\mathbf{B}$ using the NMC method, and compared it to the implied covariances used in the 3D-Var scheme at the Met Office. The implied covariances well represented the main features of the covariances

calculated from forecast differences, like the dominant barotropic mode in the extra-tropics, although the correlation length-scales were shorter.

The output from both VAR systems is an analysis increment. The analysis increment is added to the background state when the model is initialised. For the 1200 UTC forecast run, this time corresponds to 0900 UTC for the 4D-Var analysis increment, and 1200 UTC for the 3D-Var analysis increment, i.e. the 4D-Var analysis increment is added at the beginning of the time window, and the 3D-Var analysis increment is added in the middle of the time window. This means that what is considered the 1200 UTC analysis from the 4D-Var scheme is actually a three-hour forecast.

2.4   Forecast Model

The MetUM, version 6.1, is used as the forecast model in all the experiments presented in this thesis. The model is non-hydrostatic and fully compressible, with a semi-implicit, semi-Lagrangian numerical scheme, and uses the new dynamics formulation as its dynamical core (see Davies et al. (2005) for a detailed description). This includes the boundary-layer parametrization of Lock et al. (2000), a mass-flux convective parametrization (Gregory and Rowntree, 1990) and a mixed-phase cloud mircophysics scheme that is coupled to a short-wave radiation scheme (Wilson and Ballard, 1999). A staggered Arakawa-C grid (Arakawa and Lamb, 1977) is used in the horizontal, and for the experiments in this thesis a grid-spacing of 24km was used. A staggered Charney-Phillips grid is used in the vertical, and the model was run with 38 levels on a stretched grid, with the model top at 10hPa. The vertical co-ordinate is height-based and is terrain-following near the ground.

The MetUM was run in limited-area mode, with lateral boundary conditions provided by global forecast runs at a lower resolution. In limited-area mode the grid is rotated such that the Equator lies within the limited-area domain, which achieves a more horizontally uniform grid. Two different limited-area domains were used for the experiments performed in this thesis. The GFDex targeting experiments used the standard NAE domain (Figure 2.5) that was operational at the time of the GFDex field campaign in 2007. For the Norwegian targeting experiments this domain was not suitable, so the CAMM was used instead with a new limited-area domain specified (Figure 2.5). The CAMM system is designed so that a new domain can be put over any region of the globe. It has a similar set-up to the NAE; it takes

its lateral boundary conditions from a global forecast and can run four times daily. It differs to the operational NAE in that it uses 3D-Var instead of 4D-Var for the assimilation. Unlike the operational NAE model which has a continuous forecast cycle, the CAMM was not run operationally during the Norwegian field campaign. Therefore the initial start dump for the CAMM came from the global model (for the first analysis cycle; a short forecast from this cycle was then used as the background for the subsequent cycle).
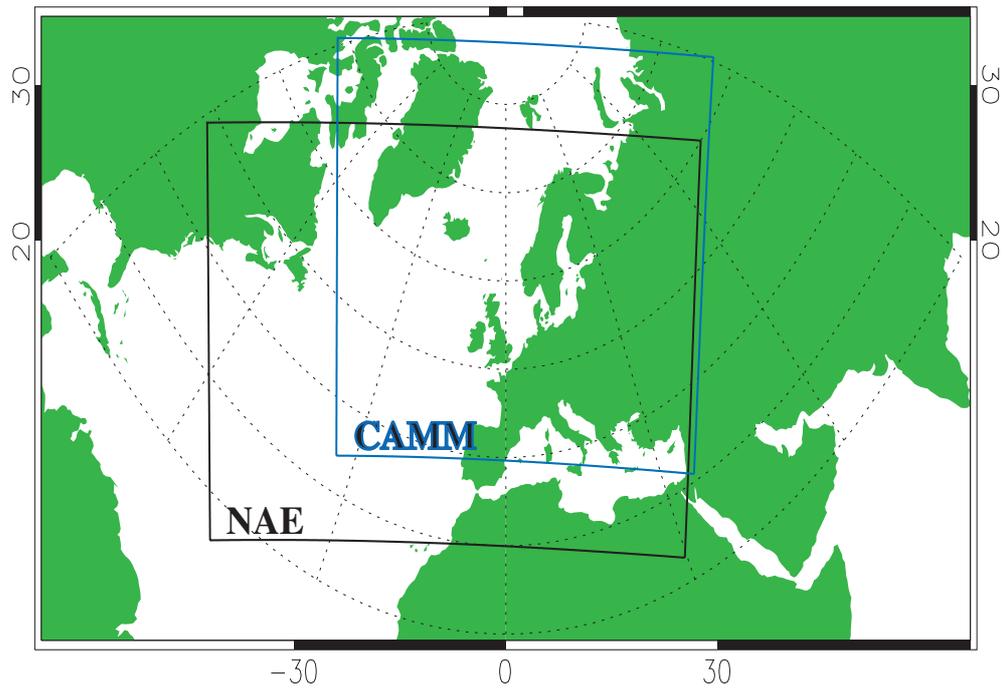

Figure 2.5: *The limited-area domains used in this thesis: the standard operational NAE domain (black box) used for all GFDex targeting experiments, and the CAMM domain (blue box) designed specifically for the Norwegian targeting experiments (i.e. non-operational).*

CHAPTER 3

Background: Targeted Observations

This chapter gives a general background to targeted observations. The sensitivity methods used to identify the region where the targeted observations should be made are detailed in Section 3.1. Section 3.2 looks at the different types of observations which have been targeted. The results from idealised modelling studies of targeted observations and previous trials of targeted observations are described in Sections 3.3 and 3.4 respectively.

3.1 Sensitive Area Predictions

An important part of the observation targeting process is defining the *target region*, where the additional observations will be made. Sensitive area prediction (SAP) plots are used to determine this. They predict where the forecast for a specified *verification region* will be most sensitive to errors in the initial conditions for a specific *optimisation time*, i.e. the time interval between performing the targeting and verifying the forecast (Figure 3.1). Here the word sensitive is used to mean that a large forecast error may result in the verification region at the verification time as a result of an initial error elsewhere in the forecast domain. To identify this sensitive region, two objective methods were used during the GFDex: total energy singular vectors (TESV; Buizza and Montani (1999)) and the ensemble transform Kalman filter (ETKF; Bishop et al. (2001)).

3.1.1 Singular Vectors

A brief definition of singular vectors (SV) are given here, following that given in Buizza and Palmer (1995). The reader is referred to this paper for a more detailed description. A simple approximation to a numerical weather prediction model can be given by:

$$\frac{d\mathbf{x}}{dt} = A[\mathbf{x}], \tag{3.1}$$
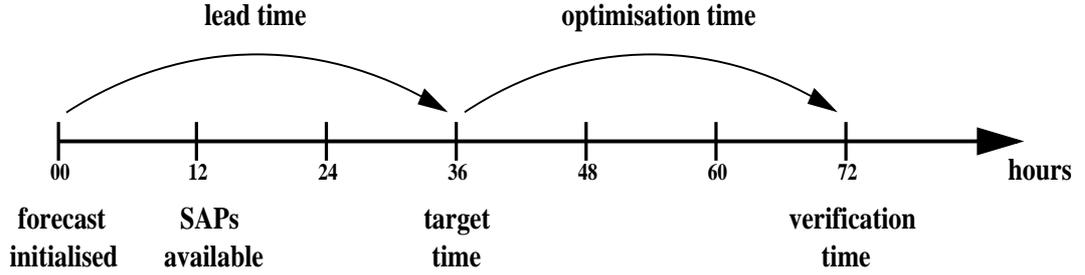
Figure 3.1: *An illustration of the timeline involved in making targeted observations. In this example the lead time is 36 hours and the optimisation time for the sensitive area predictions is also 36 hours.*

where **x** is the state vector containing information about the model variables that describes the state of the atmosphere at a given point in time, and A is an operator which describes the evolution of **x** throughout the forecast. For the purposes of targeting, we are interested in the growth of initial errors, which can be thought of as perturbations to the state vector at initial time, $\mathbf{x}'$. The evolution of these perturbations from (initial) time t_0 to time t can be described using

$$\mathbf{x}'_t = \mathbf{L}\mathbf{x}'_{t_0}, \tag{3.2}$$

where **L** is a matrix operator called the forward tangent propagator (it propagates the perturbations forward in time through the forecast) of the discrete linearised model.

The adjoint of **L** calculated with respect to the total energy norm is $\mathbf{L}^{*E}$. The singular vectors can be calculated via a singular value decomposition, which allows the factorisation of a matrix in terms of its singular vectors and singular values. The eigenvalue problem that we are trying to solve is of the form $\mathbf{Kr} = \sigma\mathbf{r}$ where **r** are the eigenvectors of **K** and σ the eigenvalues. **r** are a set of orthonormal vectors, called singular vectors, and their magnitude is described by σ, the singular values. **K** is defined such that **r** have maximum total energy within the verification region at final time:

$$\mathbf{K} = \mathbf{E}^{-0.5}\mathbf{L}^{*E}\mathbf{T}\mathbf{L}\mathbf{E}^{-0.5}, \tag{3.3}$$

where **E** is the inner product of the total energy norm. **T** is a local projection operator which allows us to find singular vectors with maximum amplitude within the verification region (i.e. **T** has value 1 inside the verification region and 0 outside).

The largest singular value is associated with the leading singular vector, and subsequent singular vectors have singular values of decreasing amplitude. In order to obtain a good estimate of the sensitivity of the atmosphere, Buizza and Montani (1999) showed that it was sufficient to use the leading four to ten singular vectors to define the target area, and that there was little difference in the target areas obtained. To obtain the target area from the singular vectors, Buizza and Montani (1999) took a vertical average of the total energy of each initial singular vector, f_i^E (where i is the number of the SV), is taken at each point s, weighted by the ratio of the singular value, σ to the leading singular value, σ_1 and summed over the N singular vectors:

$$F^E(s) = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_1} f_i^E(s). \tag{3.4}$$

This function F^E has a maximum at point s_0, and the target area is then defined where:

$$F^E(s) \geq 0.5 F^E(s_0), \tag{3.5}$$

where the factor of 0.5 appears to be a good compromise between identifying too large an area (for a smaller factor) or not identifying a large enough area (for a larger factor) (Buizza and Montani, 1999). The TESVs therefore identify the fastest growing perturbations over the optimisation time that would impact the verification region, where the growth is measured using a total energy norm. This is an appropriate metric to use for targeting purposes as it provides a good approximation to the analysis error covariance matrix used by data assimilation schemes (Palmer et al., 1998).

Another norm that has been used for targeting is the Hessian of the cost function J, for 4D-Var (2.7). The Hessian is the second derivative of J (with respect to $\mathbf{x}$), and is equal to the inverse of the analysis error covariance matrix (Rabier and Courtier, 1992):

$$\nabla \nabla J = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} = \mathbf{A}^{-1}, \tag{3.6}$$

where $\mathbf{H}$ is now the tangent linear version of $\mathbf{h}$, and $\mathbf{A}$ is the analysis error covariance matrix. Hessian singular vectors (HSV) are different to TESVs because the norm used at initial time is the Hessian of the cost function instead of total energy, but the final time norm is

total energy; as with TESVs. HSVs should give more accurate representations of where the atmosphere is sensitive to errors in the intial conditions because the initial time norm is related to the analysis error covariance matrix, as shown by (3.6), and so has knowledge of the error statistics used in the data assimilation scheme. HSVs calculated using the Hessian of the 3D-Var cost function and background error statistics using the NMC method (Parrish and Derber, 1992) exhibit structure on larger scales and have a slower growth rate than TESVs (Barkmeijer et al., 1999). Lawrence et al. (2009) investigated the properties of HSVs that used more realistic background error statistics calculated using an ensemble of 4D-Vars (Fisher and Andersson, 2001) where random noise is added to observations to create each ensemble member, instead of the NMC method. Lawrence et al. (2009) found that the HSVs display similar characteristics to TESVs, confirming that TESVs are appropriate for targeting purposes.

Moist adjoint SVs have also been used to identify regions of sensitivity for targeting, and were used along with TESVs and HSVs during the Atlantic THORPEX Regional Campaign (ATReC). The computation of moist adjoint SVs uses a full-physics model, instead of a dry dynamics model as is used to calculate TESVs and HSVs. This inclusion of moist processes has been shown to be important to correctly represent severe cyclonic development (Hoskins and Coutinho, 2005).

Singular vectors are calculated operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF), to determine the perturbations for their ensemble prediction system (EPS). A comparison of the growth rates of singular vectors at different times of the year shows that they generally have largest growth rates in winter and smallest in summer (Buizza and Palmer, 1995). The dominant singular vectors are typically located in the principal areas of cyclogenesis in the northern hemisphere, in regions of baroclinic instability. A measure of the growth rate of the most unstable baroclinic mode, or Eady mode, is the Eady index σ_e, which is defined following Hoskins et al. (2000) as:

$$\sigma_e = 0.31 \frac{f}{N} \frac{du}{dz}, \qquad (3.7)$$

where f is the Coriolis parameter, N is the static stability and $\frac{du}{dz}$ is the vertical shear of the wind component u. Mapping the distribution of the Eady index was found to be a good proxy for the location of the dominant singular vectors, although the growth rate calculated by the

Eady index often underestimated that of the singular vectors (Buizza and Palmer, 1995).

Singular vectors are nonmodal, that is they do not preserve their shape between initial and final time. The structure of the singular vector is dependent on the choice of the metric. However both TESVs and HSVs show an upscale energy cascade during the optimisation period. At initial time the energy is dominated by potential energy on sub-synoptic scales, and at final time by kinetic energy on synoptic scales (Buizza and Palmer, 1995). Singular vectors with maximum amplitude are typically located in the lower troposphere at initial time, and tilt backwards with height, against the direction of vertical wind shear. This characteristic baroclinic structure changes during the optimisation period, so that the maximum amplitude of the SV is located near the jet level in the upper troposphere at final time, and is vertically stacked, so its structure is more barotropic.

3.1.2 The Ensemble Transform Kalman Filter

The ETKF can also be used to estimate which regions of the atmosphere are sensitive to uncertainties in the initial conditions. A full description of the theory behind the ETKF is given in Bishop et al. (2001). The ETKF is an extension to the Kalman Filter. The Kalman filter is the procedure by which an estimate of the state of the atmosphere and analysis error covariance $\mathbf{P}^a$ are obtained at some time t after observations have been made of the state of the atmosphere (written according to Daley (1991))

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T [\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}]^{-1} \mathbf{H}\mathbf{P}^f, \qquad (3.8)$$

where $\mathbf{P}^f$ is the forecast error covariance matrix, given by:

$$\mathbf{P}^f = \mathbf{M}\mathbf{P}^a\mathbf{M} + \mathbf{Q}, \qquad (3.9)$$

where $\mathbf{M}$ is the linearised forecast model and $\mathbf{Q}$ is the model error covariance matrix. A major problem in implementing the Kalman filter in NWP models is the size of the error covariance matrices. $\mathbf{P}^f$ is an array of approximately 10^7 by 10^7 elements, therefore we cannot store $[\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}]$, let alone invert it. A way to overcome this is to approximate the error covariances by transforming perturbations from an ensemble of forecasts instead of using the error covariance matrix itself. The ETKF assumes that you can linearly combine

nonlinear ensemble perturbations, and then transform them in such a way that they describe the forecast error covariance that would be obtained by making targeted observations (Bishop et al., 2001).

The ETKF therefore requires only an ensemble of K forecasts $\mathbf{x}_i$, with ensemble mean $\overline{\mathbf{x}}$. K-1 independent ensemble perturbations $\mathbf{X}$ are created by differencing each forecast from the ensemble mean:

$$\mathbf{X} = \frac{\sum_{i=1}^{K} \mathbf{x}_i - \overline{\mathbf{x}}}{(K-1)^{0.5}}, \tag{3.10}$$

where the forecast error covariance matrix $\mathbf{P}^f(t)$ at time t is given by the outer product of the ensemble perturbations:

$$\mathbf{P}^f(t) = \mathbf{X}(t)\mathbf{X}(t)^T. \tag{3.11}$$

The analysis error covariance matrix at some time $t + m$, where $t + m \geq t$ is then expressed by a transformation of $\mathbf{X}$:

$$\mathbf{P}^a(t + m) = \mathbf{X}(t + m)\mathbf{T}\mathbf{T}^T\mathbf{X}(t + m)^T, \tag{3.12}$$

where $\mathbf{T}$ is a transformation matrix that transforms the ensemble perturbations into perturbations that are orthonormal with respect to an inverse analysis error variance norm (Majumdar et al., 2002b):

$$\mathbf{X}^T\mathbf{T}^T\mathbf{D}^{-1}\mathbf{X}\mathbf{T} = \mathbf{I}, \tag{3.13}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{D}^{-1}$ is a matrix which contains the inverse of the analysis error variance associated with the routine observation network. For the ETKF used to make the SAP plots, $\mathbf{X}$ contains T, u and v at 850, 500 and 250hPa levels at 2.5° resolution. P_{t+m} is first evaluated for the routine observations only, where the radiosonde network is used as a proxy for the routine observing network. This is then used as a first guess to calculate P_{t+m} for the routine observing network plus the particular deployment of targeted observations. For the targeted observations, test sets of observations consisting of 9 observations at different

horizontal positions are used, where each observation contains u, v and T at 850, 500 and 250hPa levels. P_{t+m} is calculated for a particular deployment of observations, and then this is used as the first guess for P_{t+m} for the next deployment of observations. This use of serial observation processing makes the ETKF efficient.

The ETKF can also be used to give the forecast signal variance, where the forecast signal from the targeted observations $\mathbf{s}^q(t)$ is:

$$\mathbf{s}^q(t) = \mathbf{x}^q(t) - \mathbf{x}^r(t), \tag{3.14}$$

where $\mathbf{x}^r(t)$ is the forecast containing only routine observations, and $\mathbf{x}^q(t)$ is the forecast containing both routine and targeted observations. The signal covariance matrix $\mathbf{S}^q$ is then the covariance of $\mathbf{s}^q$:

$$\mathbf{S}^q(t) = \langle \mathbf{s}^q \mathbf{s}^{qT} \rangle = \langle (\mathbf{x}^q(t) - \mathbf{x}^r(t))(\mathbf{x}^q(t) - \mathbf{x}^r(t))^T \rangle. \tag{3.15}$$

Without having to assimilate the targeted observations and run the forecasts, $\mathbf{S}^q$ can be calculated using the transformed ensemble perturbations. This signal variance has been shown to be equivalent to the reduction in forecast error variance obtained by assimilating targeted observations. Majumdar et al. (2001) found an positive although nonlinear relationship between the ETKF signal variance run on an National Centers for Environmental Prediction (NCEP) ensemble, and resulting reduction in forecast error variance due to additional observations assimilated into the NCEP deterministic forecast. The predicted error reduction may not be realised due to differences between the assumed error statistics of the ETKF and the data assimilation scheme. The ETKF assumes that the data assimilation scheme used to assimilate the targeted observations into the forecast uses similar error statistics to the Kalman filter (Bishop et al., 2001). This is because the ETKF assumes the error covariance of an ensemble Kalman filter (EnKF) (Majumdar et al., 2002b). Due to this, using 4D-Var to assimilate the additional observations provides better results than using 3D-Var, as the error statistics used in 4D-Var are flow-dependent, like those of the EnKF (Bergot, 2001).

An advantage of using the ETKF is therefore that it can provide information on the degree of sensitivity in the forecast on a given day. This is achieved by plotting non-normalised signal

variance, so that the magnitude of forecast sensitivity can be compared from one day to the next. This is particularly useful for operational targeting where decisions have to be made on whether or not to make targeted observations, to optimise the use of the available resources.

The ETKF perturbations do not typically exhibit a tilt with height (Petersen et al., 2007). Thus vertically averaging the signal variance produced at different model levels throughout the atmosphere to define the target region is justified. Petersen et al. (2007) showed that an ensemble of 10 members is sufficient to provide an accurate summary map for targeting guidance; increasing the number of ensemble members does not change the target regions identified by the ETKF. There is some debate as to the locations identified by the ETKF as target regions. One view is that the ETKF identifies regions of the upper troposphere where a baroclinic wave has already developed (Langland, 2005) which is in contrast to the findings of Petersen et al. (2007) who found ETKF targets to lie more often near the jet-stream in the upper-troposphere, and areas of baroclinic instability in the mid-troposphere.

3.1.3 Comparison of SV and ETKF Techniques

The different characteristics of SV and ETKF perturbations mean that the sensitive areas identified by each often differ. In ATReC it was found that in only 46% of cases did the sensitive areas calculated using the ETKF and TESVs overlap by more than 0.5 (Leutbecher et al., 2004). Conflicting guidance from different SAPs can make it more difficult to make decisions on where to make the additional observations. This is in contrast to the findings of Majumdar et al. (2002a) who found the ETKF and ECMWF TESV summary maps to be well-correlated in seven out of ten North Pacific Experiment (NORPEX) cases studied, and ETKF and Naval Research Laboratory TESV maps well-correlated in nine out of ten cases. The differences in result here may be related to differences in method and definition of similarity. Leutbecher et al. (2004) said that two target regions were similar if they had a geographical overlap between two regions of greater than 0.5, whereas Majumdar et al. (2002a) use the terminology when a possible flight track to sample each target region has a Spearman rank correlation coefficient greater than 0.46 and a positive modified equitable threat score.

The areas highlighted as sensitive by the different types of singular vectors do not always agree. A simple measure of the agreement between two SAPs is to look at the ratio of the overlap between the two sensitive areas S_j and S_k to the size of the sensitive area, defined in

Leutbecher et al. (2004) as

$$O_{jk} = \frac{area(S_j \bigcap S_k)}{area(S_j)},\tag{3.16}$$

assuming that the size of the sensitive areas are equal. Thus an overlap of 1 means the target regions are identical, and an overlap of 0 means the target regions do not intersect at all. However it was shown during ATReC that for 83% of cases the HSVs and TESVs had a geographical overlap of at least 0.5, which rose to 98% for moist-adjoint SVs and TESVs (Leutbecher et al., 2004), calculated using the number of gridpoints instead of an area as given by the above equation. This is a larger overlap than between TESV and ETKF sensitive areas. The methods will identify slightly different areas of sensitivity because of differences in the way they calculate the SVs. An advantage of using moist adjoint SVs is that they use full-physics dynamics rather than only dry dynamics and so they account for large-scale latent heat release. This is important for severe cyclonic development (Hoskins and Coutinho, 2005). However analyses of the Fronts and Atlantic Storm Track Experiment (FASTEX) data showed that TESVs correctly identified regions of sensitivity in the forecast, even though moist processes were shown to have been important in those cases (Gelaro et al., 1999).

There are differences in the types of region identified by ETKF- and TESV-SAPs. The reader is referred to Majumdar et al. (2002a) for a detailed discussion of the reasons for this. Singular vectors identify regions where a small initial error will grow rapidly over the forecast time, thus singular vectors indicate areas of large potential for error growth. By contrast, the ETKF identifies regions that result in a large error at verification time, thus the ETKF identifies both regions where an initial error has grown rapidly to be large at verification time and regions where the initial error was large but did not grow. Singular vectors often identify sensitive regions over data-rich regions, as the position of the routine observing system is not taken into account during their calculation. As the ETKF takes into account the position of the radiosonde network, it tends to identify only regions away from the routine observing network. However, it may also identify spurious sensitive regions due to a lack of covariance localisation (Majumdar et al., 2002b). It has been noted during field campaigns such as ATReC (G. N. Petersen, personal communication) and GFDex, that SVs tend to pick out the leading edge of a developing baroclinic system as a target region, whereas the ETKF will

identify the centre of the same system as being most sensitive. This may be related to the ability of current NWP models to accurately predict the evolution of a baroclinic system. As the ETKF is concerned with the magnitude of forecast error, if the largest difference between individual ensemble forecast members is the intensity of a developing low pressure system then the centre of the low will be identified as this is where the analysis error variance is largest.

Majumdar et al. (2006) compared TESV and ETKF guidance for tropical cyclone targeting. They found their results to be dependent on the scale of the target. For major hurricanes (defined as category 3 and above) the guidance for larger-scale targets was similar for 86% of cases, but this reduced to 25% for weaker hurricanes. For smaller scale targets, such as in the vicinity of the cyclone itself, the guidance often differed between the two methods. The results of this study may have been affected by not taking into account gridpoints within 333km of the cyclone centre. This was done because targeting in this area is currently made redundant by the inability of the data assimilation schemes to extract useful information from observational data so close to the cyclone centre. Aberson (2008) found that assimilating dropsonde observations that sample the core of tropical cyclones can lead to large forecast degradations.

3.2 Types of Targeted Observations

Targeting studies to date have largely focussed on the use of dropsonde observations as targeted observations. The advantage of using direct measurements of the atmosphere is that the data is high-resolution in the vertical and has small errors. The current observing network lacks high-resolution profile data in data-sparse regions such as the North Atlantic (Section 2.1). The disadvantage of using dropsonde measurements is that to obtain the data requires an aircraft equipped to release the sondes. This is both expensive and allows little flexibility in the location of the target region as it requires that such an aircraft is already within range of the target region. The target region must also be identified up to 24 hours in advance of the observation deployment, especially in regions where there is a lot of commerical air traffic, due to air traffic control restrictions.

Satellite data has the potential to be used as targeted data, due to its good data coverage. Only 3% of satellite data available for assimilation is actually assimilated by the Met Office

(pers. comm. Roger Saunders), therefore extra 'targeted' data is free. This also means that the sensitive regions do not need to be identified far in advance, as targeted data can easily be obtained for any region of the globe. The disadvantage of satellite data as targeted observations is that the data is remotely-sensed, resulting in lower vertical resolution of data and larger errors. Profiles from infra-red sounders, like ATOVS data, are also only available above cloud tops. This is a problem for targeting observations because targeted observations are typically taken in baroclinically active regions, which are often cloudy (McNally, 2002). The satellites may therefore not be able to sample shallow sensitive regions indicated by singular vectors to be located within the cloud layer (Lawrence et al., 2009). Microwave sensors are less sensitive to the presence of cloud, however Dando (2007) found only small forecast improvements by assimilating targeted ATOVS observations. To date there have been few studies that have assessed the potential of using additional satellite data as targeted observations; improved forecast products and data assimilation schemes may increase the potential benefit of satellite data as targeted observations in data-sparse regions such as the Arctic.

During ATReC, lidar observations were used as targeted observations, in addition to dropsonde observations. Weissmann and Cardinali (2007) assimilated airborne Doppler lidar observations as targeted observations, achieving an average 2-4 day forecast error reduction in geopotential height of 3%. The 2-μm Doppler lidar was mounted on the DLR aircraft, and measured profiles of wind speed and direction with a vertical resolution of 100m, and a horizontal resolution of 5-10km (Weissmann et al., 2005). The disadvantage of lidar measurements compared to dropsonde measurements is that the lidar cannot see through cloud, and therefore has the same problems as satellite data.

The Eurorisk-PREVIEW (PREVention, Information and Early Warning) programme is trialling targeted observations in an operational manner. When the target region is identified, extra observations are gained by requesting additional AMDAR observations and additional radiosonde ascents (more information can be found at http://www.preview-risk.com). The benefit of this approach is that it uses existing observation platforms, rather than flying a plane into the sensitive region and then releasing instruments, and therefore is a cheaper and more flexible approach to targeted observing.

3.3   Overview of Idealised Targeting Studies

Field campaign trials of targeted observations are limited by their small number of cases, which make it difficult to draw general conclusions about the success of the targeting and assess whether the results are statistically significant. Buizza et al. (2007) (see also Cardinali et al. (2007) and Kelly et al. (2007)) addressed these issues by running large numbers of hindcast experiments for cases where the extra observations were placed over the Atlantic or Pacific oceans and were either targeted in TESV sensitive regions, or random regions. They found that the impact of targeted observations in TESV sensitive regions was larger in winter than summer, and larger for the Pacific than the Atlantic region. The routine observing network in the target region was also found to greatly influence the forecast impact from targeted observations, with an average forecast error reduction of 19.1% if only targeted observations were assimilated over the ocean (i.e. no routine observations assimilated over the ocean) reducing to only 2% if the observing network in the targeting region could not be considered sparse (routine observations assimilated over the ocean). They also concluded that targeted observations in sensitive regions were more valuable than in random regions. Kelly et al. (2007) assessed the effect of the data assimilation scheme on the potential value of targeted observations by removing observations from the north Pacific and north Atlantic and using 3D-Var and 4D-Var to assimilate the observations. Their results showed that 4D-Var was better able to cope with gaps in the observing network.

Cardinali et al. (2007) showed the dependence of the forecast error reduction from targeted observations on the weather regime and time of year. Modelling experiments showed that a greater improvement from targeting should be seen in the winter, when the circulation is strongest, and the growth of the singular vectors is fastest. This may be related to the ability of the singular vector method to correctly identify target regions, which may therefore be degraded in summer (Buizza et al., 2007). The study focussed on periods when a tropical cyclone moved into the extra-tropics and weakened as it moved north. Targeted observations in sensitive areas during these so-called extra-tropical transition periods were more effective than outside of these periods, particularly when there was non-zonal and unstable flow in the north Atlantic.

3.4   Overview of Previous Field Campaign Targeting Studies

There have been a number of trials of targeted observations during field campaigns (Langland (2005) provides an overview of these campaigns). Table 3.1 details the chronology of targeting studies, and the average forecast error reduction that these have achieved using dropsondes as targeted observations. Table 3.1 is not a comprehensive list of all of the targeting studies described in this chapter; it details only those studies which reported the average forecast error reduction. The forecast error reduction given in Table 3.1 is for forecasts up to 2 days, measured as a reduction in the root mean squared error in 500 and/or 1000 hPa geopotential height. It is clear from Table 3.1 that field trials of targeted observations have had mixed success in their attempts to use targeted dropsonde data to improve forecasts. FASTEX in 1997 (Bergot, 1999) was the first field campaign to objectively target sensitive areas to improve the forecasts of landfalling systems downstream of the target area (the north Atlantic). Both the Ensemble Transform method (ET; Bishop and Toth (1999)) and TESV (Section 3.1) were used to identify the sensitive region, where the observations should be targeted. The signal from the targeted sondes was shown to move and amplify from the target region to the verification region, suggesting that both methods successfully identified sensitive regions (Szunyogh et al., 1999; Montani et al., 1999). The largest average reduction in the forecast error was a reduction in the two-day ECMWF forecast error of 15%, averaged over five TESV targeting cases (Montani et al., 1999). Szunyogh et al. (1999) analysed seven ET targeted cases using the NCEP model and found an average forecast error reduction of 10%. Bergot (1999) analysed all 20 targeting cases from FASTEX (Montani et al. (1999); Szunyogh et al. (1999) analysed subsets of these cases using different models) , and found that on average (over all forecast cases) the forecast impact was small, with only weak improvement. Targeted observations taken closer to the north American continent, where there is good data coverage, had less impact than targeted observations taken in the data void of the north Atlantic.

The large forecast improvement seen in some cases during FASTEX encouraged further trials of targeted observations, in NORPEX in 1998 (Langland et al., 1999; Liu and Zou, 2001) and Winter Storm Reconnaissance program (WSR) (Szunyogh et al., 2000, 2002) which has run every year since 1999. In both of these field campaigns a reduction in the two-day forecast error of 10-20% was achieved as a result of assimilating targeted sonde data. The

Table 3.1: *Previous Targeting Experiments*

Study (Project)	Target Region	Assimilation Scheme	Grid-length (km)	Mean Error Reduction (%)	Cases
Montani et al. (1999) (FASTEX)	N Atlantic	3D-Var	60	15	5
Szunyogh et al. (1999) (FASTEX)	N Atlantic	3D-Var	200	10	7
Bergot (2001) (FASTEX)	N Atlantic	3D-Var	140	0.9	20
		4D-Var	140	6.4	20
Langland et al. (1999) (NORPEX)	NW Pacific	OI	200	10	27
Szunyogh et al. (2000) (WSR)	NW Pacific	3D-Var	200	10-20	25
Petersen et al. (2007) (ATReC)	N Atlantic	4D-Var	60	≥ 5 ($\frac{1}{3}$ cases)	38

aim of the WSR was to improve the forecasts of extra-tropical cyclones making landfall on the west coast of the USA. The largest impact was achieved when landfalling systems were directly targeted with dropsondes (Szunyogh et al., 2002). The targeted data always sampled dynamical features such as upper tropospheric waves or cyclones, areas of baroclinic energy conversion. The signal from the targeted data (defined as the difference between a forecast containing routine and targeted data, and one containing only routine data) was shown to propagate downstream at 30° longitude per day (Szunyogh et al., 2000).

The data assimilation scheme used to assimilate the targeted data also affects the impact from targeted observations. Liu and Zou (2001) and Bergot (2001) assimilated the same sets of observations using both 3D-Var and 4D-Var and showed that a greater reduction in forecast error was seen when 4D-Var was used to assimilate the observations. These results are in contrast to those of Kelly et al. (2007) (discussed in Section 3.3) who found that a smaller impact should be expected when using 4D-Var, due to the ability of 4D-Var to cope with gaps

in the observing system by propagating data from data-rich to data-sparse regions. There are several possible reasons why these studies report conflicting results. The Kelly et al. (2007) study looked at a much larger number of cases, meaning that the results have more statistical significance. An important difference is the length of the assimilation window used in the 4D-Var scheme. Both Liu and Zou (2001) and Bergot (2001) use 4D-Var schemes with a 6-hour window for observations, at relatively coarse resolution (T62 and T95) whereas Kelly et al. (2007) uses a 12-hour window and T511 model (T159 in the inner loop). This may account for the increased ability of the 4D-Var scheme to propagate information in Kelly et al. (2007). Bergot (2001) also found that the increase in improvement seen in the short and medium-range forecasts using 4D-Var did not increase the spread of the medium-range forecasts, as was seen when 3D-Var was used to assimilate the observations. This study also compared the effectiveness of the location of the additional observations relative to the location of the sensitive area. It found that when additional observations were taken at the edge of the sensitive areas, e.g. not in the most sensitive regions, assimilating these using 3D-Var caused a deterioration in the forecast, but that assimilating the same data using 4D-Var showed a small improvement in the forecast. This demonstrates the dependence of the results on the assimilation method.

Recently, tropical cyclones have also been targeted with dropsonde observations to try to improve the forecast of the tropical cyclone track. Targeting of Atlantic hurricanes has achieved 2-day forecast track error reductions of 25% (Aberson, 2003). The program Drop-windsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) carried out 10 targeting experiments of different tropical cyclones and found that the assimilation of targeted sondes improved the 72-hour forecast of the tropical cyclone tracks by 22%, averaged over three different global models.

These earlier studies used global forecast models with gridlengths of 100 km and greater - now considered to be relatively coarse resolution (the possible exception to this is the study of Montani et al. (1999), who used a spectral model at T213 truncation, which is approximately 60 km resolution), and used data assimilation schemes that did not take into account the exact time at which the observations were made (i.e. were not four-dimensional schemes). More recently, ATReC showed a mixture of small improvement and degradation due to targeted observing, using four-dimensional schemes to assimilate the targeted data (Petersen and Thorpe,

2007; Fourrié et al., 2006). The reduced impact of the targeted data during ATReC, compared to earlier targeting studies (Table 3.1), could be due to improved data assimilation and model formulation, along with an increased number of routine observations, which together have limited the scope for improvements from targeting. This does however highlight the need for a study of targeted observations using an up-to-date model of higher-resolution than has previously been used. This issue is addressed in this thesis by using a 24 km resolution regional model to run the hindcast studies.

CHAPTER 4

The Impact of GFDex Targeted Observations

This chapter presents the impact of the GFDex targeted sonde data on the 1-3 day forecasts for Northern Europe, for five cases. The impact seen in two cases is presented in greater detail, examining the mechanisms for the forecast impact: in the first case the improvement propagates into the verification region with a developing polar low; in the second case the improvement is associated with an upper-level trough. The impact of cycling targeted data in the background of the forecast (including the memory of previous targeted observations) is investigated. Finally, the robustness of the results is assessed using a small ensemble of forecasts.

4.1 Introduction

The aim of targeted observing is to improve the short-term forecast for a specified region and time by the assimilation of additional observational data into operational weather forecasts. For a limited-area model, there are three potential sources of forecast error: model deficiencies, errors propagating in from lateral boundary conditions and errors in the initial conditions. The aim of targeted observing is to reduce the initial condition error in a sensitive region, where errors have the potential to grow (or are large at the targeting time) and therefore degrade the forecast downstream of the target region a short time later. Targeted observing should therefore reduce forecast 'busts' which occur as a result of the growth of initial condition error. Field programmes which have trialled targeted observing have taken these observations in oceanic regions, away from the radiosonde network, with the rationale that initial condition errors are most likely to occur in regions where there are fewer observations available to constrain the background field.

Targeted observations have the potential to be used as part of an adaptive observing system, where additional observations are targeted according to the synoptic situation. This was trialled during 2008 under the Eurorisk-PREVIEW programme (http://www.preview-

risk.com). Here we aim to contribute to the debate on the effectiveness of targeted observing, by analysing the impact of targeted observations made during the field campaign of the Greenland Flow Distortion experiment in February-March 2007 (GFDex; see Renfrew et al. (2008)).

The GFDex was a three-week field campaign which used the Facility for Atmospheric Airborne Measurement (FAAM) BAE-146 aircraft based out of Keflavik, Iceland from 19 February to 12 March 2007. The GFDex carried out targeted observing flights in the region around Iceland and southern Greenland using the FAAM aircraft to release dropsondes into sensitive regions. Here we present an analysis of the impact of these targeted observations. The set-up of the targeting experiments and the details of the hindcasts performed to analyse the results are given in Sections 4.2 and 4.3. An overview of the impact seen in each of the five targeting cases is given in Section 4.4. Two cases where the targeting was successful are analysed in more detail in Sections 4.5 and 4.6. As sondes were released for several targeting flights over only two and a half weeks, this offers us the chance to study the impact of memory of previous targeted observations on the forecast, examined in Section 4.7. Ensemble methods are used to quantify the significance of the forecast improvement in Section 4.8, followed by discussion and conclusion in Section 4.9.

4.2 Set-up of the GFDex Targeting Experiments

This section describes the set-up of the targeting experiments during GFDex: the sensitive area predictions, the timelines involved in the targeting and finally an overview of the targeting flights made.

4.2.1 Sensitive Area Predictions

The sensitivity calculations were performed for two fixed verification regions. These were approximately square regions of side 2000 km. One region was positioned over north-west Europe centred over the UK and the other over Scandinavia (Figure 4.1). The verification regions were placed downstream of Greenland, in areas where flow perturbations from the target region might be expected to move to within the specified optimisation times. The regions were similar in size to those used by previous targeting experiments such as FASTEX (Montani et al., 1999) and ATReC (Petersen and Thorpe, 2007).
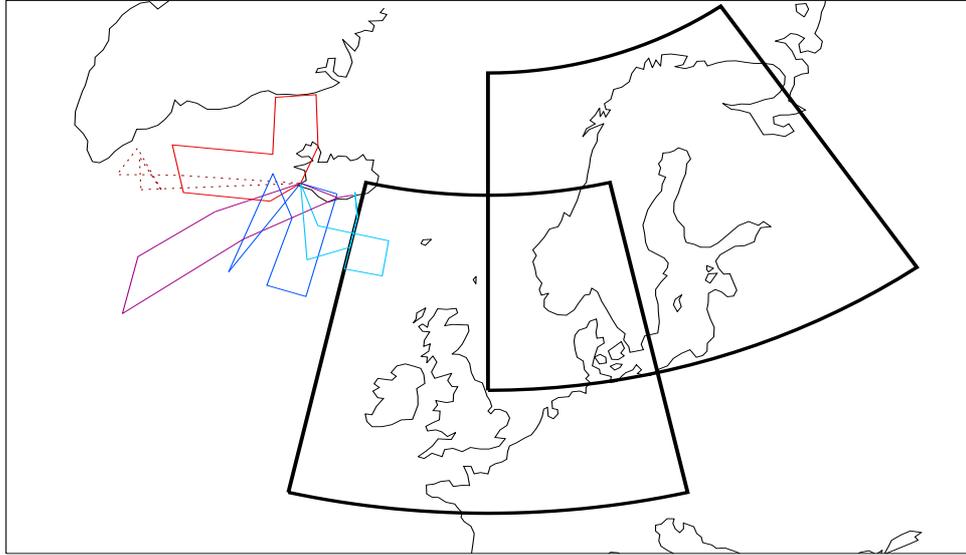
Figure 4.1: *Map of the study region, showing the flight tracks of the targeting (solid lines) and null (dotted line) flights during the GFDex. The flights started from Iceland. The two verification regions, one centred over the UK (referred to in the text as north-west Europe) and one over Scandinavia, are marked by bold lines.*

The TESV-SAPs (Section 3.1) were generated by ECMWF, at T42 resolution, using output from their deterministic global forecast model. Total energy singular vectors are used to identify the fastest growing perturbations over the optimisation time that would impact the verification region. The growth of the singular vectors was measured using a total energy norm. This is an appropriate metric to use for targeting purposes as it provides a good approximation to the analysis error covariance matrix used by data assimilation schemes (Palmer et al., 1998). The SAP plots are generated by vertically averaging the ten leading singular vectors, which is sufficient to gain a good representation of the target region (Buizza and Montani, 1999).

The ETKF-SAPs (Section 3.1) were provided by the Met Office, at a resolution of 2.5 degrees, using output from the 23 member 15-day global Met Office Global and Regional Ensemble Prediction System (MOGREPS; Bowler et al. (2008)). The ETKF takes the MOGREPS perturbations (ensemble member minus the ensemble mean) and transforms them to obtain an estimate of the analysis error covariance within the verification region from assimilating routine observations only, or routine observations plus targeted observations. Plots of

non-normalised signal variance, roughly equivalent to the reduction in forecast error variance from assimilating real targeted observations (Majumdar et al., 2001), were also provided by the Met Office during the campaign. This allowed the magnitude of forecast sensitivity to be compared from one day to the next; this information was not provided by the TESV sensitivity calculations.

Little overlap between the ETKF- and TESV-SAPs was observed during the campaign. This was also the case during the ATReC campaign, where the ETKF and singular vector sensitive areas overlapped by more than half in only 46% of cases (Leutbecher et al., 2004). There was no overlap in the regions of maximum sensitivity for the targeted cases. This meant that flights were tasked using either TESV- or ETKF-SAPs, but not both. It was noted that the TESV sensitive regions tended to be located further north and on the eastern side of low pressure systems, whereas the ETKF-SAPs were located further south and towards the centre of lows. It is beyond the scope of this study to analyse the reasons for these differences in locations; further analysis would be required to determine whether the differences are systematic and their cause.

4.2.2   Timelines Involved in Targeted Observations

Targeted observations were made over a period of approximately four hours, centred on 1200 UTC (hereafter referred to as the *target* time, Figure 3.1). A constraint on operational targeted observing is the need for a time-lag between running the sensitivity calculations and making the targeted observations. This so-called *lead time* allows the resources to be mobilised and flight plans to be drawn-up. During the GFDex, air traffic control restrictions meant that flight plans had to be filed 24 hours before take-off. Thus SAPs with a 36-hour lead time were used to plan the targeted observation flights. SAPs were also received with lead times of 48 and 60 hours, which were used to make preliminary flight plans. SAPs with a 24-hour lead time were used as a final check on the location of the sensitive region. It was observed during the campaign that SAPs with a 24-hour lead time were consistent with the SAPs with a 36-hour lead time. This agrees with the findings of Petersen et al. (2007) that there is good spatial coherence between SAP plots generated 12 hours apart. SAPs were generated for three optimisation times: 24, 36 and 48 hours. A range of optimisation times was necessary to allow the signal from the dropsondes to move from the target region into the

verification region. This distance varied from case to case, as the target area ranged from the Denmark Strait, to the south-east of Iceland. The speed with which the impact from targeted observations propagates downstream was shown in Szunyogh et al. (2000) to be 30° longitude per day suggesting that downstream development may be important, as the speed is consistent with the group velocity of a dispersive Rossby wave in the atmosphere (Persson (1999) and references therein).

The SAPs were generated from forecasts run at 0000 UTC and 1200 UTC daily. The SAP plots were available for us to analyse 7-8 hours (ECMWF SAPs) and 10-11 hours (Met Office SAPs) after this. This means that the total time from forecast initialisation to verifying the forecast containing the targeted observations varied from 48 hours to 108 hours. This tests the limits of the assumptions of linearity made by both ETKF and SV methods as a growing perturbation is likely to be in a non-linear regime after 48 hours (Gilmour et al., 2001). However during the field campaign it was observed that there was little variation in the location of the sensitive regions identified by either the ETKF or SV methods as the total time decreased.

4.2.3 Overview of GFDex Targeting Flights

Targeted observing was carried out on four out of twelve flights during the field campaign; three of these flights were into TESV-predicted sensitive regions and one into an ETKF-predicted sensitive region. Targeting flights were made nearly every time a sensitive region was predicted to be in-range of the aircraft. This is a different strategy to that which would be used for operational targeting, as due to the cost of targeting, it could only be justified to reduce the uncertainty in forecasts which indicate the possibility of a high-impact weather event. After the campaign a null case was chosen from the other flights. This was a flight into a Greenland lee cyclone, into an area which neither the ETKF- nor TESV-SAPs predicted to be sensitive to initial condition errors. The assimilation of additional data here should not significantly impact the forecast within the verification regions. Details of the flights are given in Table 4.1, and the flight tracks are shown in Figure 4.1.

Global Positioning System (GPS) dropsondes (Vaisala RS92) released from the aircraft were used to make the observations. These measure vertical profiles of height, pressure, air temperature, humidity and wind, transmitting the data back to the aircraft. The data were sent

Table 4.1: *Specification of targeting and null case studies*

Date	Target Region	SAP type	Sensitivity for NW Europe for optimisation times (hrs)			Sensitivity for Scandinavia for optimisation times (hrs)			Targeted /Total Sondes deployed
			24	36	48	24	36	48	
24/02/07	S of Iceland	TESV	Y	Y	Y	N	N	Y	11/11
26/02/07	SE of Iceland	TESV	Y	Y	Y	Y	Y	Y	7/14
01/03/07	Denmark Strait	TESV	N	Y	Y	Y	Y	Y	9/18
03/03/07	SW of Iceland	NULL	N	N	N	N	N	N	5/16
10/03/07	SW of Iceland	ETKF	Y	Y	Y	Y	Y	Y	9/15

onto the GTS via a satellite communications device, during the flight. The data were therefore available for meteorological centres to use in the operational 1200 UTC forecasts. Data from all GFDex sondes (targeted and non-targeted) were assimilated into operational forecasts.

Each targeted observing flight during the GFDex released 7-11 targeted sondes (out of 11-18 total sondes per flight, see Table 4.1), covering an area of approximately 1×10^5 km^2. For the null case, five of the 16 sondes that were launched during the flight are treated as targeted sondes. Note that some flights released additional sondes, some inside and some outside the predicted sensitive regions, for additional observational aims. The treatment of these non-targeted sondes is discussed in Section 4.3.2. The spatial coverage of the target area and number of dropsondes released within it has been shown to affect the success of the targeting. Idealised modelling studies suggest that releasing between 10 and 40 dropsondes within an area of 3×10^6 km^2 should be sufficient to improve the 2-day forecast over Europe (Leutbecher et al., 2002). Buizza and Montani (1999) showed that insufficient sampling of a target area can result in only 65% of the possible theoretical forecast improvement being realised. However sampling the large area indicated necessary by the modelling studies with one aircraft like the BAE-146 used in the GFDex is logistically impossible. This is due to

the limited range of the plane and the need to release all the sondes within a 4.5-hour time window so that they do not miss the cut-off time for the operational Met Office global forecast model (1345 UTC for the 1200 UTC forecast).

The spacing of the dropsondes can also be important to the success of the targeting. Numerical modelling studies indicate that the spacing of the dropsondes should be 1-2 times the horizontal correlation length-scale of the variables, such as wind and temperature, being assimilated (Leutbecher et al., 2002). As the analysis increments are spread out according to the error correlation specification in the data assimilation system, it is evident that dropsondes placed closer than a correlation length will lead to increments which overlap. The horizontal correlation length-scale for temperature for example varies with height, from around 280 km at the height the dropsondes were released from to approximately 200 km near the ground (Ingleby, 2001). Taking these results into account, and allowing for the range of the plane, the spacing between targeted dropsondes in the GFDex was always greater than 180 km and for most flights was around 220 km.

4.3   Design of Hindcast Studies

This section describes the model used to run the hindcasts, and the different hindcast runs performed with the model.

4.3.1   Description of Model

To perform the hindcast studies the MetUM (Section 2.4) was run as a limited-area model, over the operational NAE domain (Figure 2.5), with 24 km grid spacing and 38 vertical levels. This is half the horizontal resolution of the operational NAE forecasts, due to the large computational cost of running the hindcasts. The justification for running a limited area model, rather than a global model, was that both the target and verification regions were inside the limited area domain, although near the northern edge, and this was sufficient to study the impact of sonde data on the forecasts for the verification regions. Using a limited area model also allowed the hindcasts to be run at higher resolution. Lateral boundary conditions were provided by the operational global control run of MOGREPS. A 4D-Var data assimilation scheme with a six-hour window was used for the assimilation of observations (Rawlins et al., 2007). The hindcast studies used the same assimilation window as the operational NAE forecast cycle,

Table 4.2: *Definition of the Hindcast Runs*

Run	Routine Observations	Non-targeted GFDex sondes	Targeted GFDex sondes	Continuous Assimilation
CONTROL	Y	N	N	Y
TNOMEM	Y	N	Y	N
TMEM	Y	N	Y	Y
ALL DATA	Y	Y	Y	Y

which for the 1200 UTC forecast was 0900 UTC to 1505 UTC, but a 48 km gridlength for the 4D-Var scheme (half the operational resolution).

Before observations are passed to the data assimilation scheme, they are first subject to a quality control process, based on Bayesian methods, to remove 'bad' observations (Section 2.2). With the exception of some wind data from one malfunctioning sonde, the GFDex observations were accepted by the quality control system and passed to the data assimilation scheme; only 0.2% of GFDex observations were rejected.

4.3.2   Description of the Hindcast Runs

To assess the impact of the targeted dropsondes on the forecast, the complete forecast cycle of four forecasts per day was re-run for the period 19 February to 13 March 2007, with different combinations of observations excluded from the forecast. Four separate forecast cycles were run: labelled CONTROL, TNOMEM, TMEM and ALL DATA (Table 4.2).

The CONTROL run assimilated only routine observations. Here 'routine observations' is taken to mean the non-adaptive component of the observing network, such as radiosonde, AMDAR, satellite and surface station observations. The CONTROL forecast run therefore produces a set of forecasts which have not been influenced by the additional observations made during the GFDex. Note that during the GFDex there were some additional radiosonde launches from land stations in Greenland and Iceland at 0600 and 1800 UTC and from three ships when they were within a pre-defined area around Iceland and southern Greenland (Renfrew et al., 2008). These have been classified as routine observations for this study, as the additional land launches were outside the observation window for the data assimilation scheme

for the 1200 UTC forecast and the ship releases were not within sensitive areas. Two other runs assimilated targeted observations along with routine observations. The TNOMEM hindcasts took the background from the CONTROL forecast and assimilated targeted observations for the 1200 UTC forecast on that *one particular targeting day*. This was done separately for each case study, so that the background for each TNOMEM forecast contains only routine observations. The set-up of the hindcasts means that any differences between the TNOMEM forecast and the CONTROL forecast are due to the targeted observations made that day. In contrast with the TNOMEM run, the TMEM run assimilated targeted observations *every time they were available*, so that the targeted data was cycled in the forecast. This allows the effect of memory of previous observations in the forecast to be investigated. Therefore any differences between the TMEM and TNOMEM forecasts are due to differences in the background of the two forecasts. As the model was run with a 24 km resolution instead of the 12 km used operationally, a run with the same observations as the operational run was also produced. This included routine observations and all the GFDex sondes, including non-targeted sonde data. This is the ALL DATA run, and was used to produce the analyses needed to verify the forecasts. The difference between the TMEM and ALL DATA runs is that the TMEM run does not include non-targeted sonde data. Each forecast was run for 63 hours, starting at 0900 UTC on the targeting day and outputting data every 6 hours.

4.4   Impact of the Targeted Observations

The targeted sonde data are taken in regions where there are few high-density profile measurements to constrain the background field. This may result in larger initial condition errors in this region, which may then propagate into the verification region. In order to correct these errors, if such errors exist, the sonde data must be different from the background field and its inclusion in the forecast should lead to an adjustment of the background field towards the sonde observations. Figure 4.2(a)-(e) shows the *difference* between the *v*-component (north-south) of the wind measured by the sonde and the CONTROL model *v* (at the location where the sonde was released), defined as

$$difference = | CONTROL - SONDE | . \qquad (4.1)$$
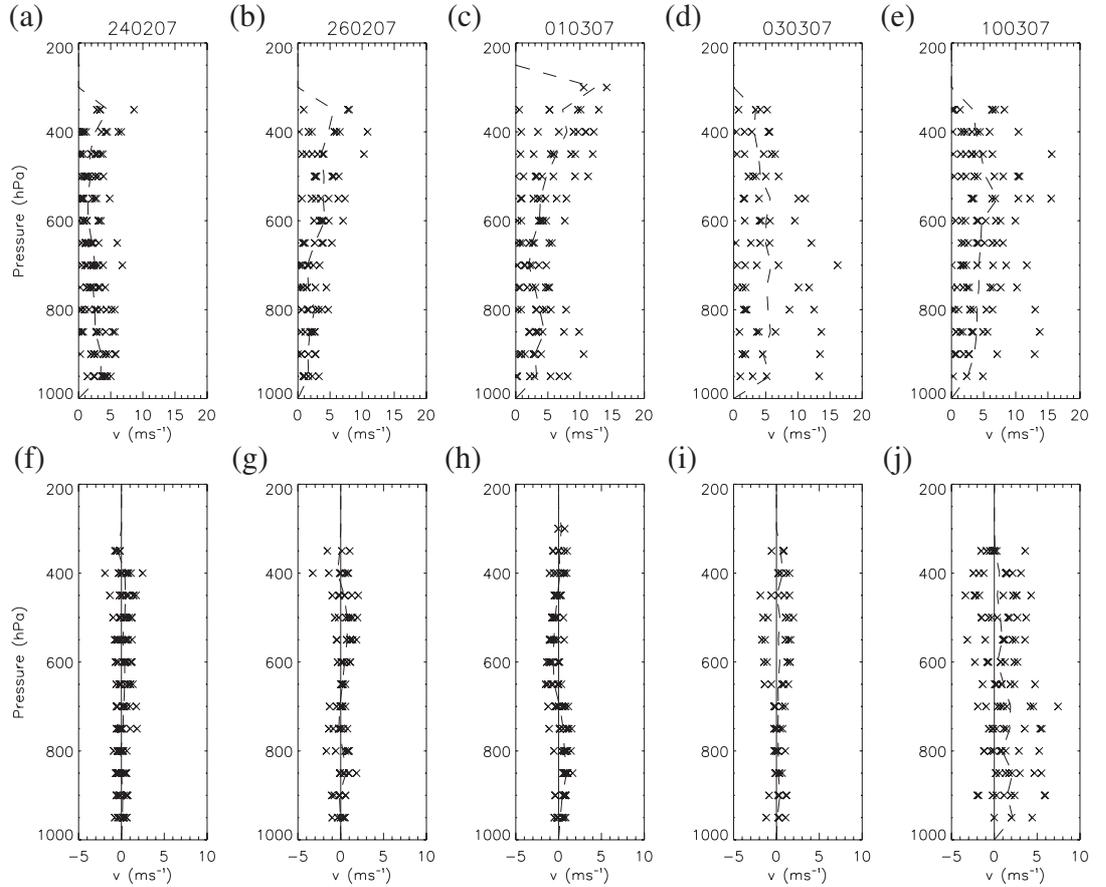
Figure 4.2: *Sonde-model comparison for v-component of the wind for each of the five cases (dates given above the top panels). The top panels (a) to (e) show the difference (4.1) between the CONTROL forecast and sondes for each case. The bottom panels (f) to (j) show the adjustment (4.2); this is positive if the forecast containing the dropsonde information is closer to the dropsonde data than the CONTROL forecast and negative if it is further away. Each cross represents data from one sonde at that pressure level, and the dashed line represents the average value. Note that the altitude of the aircraft varied both within and between flights therefore sondes were released from different altitudes.*

The CONTROL model data have been interpolated to the position that the dropsonde was released from, and both model and dropsonde data have been interpolated onto common pressure levels for ease of comparison. Here, hourly model output was used, taking the model output closest to the time of the dropsonde release, so that the background field should be representative of the true atmospheric state at the time the sonde data was collected. It is seen that the *difference* is larger for the last three cases, and in particular the last case, where the *v*-wind difference *difference* is as high as 15ms^{-1} (figure 4.2(e)), and the upper-level temper-

ature difference is as high as 6°C (not shown). It is likely that a larger *difference* will have a larger impact on the forecast.

The model *adjustment* towards the sonde data, shown in Figure 4.2(f)-(j), is defined as

$$Adjustment = \mid CONTROL - SONDE \mid - \mid TNOMEM - SONDE \mid .$$

Positive values of *adjustment* mean that the model profile from the forecast containing the targeted observations (TNOMEM) is closer to the sonde observations than the profile from the CONTROL forecast which does not include targeted observations. It implies that the data assimilation scheme has made use of the data to modify the background field, and negative or zero values of *adjustment* imply that it has not. It is seen that for most cases the *adjustment* is near zero, although with a slight positive bias, which for 01 and 03 March cases where the *difference* is large may indicate that the data assimilation system has been unable to use the information contained in the targeted observations. For the 24 February and 26 February cases, the average *difference* is below 5ms^{-1} and the *adjustment* is correspondingly small. This suggests that the targeted sondes will have little impact on the forecast in these cases. The 10 March case shows a relatively large *difference* and large *adjustment* with an obvious positive bias.

To quantify whether the dropsonde observations had a positive or negative impact on the forecast for the verification regions, the forecasts must be verified against some measure of the true state of the atmosphere at verification time. Model analyses from the ALL DATA forecast run were used as the best estimate of the 'truth'. Following most previous targeting studies, model analyses were used instead of observations due to the poor and irregular spatial coverage of profile observations, such as radiosonde data, available to verify the forecasts against. The forecasts for one case were also verified against operational ECMWF analyses, interpolated to the same grid and resolution, as a comparison. It was found that the improvement was similar in magnitude and location to the improvement seen when MetUM analyses were used to verify the forecasts.

The results are presented here in terms of total energy. Total energy was chosen to measure the forecast improvement, because it is used as a metric in calculating the TESV- and ETKF-SAPs. The total energy of the deviation of a forecast from the analysis can be calculated

as

$$TE_{F-A} = \frac{1}{2}(u_{F-A}^2 + v_{F-A}^2) + \frac{1}{2}\frac{c_p}{T_{ref}}(T_{F-A}^2), \qquad (4.2)$$

where u is the east-west oriented wind component, T is the temperature, and the subscript $F-A$ indicates a forecast minus analysis difference. Here the reference temperature T_{ref}=300K and c_p is the specific heat capacity. The quantity TE_{F-A} was calculated at 850, 500 and 250hPa levels, and the results summed. These levels were chosen to be consistent with the levels used in calculating the ETKF-SAPs. An impact metric I is then defined, in terms of total energy, such that $I > 0$ implies that the targeted observations have improved the forecast but $I < 0$ implies that they have degraded it:

$$I = TE_{C-A} - TE_{T-A}, \qquad (4.3)$$

where C is the CONTROL forecast, A is the verifying analysis (from the ALL DATA hindcast run, see Section 4.3.2) and T is the forecast containing targeted observations, which for the results in this section is TNOMEM.

Figure 4.3 shows the impact as a result of assimilating the targeted data, using (4.3) averaged over the verification regions. For all cases, and in both verification regions, the average impact of targeted data on the forecasts is small (Figure 4.3(a) and (c)), which may be related to deficiencies in the ability of the data assimilation to make use of the targeted data, as previously discussed. There is no improvement to the forecast in the north-west Europe verification region (Figure 4.3(a)) at the forecast times that the SAPs indicated sensitivity for (24-48 hours) for either the 24 and 26 February or 10 March cases. The 01 March and 03 March cases show some improvement at longer forecast times, but on inspection of the individual forecasts it is not possible to conclusively attribute this to the targeted observations. For the Scandinavian verification region, Figure 4.3(c), there is clear improvement ($I > 0$) in the 01 and 10 March cases at the 24 hour forecast time. The forecast improvement for these cases grows from near zero at target time to a peak at around 24 hours forecast time after which I decreases, so that the forecast is degraded after approximately 30 hours forecast time. Upon inspection of the spatial variation in impact, it is only for these two cases that it is possible to track the impact from the target region to the verification region. The mechanism
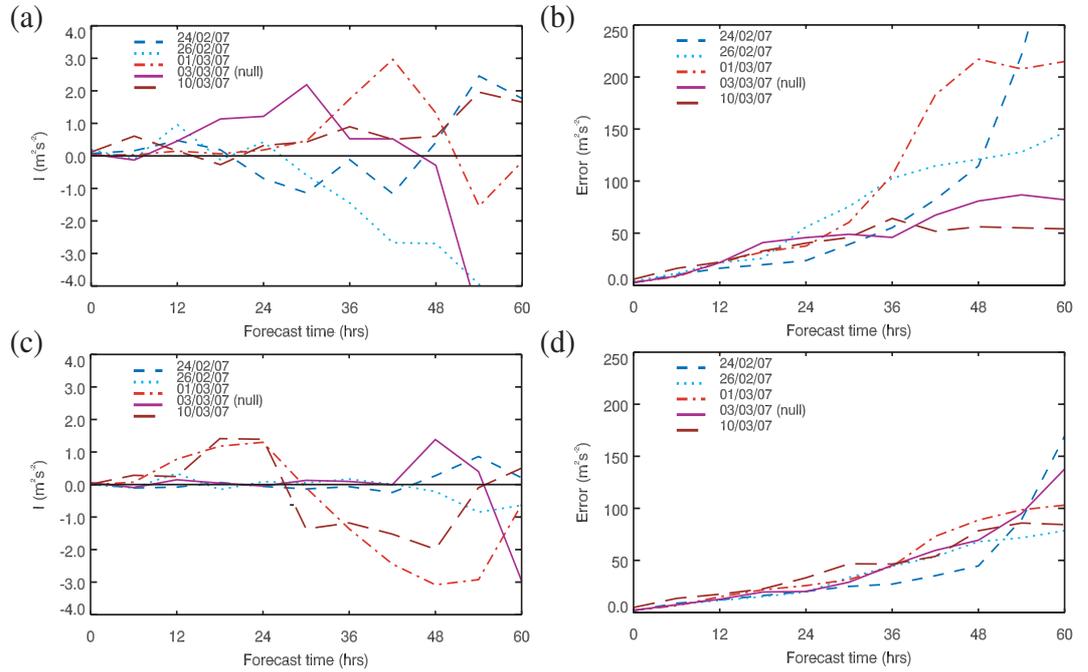
Figure 4.3: *Impact, I (4.3), of the targeted sondes for the TNOMEM forecast (panels (a) and (c)) and the forecast error (of the CONTROL forecast) during the GFDex (panels (b) and (d)). The top panels show the results for the north-west Europe verification region, and the bottom panels for the Scandinavian verification region. Note the difference in scale used for the y-axis for I and Error.*

for this forecast impact for these cases are examined in detail in Sections 4.5 and 4.6. Small forecast impact is seen for the other two targeted cases and the null case for this verification region. Recall that for the 24 and 26 of February the *difference* was smaller than for the other cases, suggesting that there were only very small errors in the initial conditions in the target region (the forecast background well-represented the state of the atmosphere). It is therefore unsurprising that there is only small forecast impact for these cases.

To put into perspective the magnitude of the impact of the targeted observations, the magnitude of the forecast error (defined as (4.4)) within the verification regions, which the targeted observations aim to reduce, is examined.

$$Error = TE_{C-A}. \tag{4.4}$$

Figure 4.3(b) and (d) shows that the forecast error increases with time, to 30-40m^{2}s^{-2} after 24 hours optimisation time for the 01 March and 10 March cases, which is an order of magnitude

larger than the modulus of the forecast impact (Figure 4.3(a) and (c)) at that time. The forecast error is non-zero at target time because the CONTROL forecast is verified against analyses from the ALL DATA run. Therefore there are differences in the number of observations both assimilated and in the background fields. The operational NAE forecasts generally verified well during this time and the forecast error and therefore potential to improve the forecast was small. The maximum reduction in forecast error due to the targeted observations was 5% at 24 hour forecast time for the 01 March case, and the forecast degradation reached around 3% in several cases.

4.5   Targeting Case Study: 01 March 2007

In this section the impact of targeted observations for the 01 March case is examined in greater detail. An overview of the case is given, including the synoptic situation at target time. The initial increments to the background field are then examined from a dynamical perspective, to understand how the targeted observations modified the background field, and how the impact of the targeted observations propagated through the forecast.

4.5.1   Overview

At targeting time on 01 March, the surface pressure chart (Figure 4.4(a)) showed a small Greenland lee cyclone (labelled 'G' on Figure 4.4(a)) to the south-west of Iceland, and weak barrier flow through the Denmark Strait. The barrier flow was forecast to strengthen as a low pressure system, south-east of Cape Farewell at this time, moved north towards Iceland and deepened (see Petersen et al. (2009) for further details). At upper levels there was westerly flow on the south-side of a trough positioned over north-east Greenland and extending out into the Greenland Sea (not shown). This, with an associated upper-level PV anomaly, was providing upper-level forcing for polar low development (labelled 'P' on Figure 4.4(a)) in the Greenland Sea, within the cold-air outbreak that extended as far south as the UK.

The TESV-SAPs indicated the sensitive region for short-range forecasts over north-west Europe and Scandinavia to be west of Iceland, over the Denmark Strait and Irminger Sea (Figure 4.4(a)). This corresponds with the location of the weak barrier winds and lee cyclone. By contrast, the ETKF-SAPs (Figure 4.4(b)) showed little forecast sensitivity to additional observations in this area, with signal variances of only $0.2\text{m}^2\text{s}^{-2}$ for both verification regions for
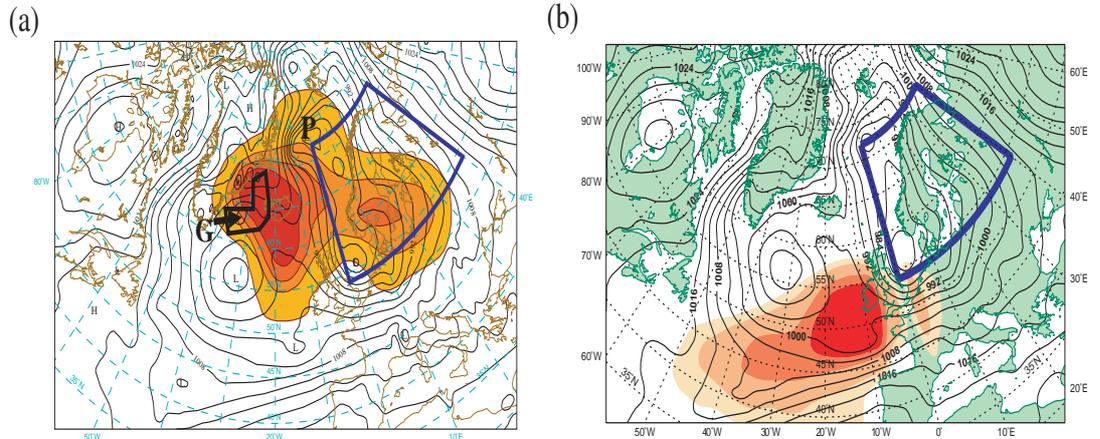
Figure 4.4: *The (a) ECMWF TESV-SAP used to plan the 01 March flight and (b) Met Office ETKF-SAP for the same day. The SAPs show the locations of the sensitive regions, where the addition of sonde data could improve the forecast for the Scandinavian verification region (36hr lead time, 24hr optimisation time). The shading indicates the sensitive regions, with the most sensitive 1, 2, 4 and $8\times10^6 km^2$ of the forecast domain shaded. The darkest shading indicates the region of maximum sensitivity. The black contours are the 36-hour forecast of mean sea-level pressure valid at targeting time. The flight track and Scandinavian verification region are outlined. In (a) P indicates the position of the developing polar low at targeting time, and G the location of the Greenland lee cyclone.*

SAPs with a 24-hour lead time. This was below the average seen during the field campaign, and so the ETKF-SAPs were not used for flight planning for this case. The flight released 9 dropsondes for targeting, equally spaced along the flight track. Of these 9 sondes, one sonde failed completely and therefore data from only 8 sondes was transmitted onto the GTS. An additional 9 sondes were released during the flight to study the barrier wind and lee cyclogenesis. This additional sonde data was assimilated into the ALL DATA run, along with the targeted sonde data.

4.5.2   Use of Targeted Sonde Data by the Assimilation Scheme

The 4D-Var assimilation scheme used to assimilate the sonde data seeks to find an increment to the background state that best minimises a cost function over a six-hour period. Increments are then applied to model-grid oriented wind components (u_m and v_m), potential temperature, exner pressure, total water and ozone, at the start of the forecast, 0900 UTC. The analysis increment therefore indicates how much 'new' information was given to the data assimilation scheme by the targeted data. In this context, 'new' information may be informa-
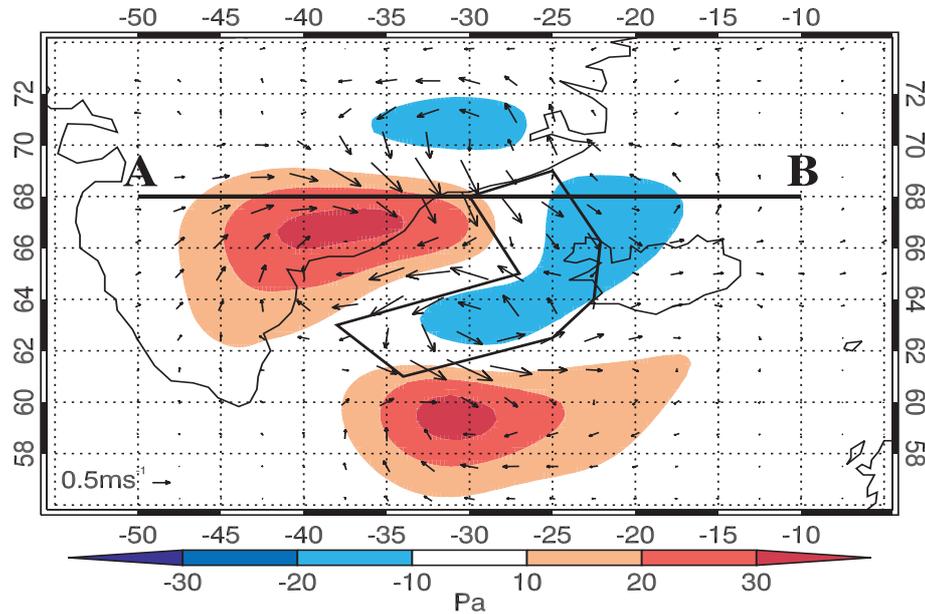
Figure 4.5: *TNOMEM minus CONTROL analysis increment for the 01 March case: pressure (shaded) and wind strength and direction (arrows). The wind differences are a pressure-weighted vertical average over all 38 model levels, and the pressure increment difference is shown at model level 16 (approximately 500hPa). The flight track is overlaid. The position of the cross-section displayed in Figure 4.6 is shown by the line connecting points A and B.*

tion which is not already contained in the background field, or information that is not provided by routine observations used to constrain the background field. The size of the analysis increment is dependent on the magnitude of the difference between the background field and observations.

The difference between the CONTROL and TNOMEM analysis increments (Figure 4.5) shows the initial impact of the targeted sonde data. The background used in both forecasts is the same, and the only differences are that the TNOMEM forecast has assimilated targeted sonde data in addition to the routine observations assimilated by the CONTROL forecast. Therefore any differences between the analysis increments for these two forecasts identify where the targeted sonde data is being used to constrain the background field. The impact of the targeted data is localised around the region where the sondes were released. There is a cyclonic wind increment in the southern portion of the flight track, in the region of the small lee cyclone. The effect of the targeted observations on the background at this location is to strengthen this lee cyclone. This is consistent with the pressure increment, which is negative

over the lee cyclone indicating that the cyclone is too weak in the CONTROL forecast.

The vertical structure of the TNOMEM-CONTROL analysis increments and CONTROL forecast background fields were examined along an east-west section at 68°N. The background fields (Figures 4.6(a) and (c)) show an approximately zonally oriented jet at 350hPa near the coast of Greenland and a weak barrier flow at 950hPa. The size of the analysis increments (Figures 4.6(b) and (d)) is small, relative to the background winds. The TNOMEM increment is strengthening the *u*-component of the flow and enhancing the southward component of the low-level flow near the coast of Greenland. The *v*-component of the flow is weak, and thus small analysis increments have a larger impact on the background fields.

There is no strong evidence of upstream vertical tilting in the increments, which would be an indication of a structure conducive to baroclinic growth. The lack of tilt may therefore indicate a barotropic structure, as obtaining a tilted increment would require that the observation minus first guess field tilt with height. This pattern of increments also reflects the assumed background error covariance structures used in the data assimilation scheme, which vary little with height.

4.5.3   The Impact of Targeted Sonde Data on the Forecasts

It was seen in Section 4.4 that the impact of sonde data on the forecast of the vertically summed total energy was small when averaged over the verification region. Figure 4.3(c) showed that for the 01 March case, a small improvement was seen in the first 30 hours. To be able to say that the improvement was due to the influence of the targeted sonde data, it is necessary to track the growing improvement from the target region at target time to the verification region at verification time.

Figure 4.7 shows the TNOMEM forecast impact during the first 36 hours of the forecast, calculated using (4.3). The forecast impact can be split into two components: a component associated with the lee cyclone which grows in-situ, and a second component associated with an upper-level trough and jet streak and which is advected into the verification region with a developing polar low, initially located at 72°N, 13°W. Small improvement is seen in the target region at analysis time (Figure 4.7(a)). Thus the targeting has succeeded in reducing the initial condition error in the target region. Most of the initial improvement is in the southern part of the target region and is associated with the lee cyclone. This impact grows after 12 hours
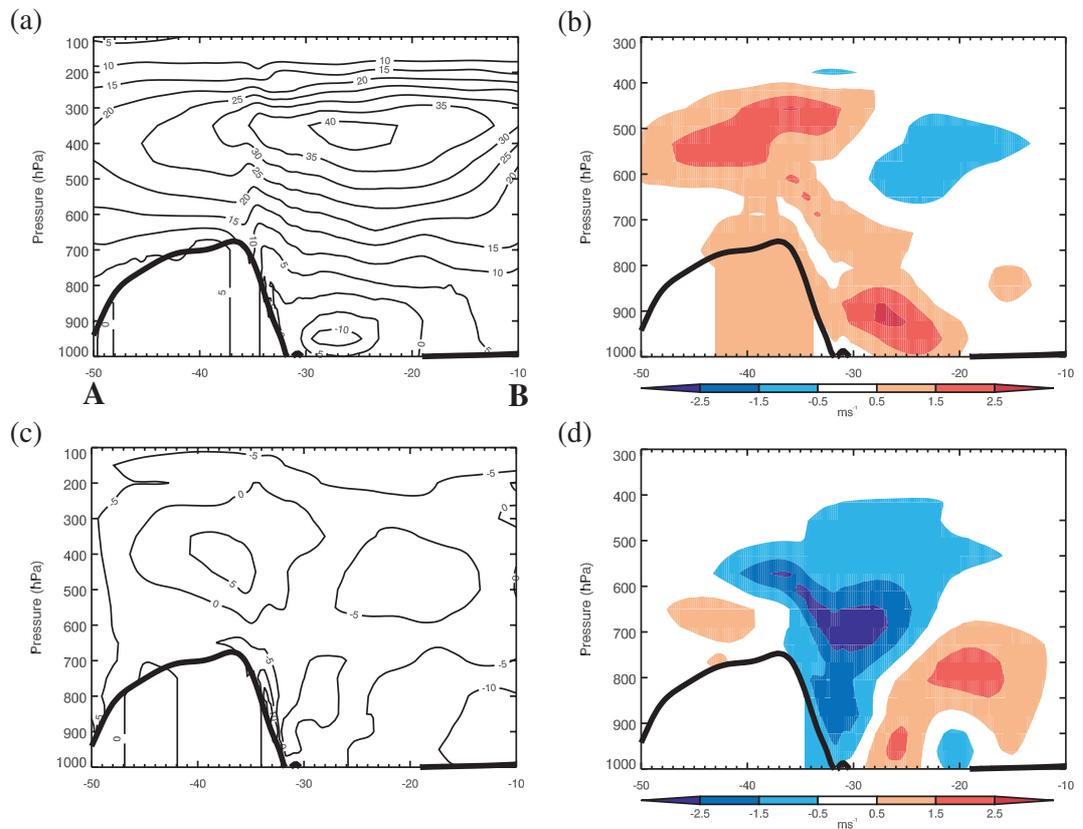
Figure 4.6: *Cross-sections through the Denmark Strait at 68°N from 50°W to 10°W for 01 March case. The left panels show the background field used in the data assimilation system for (a) the u-component of wind and (c) the v-component of wind. The right-panels show the TNOMEM minus CONTROL analysis increments of (b) u-component of wind and (d) v-component of wind. The thick black line is the surface pressure, used to show the orography; the cross-section cuts through Greenland on the left. The position of A and B relative to the flight track is shown on Figure 4.5*

but does not propagate into either verification region. Instead it improves the forecast of a cyclone which, initially located south of the lee cyclone (Figure 4.4(a)), moves into the lee of Greenland and deepens rapidly. Figure 4.3(a) showed no impact in the north-west Europe verification region until 36 hours into the forecast. It is seen in Figure 4.7 that no impact moves from the target region into the north-west Europe verification region, which explains the lack of impact initially. There is some impact at 36 hours to a baroclinic wave as it moves eastwards to the south of the UK. However this impact does not appear to be linked to the target region, and therefore the impact seen may be due to random noise in the initial field and cannot be attributed to the targeted observations.
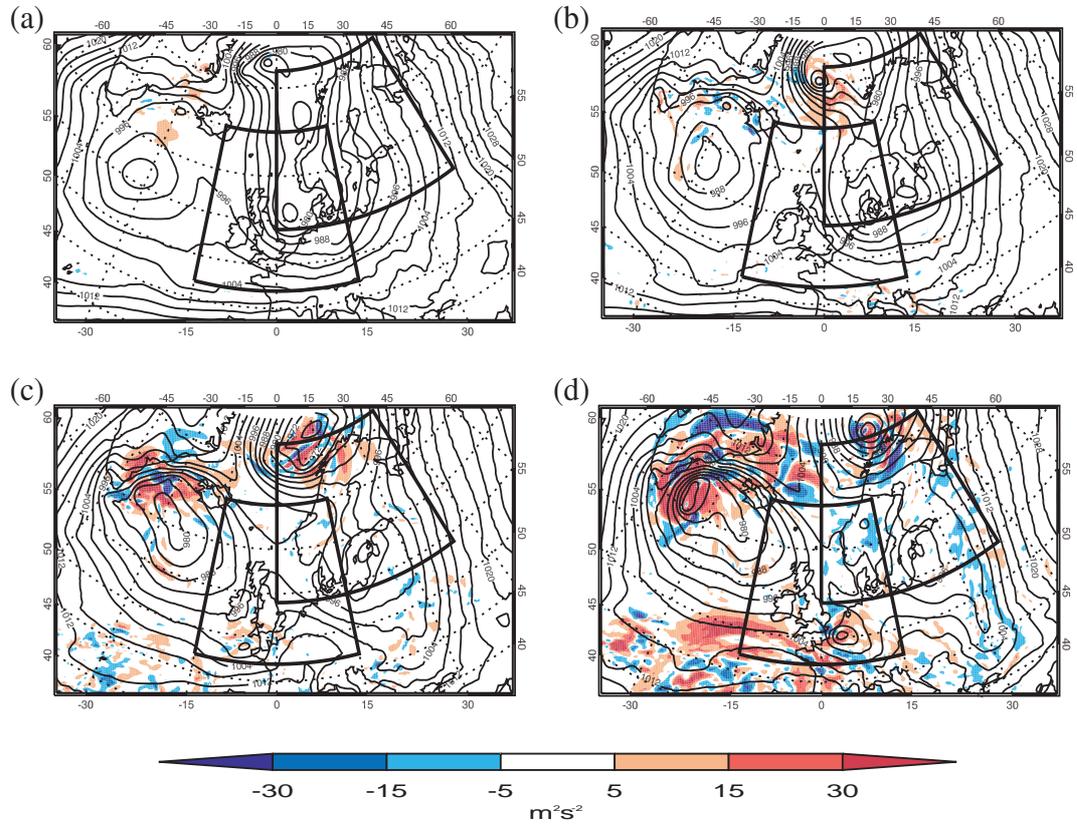
Figure 4.7: *Impact, I (4.3), of the targeted sondes for the TNOMEM forecast of the 01 March case at (a) 0, (b) 12, (c) 24 and (d) 36-hours optimisation times, shown by shading. The forecast is improved if the impact is positive (red shading) and degraded if the impact is negative (blue). The CONTROL forecast of mean sea-level pressure is contoured. The verification regions are outlined.*

At target time there is also small improvement on the eastern coast of Greenland (Figure 4.7(a)). This propagates east into the verification region with the developing polar low, with the improvement mostly confined to the warm sector of the polar low (Figure 4.7(b)). The improvement is largest 24 hours after targeting (Figure 4.7(c)), with values greater than $30 \text{m}^2\text{s}^{-2}$. The impact then moves out through the northern boundary of the verification region with the decaying polar low by 36 hours after targeting (Figure 4.7(d)), after which time the forecast is degraded.

The targeted sondes sampled the southern part of the upper-level trough and its associated upper-level PV anomaly, and the impact was advected at upper-levels by the PV anomaly, which moved with the polar low into the verification region, and caused some forecast improvement. This can be clearly seen by examining a vertical cross-section along $68°$N through
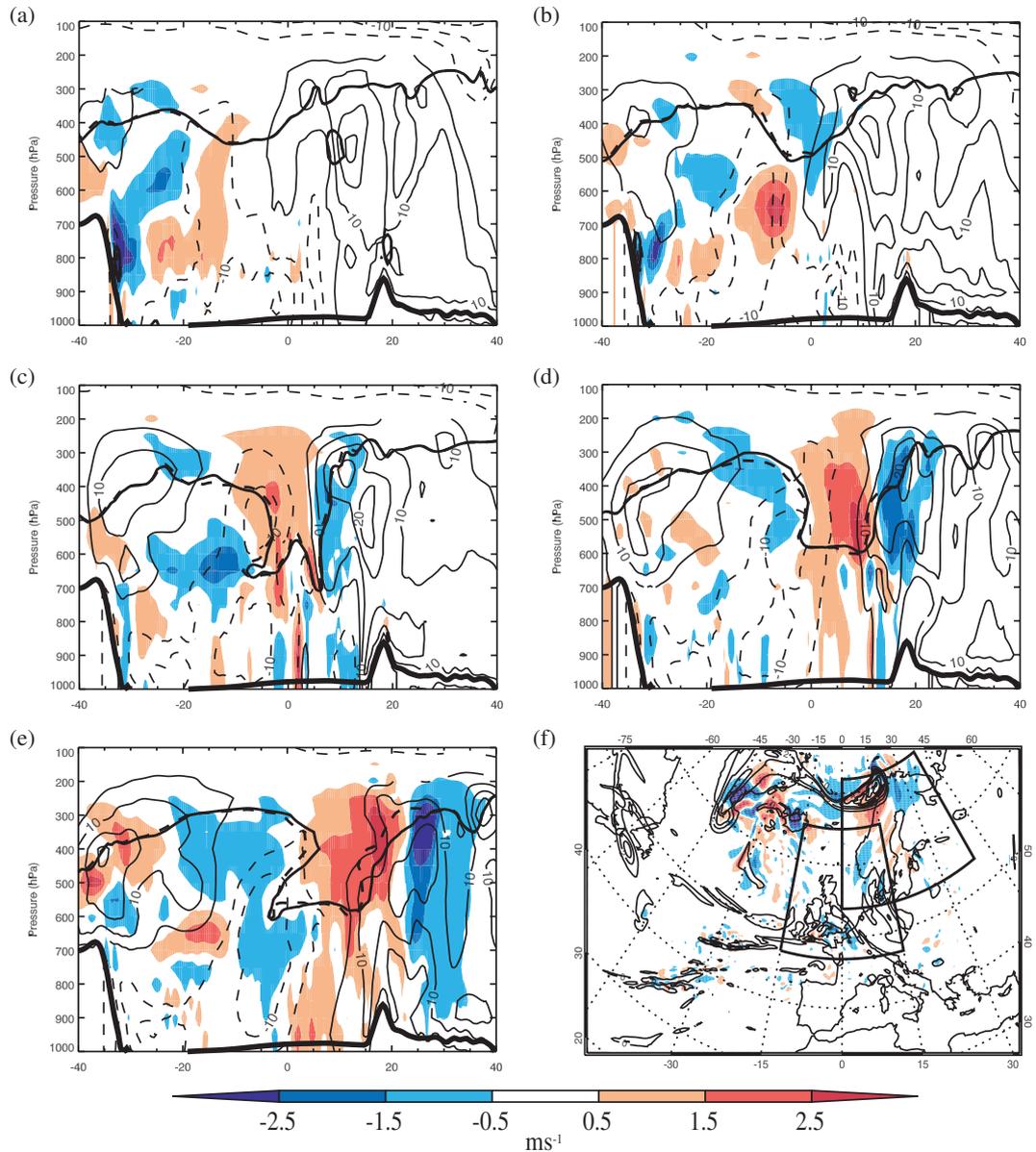
Figure 4.8: *Shading in panels (a) to (e) shows the TNOMEM minus CONTROL forecast of the v-component of wind at 68°N from -40°E to 40°E at (a) 0, (b) 6, (c) 12, (d) 18 and (e) 24-hours optimisation time for the 01 March case. The CONTROL forecast of v is contoured every $10ms^{-1}$ with thin solid (southerly) and dashed (northerly) lines. The dynamical tropopause is shown using the PV=2PVU contour, where the CONTROL forecast is the bold solid line and the TNOMEM forecast is the bold dashed line. The surface pressure (thick black line) is used to indicate the orography, showing Greenland on the left of the cross-section and Norway on the right. Panel (f) shows the TNOMEM minus CONTROL forecast of v at 500hPa at 24-hours optimisation time, using the same scale for the contours as for panels (a) to (e). The PV field at 500hPa is contoured with a contour interval of 1PVU.*

Figure 4.9: *Schematic diagram to show how targeted sondes modified an upper-level PV anomaly and associated wind field. The left panel shows the cyclonic circulation associated with an upper-level anomaly to the dynamic tropopause in the CONTROL forecast (solid line). If the eastern edge of the anomaly is moved to the west in the TNOMEM forecast (dashed line) then the associated circulation is also moved. The resulting difference in v-wind field has a dipole structure, as shown on the right panel.*

the forecast domain to see how the impact of the targeted data moves from the target region to the verification region. At target time the difference between the TNOMEM and CONTROL forecast is located mainly at lower levels at around 750hPa on the eastern side of Greenland (Figure 4.8(a)). This propagates upwards and to the east so that after 12 hours optimisation time the maximum difference is located at the dynamical tropopause and associated with the PV anomaly (tropopause fold) that is forcing the development of the polar low (Figure 4.8(c)). The difference at this time is also associated with a warm anomaly at the surface, particularly to the north of the system where the tropopause fold is shallower and there is a strong zonal gradient in wind speed from the calmer core of the polar low to the strong low-level northerly jet to its west (not shown). The difference is initially located on the western side of the PV anomaly, and moves around the southern edge of the PV anomaly into the verification region (Figures 4.8(e) and (f)) as the anomaly moves eastwards. The inclusion of targeted sonde data in the forecast caused the eastern edge of the PV anomaly to be shifted west, by up to 200km (Figure 4.8(e)). This caused the cyclonic circulation associated with the PV anomaly to be shifted to the west, so that when the CONTROL and TARGETED forecasts are differenced, a dipole structure is seen in the *v*-wind field. This is illustrated schematically in Figure 4.9. The mechanism that caused the impact in the Scandinavian verification region is therefore the modification of the position of a PV anomaly associated with an upper-level trough, in a region of baroclinic instability. The interaction of this upper-level trough with an occluded front led to the development of a polar low, and the impact moved east with this system into
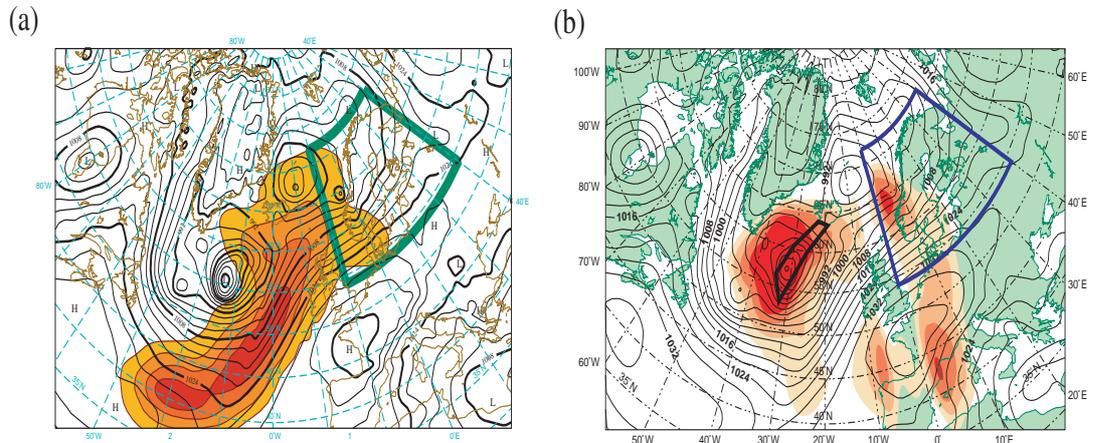
Figure 4.10: *The (a) ECMWF TESV-SAP and (b) Met Office ETKF-SAP used to plan the targeting flight. The SAPs show the locations of the sensitive regions, where the addition of sonde data could improve the forecast for the Scandinavian verification region (36hr lead time, 24hr optimisation time). The shading indicates the sensitive regions, with the most sensitive 1, 2, 4 and $8\times10^6 km^2$ of the forecast domain shaded. The darkest shading indicates the region of maximum sensitivity. The black contours are the 36-hour forecast of mean sea-level pressure valid at targeting time. The flight track and Scandinavian verification region are outlined.*

the verification region.

4.6 Targeting Case Study: 10 March 2007

A brief analysis of the 10 March case is provided here to contrast with the 01 March case. Forecast improvement of similar magnitude to the 01 March case was seen in the first 30 hours in the Scandinavian verification region (Figure 4.3(c)), but the mechanism for the impact appears to be different.

4.6.1 Overview

At target time a synoptic-scale cyclone was located south-west of Iceland (Figure 4.10(b)), and was moving rapidly towards Iceland as it developed. During flight planning it was noted that there was considerable uncertainty in the forecasts from different forecast centres as to the location and strength of the cyclone at targeting time. This made planning a flight track difficult and resulted in the aircraft flying directly over the centre of the cyclone rather than to the west of the centre as planned. The cyclone was located on the poleward side of the upper-level jetstream at targeting time, and although the centre of the cyclone was forecast to move
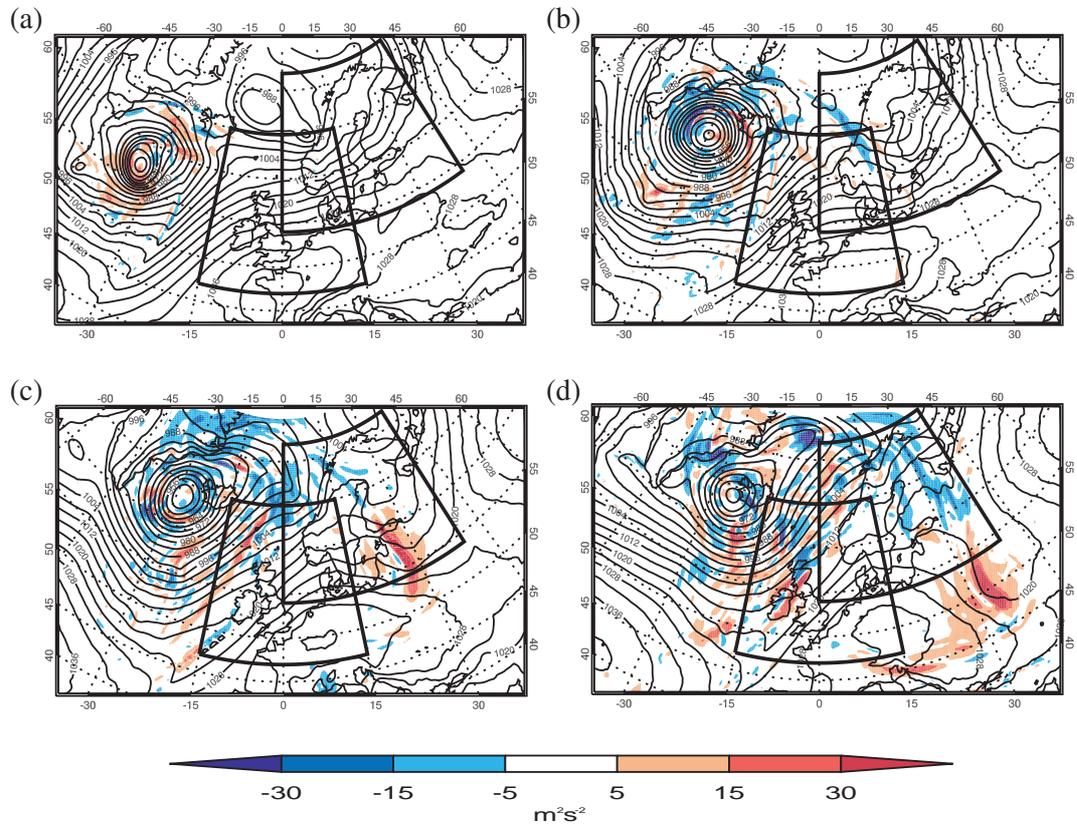
Figure 4.11: *Impact, I (4.3), of the targeted sondes for the TNOMEM forecast of the 10 March case at (a) 0, (b) 12, (c) 24 and (d) 36-hours optimisation times. The forecast is improved if the impact is positive (red shading) and degraded if the impact is negative (blue). The CONTROL forecast of mean sea-level pressure is contoured. The verification regions are outlined.*

over Iceland, the trailing fronts were forecast to move over the UK, through the north-west European verification region. An overview of this case is also given in Renfrew et al. (2008).

The ETKF-SAPs with a 36 hour lead time identified the centre of the low pressure system to be most sensitive for the forecasts for Scandinavia and north-west Europe (Figure 4.10(b)), and these were used for flight planning. The TESV-SAPs (Figure 4.10(a)) identified both the eastern edge of the low pressure system and a region to the far south, in the mid-Atlantic. The aircraft was not able to sample both sensitive regions; the ETKF-predicted sensitive region was chosen as this was the main sensitive region for both verification regions.

4.6.2   The Impact of Targeted Sonde Data on the Forecasts

The impact calculated using (4.3) for the 10 March case is shown in Figure 4.11. The impact to the 1200 UTC analysis is mostly positive (Figure 4.11(a)), and confined to the central and northern areas of the cyclone (note that the cyclone was at this stage moving slowly in a northerly direction). The cyclone stalls over Iceland as it begins to decay, and much of the impact from the targeted sondes remains with the cyclone at the later forecast times (Figure 4.11(c) and (d)). After 12-hours forecast the impact also begins to spread out along the fronts associated with the cyclone (which lie along kinks in the mean sea level pressure fields). The impact is greatest at upper-levels (not shown) and appears to propagate downstream along the jetstream in a manner similar to a jetstreak. The impact amplifies rapidly in the downstream trough and leaves the domain to the south. This case is discussed in more detail in the following section.

4.7   The Effect of Cycling Observations in the Forecast Background

The results presented thus far have focussed on the TNOMEM targeted run. This is because it provides a clean comparison with the CONTROL run, as the only differences between the two forecast runs are the addition of targeted data to the TNOMEM forecast for a particular targeting case. This also represents a situation where targeting is done infrequently, i.e. less than once a week. If targeting were done more frequently, for example if a similar area was targeted twice within a week, then perturbations from the previous targeting may still be impacting the forecast. As a 6-hour forecast from the 1200 UTC forecast is used as the background to the 1800 UTC forecast and so on, targeted observations can impact subsequent forecasts through the effect of memory in the forecast system.

The impact of assimilating the targeted data in the TMEM forecasts is shown in Figure 4.12 (Figure 4.12(a), (b) are equivalent to Figure 4.3(a), (c) but are for the TMEM forecast instead of the TNOMEM forecast). By comparison with Figure 4.3 it is seen that with memory of previous targeted observations in the forecast the magnitude of the forecast impact is increased at some forecast times, meaning that the TMEM forecasts are more different from the CONTROL forecast. The exception to this is the 24 February case, the first target-

(a)

(b)
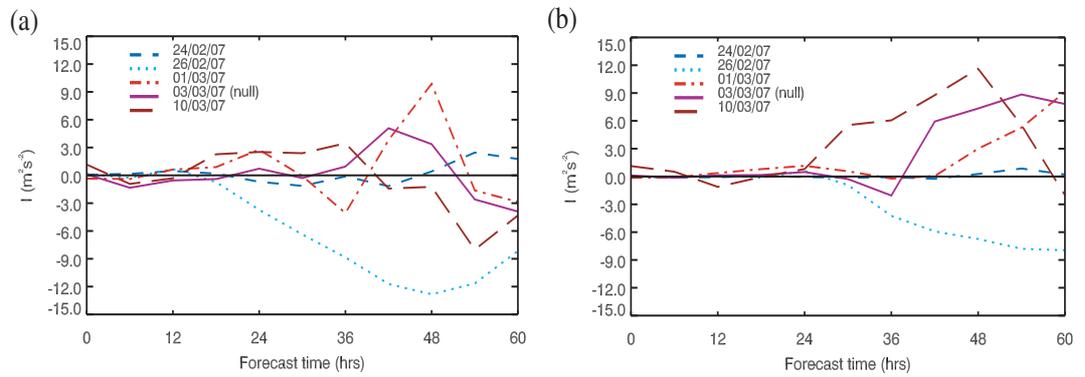
Figure 4.12: *Impact, I (4.3), of the targeted sondes for the TMEM forecast for (a) north-west Europe and (b) Scandinavian verification regions. Note the difference in scale between this figure and panels (a) and (c) of Figure 4.3.*

(a)

(b)

(c)

(d)
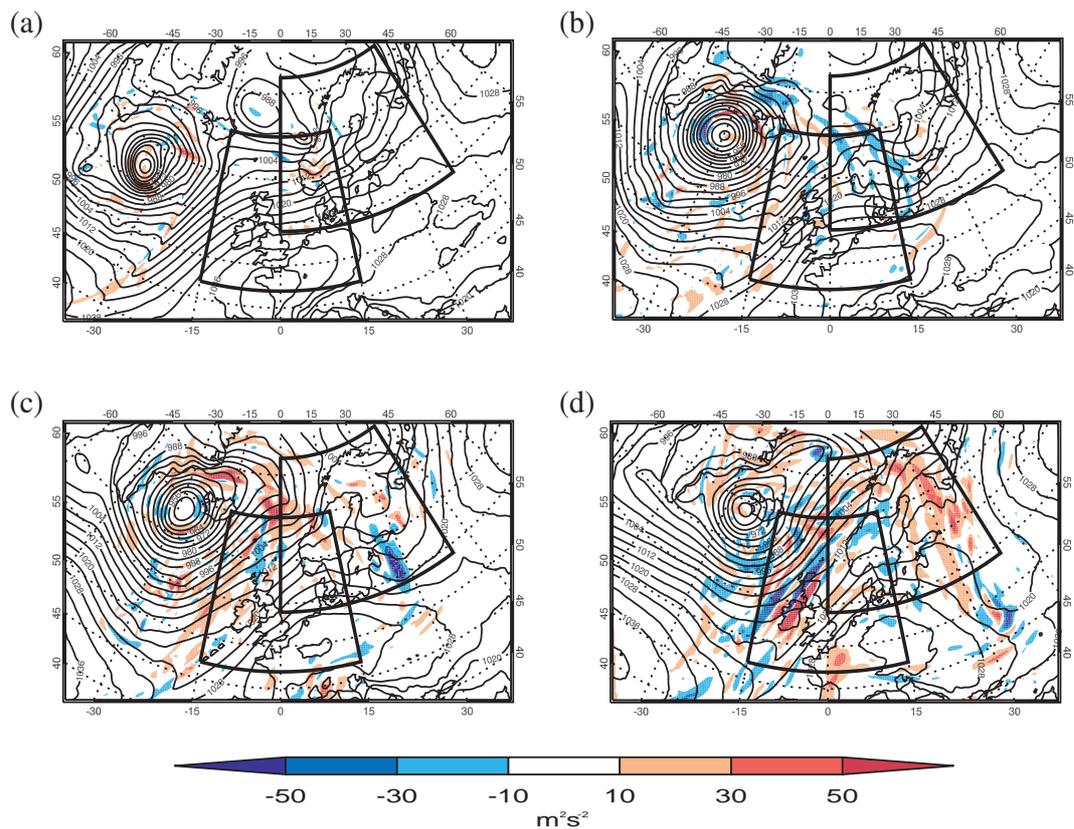
Figure 4.13: *Impact, I (4.3), of the targeted sondes for the TMEM forecast of the 10 March case at (a) 0, (b) 12, (c) 24 and (d) 36-hours optimisation times. The forecast is improved if the impact is positive (red shading) and degraded if the impact is negative (blue). The CONTROL forecast of mean sea-level pressure is contoured. The verification regions are outlined. Note the difference in scale between this figure and figure 4.11*

ing experiment of the GFDex, where the TMEM and TNOMEM forecasts are, by definition, identical. It can be seen that for the 26 February case the level of forecast degradation is increased, whereas for the final three cases the level of forecast improvement is increased in the Scandinavian verification region. It is not possible to conclude from this that the direction of the change is due to the memory of targeted observations, it is just as likely that the effect of cycling targeted observations is to introduce random perturbations into the background field, some of which grow.

Given that the hindcasts were run using a limited-area domain, it is perhaps surprising that there is such a large difference between the TNOMEM and TMEM forecast impacts for the 10 March case. As the 10 March case occurs seven days after the previous targeting case, it might be expected that any perturbations introduced into the forecast background from assimilating previous targeted observations might have moved out of the boundary of the limited-area domain, so that the TMEM and TNOMEM forecasts would show similar impact. Figure 4.13 shows the impact to the TMEM forecast calculated using (4.3) in the first 36-hours of the forecast for the 10 March case. By comparison with Figure 4.11 (note the difference in scale for the impact) it is seen that the magnitude of the impact is increased. At initial time the forecast impact is not confined to the area in which the sondes were released, but are spread along the fronts associated with the cyclone. This additional initial impact is due to differences between the TNOMEM and TMEM background fields (the TNOMEM and CONTROL background fields are the same). This shows that although the GFDex dropsonde observations had a small impact on the forecast, the perturbations from targeting seem to have persisted in the forecast background for seven days after targeting.

The inclusion of earlier sonde data in the background of the forecast does not necessarily lead to greater forecast improvement. Rather the trajectories of the CONTROL forecast and targeted forecasts differ more as the optimisation time increases when additional observational data are included in the background of the forecast as well as being assimilated into the forecast. It can be concluded from these results that differences in the background field may sometimes have a greater impact on the forecast than the analysis increment due to the assimilation of sondes. Cardinali et al. (2004) showed that up to 85% of forecast information is provided by the background field rather than observations within the 4D-Var time window. That perturbations still impact forecasts up to a week later is perhaps surprising given that the

Table 4.3: *Definition of Ensemble Forecast Members*

Forecast	Analysis Increment		
Background	TNOMEM	TMEM	ALL DATA
CONTROL	×	◇	◇
TMEM		×	
ALL DATA			×

The original simulations (Section 4.3.2) are denoted ×. These symbols are used to identify ensemble members in Figure 4.14.

perturbations are small, localised and are inserted into a limited area domain. The increased impact of targeted observations seen from the TMEM hindcast runs also has implications for studies of targeted observing. It implies that studies that only investigate the impact of targeted observations by running a 'TMEM' forecast and compare it to a control cycle containing no targeted observations may be over-estimating the direct impact of the targeted observations. This is particularly true for field campaigns, where targeting may be done many times over a short time span of a few weeks for example.

4.8 Robustness of the Forecast Improvement

It has been shown in previous sections that for two cases some improvement was seen at specific optimisation times as a result of assimilating GFDex sondes into the forecast. In this section the robustness of the improvement found in these two cases is inferred by seeing how sensitive the results are to the background field and analysis increment used to make the forecast.

An ensemble of five targeted forecasts was created (Table 4.3). In this case the CONTROL forecast is not counted as an ensemble member as this is what the perturbed forecasts are compared against. Three of the ensemble members are the ALL DATA, TMEM and TNOMEM forecasts (labelled as × in Table 4.3; recall that the TNOMEM forecast uses the background from the CONTROL forecast), detailed in Section 4.3.2. Two additional ensemble members (labelled as ◇ in Table 4.3) were created by adding the analysis increments from the TMEM and ALL DATA hindcast runs to the CONTROL background.

Figure 4.14(a) shows the impact for each ensemble member, calculated using (4.3) aver-

(a)

(b)
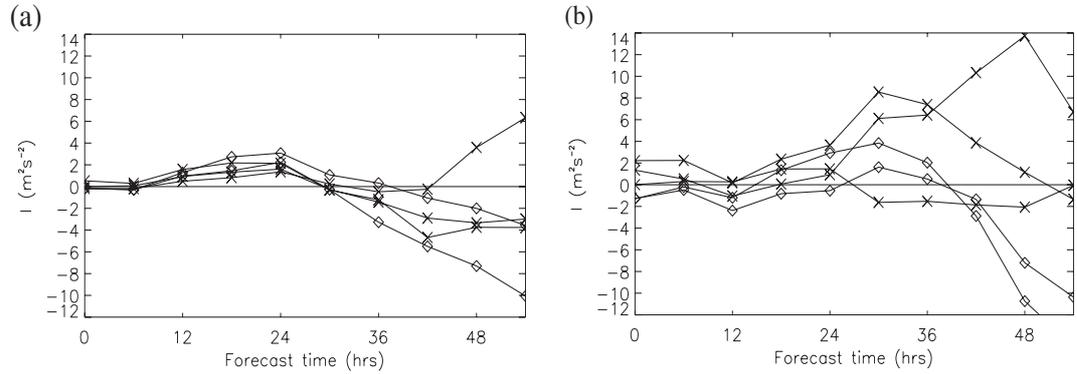
Figure 4.14: *Ensemble plume of the impact, I (4.3), averaged over the Scandinavian verification region for (a) 01 March and (b) 10 March cases. The symbols used to mark the different members are defined in Table 4.3.*

aged over the Scandinavian verification region, for 01 March case. The spread of the plume is small initially, with all members showing improvement in the short-range forecast for Scandinavia that was previously identified in the TNOMEM and TMEM forecasts. The spread of the plume exceeds the absolute value of the mean from 24 hours onwards, indicating that any improvement or degradation here is not robust, and due to differences in the background field rather than being caused by the targeted data. For the 10 March case (Figure 4.14(b)) not all members show improvement in the forecast between 18 and 24 hours after targeting which implies that the impact seen in Figure 4.3(c) is not positive definite.

The aim of the ensembles shown in Figure 4.14 is to estimate the ratio of signal to noise associated with the value of impact. For the ensembles shown in Figure 4.14, the ratio of standard deviation to ensemble mean at 24 hour forecast time is 0.3 for the 01 March case and 0.5 for the 10 March case. Although a small ensemble, its normalised spread is similar to the 23-member operational MOGREPS ensemble for these cases, which has a ratio of standard deviation to mean of 0.4 for the 01 March case and 0.6 for the 10 March case. It can be seen from Figure 4.14 that the average magnitude of impact is small. This is due to the small size of the perturbations associated with the targeted observations, which are roughly one-tenth of the size of those used to construct the MOGREPS ensemble. By design, the amplitude of a MOGREPS initial perturbation (not shown) is comparable to analysis error and its growth comparable to the growth of the forecast error over the 48 hour ETKF window. The forecast error for the GFDex cases (Figure 4.3(b) and (d)) is also comparable in magnitude to

the MOGREPS ensemble perturbations. However since the MOGREPS initial perturbations are created under assumption of linearised dynamics, which is valid at least for a 24 hour window (Gilmour et al., 2001), we know that if the initial perturbations were to be re-scaled by a constant factor (say 0.1) and re-run, the spread/mean would be unchanged. Therefore, assuming that the perturbations associated with targeted observations evolve dynamically in a similar manner to perturbations obtained by the MOGREPS ETKF system, the spread of impact for this small ensemble gives a reasonable indication of the uncertainty in the value of impact.

4.9   Conclusions

Targeted observations were made during the GFDex, in the region around Iceland and southern Greenland, with the aim of improving the 24-48 hour forecast over the UK and Scandinavia. The additional observational data were assimilated operationally into the Met Office forecasts. The impact of these targeted observations has been assessed by running hindcasts using a limited area model with a 4D-Var data assimilation scheme. There were a total of four targeted observing flights into TESV or ETKF predicted sensitive regions. One additional 'null' flight was also included in this assessment.

For two of the targeted cases, 24 and 26 February, the difference between the background and sonde data was small, therefore the dropsonde data gave little new information to the data assimilation system, resulting in little forecast impact. For the 01 and 10 March targeting cases the targeted dropsonde data gave new information to the forecast system and these data were, in general, used by the assimilation scheme to correct errors in the background field in the vicinity of the sondes. This resulted in a small improvement in the analysis at the target time in the target region. This improvement did not always propagate from the target region into the verification region, some improvement appeared to grow in-situ. However, the resulting forecast improvement was small; with a maximum forecast improvement of 5% and degradation of 3%.

The mechanisms for the forecast impact for the 01 and 10 March cases were analysed in more detail. For the 01 March case, the effect of the targeted sondes on the analysis increment was a deepening and spin-up of a low-level cyclone. It has been shown that the subsequent modification of the position of a tropopause fold and its associated wind field, which was forc-

ing the development of a polar low, caused an improvement to the forecast over Scandinavia. For the 10 March case the impact was also mostly seen at upper levels, in the downstream development of a trough and cut-off low in the Scandinavian verification region.

The magnitude of the dropsonde signal and the impact on the forecasts was larger when previous targeted observations were included in the background state. It was shown that for one case, 10 March 2007, this caused a larger forecast improvement (for the verification region) than assimilating targeted data for that day alone. It is likely that the impact of cycling observations in the forecast background is to introduce random perturbations to the background field, some of which grow, which may improve or degrade the forecast. The implication of this result for operational targeting is that previous targeting experiments may affect the success of subsequent targeting experiments on following days.

To determine the robustness of the forecast improvement seen in two cases a small ensemble of forecasts was created, by varying the background and analysis increment used to run each forecast. Using this method, the improvement to the Scandinavian verification region in the first 24 hours of the forecast for the 01 March case was shown to be a robust result. Using the same method it was shown that the forecast improvement seen for the 10 March case was not robust.

Overall, the impact of targeted observations seen in GFDex is both small and mixed, though generally positive or neutral for forecasts of lengths equivalent to the shortest optimisation time considered for the sensitive area predictions. This is consistent with other recent investigations of targeting over the north Atlantic.

CHAPTER 5

Covariance Experiments

In this chapter the methodology of the hindcast experiments presented in Chapter 4 is modified to see if the targeting experiments could be designed to make better use of the information contained in the targeted data. It was shown in Chapter 4 that the impact of these targeted observations was small, with a maximum forecast improvement in the Scandinavian verification region of only 5% at 24 hour forecast time, when verified against MetUM analyses. Here we aim to demonstrate that there is the potential to increase the impact of targeted observations. Results from two sets of experiments are shown here: in the first the dropsonde observation error covariances are reduced from their operational values to give the dropsonde observations greater weight during the assimilation process. In the second, different sets of targeted observations are assimilated, and the covariances also varied.

5.1 Background

The specification of observation errors is important because it determines the weight given to observations during the assimilation process. Recall from Chapter 2 that the cost function for 3D-Var (neglecting time dependence here for simplicity) can be written as (2.6):

$$J(\mathbf{x}) = J_b + J_o = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathbf{h}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}(\mathbf{x})). \tag{5.1}$$

Where J_b is the background term and J_o is the observation term. The observation term of (5.1) describes the penalty from the observations, $\mathbf{y} - \mathbf{h}(\mathbf{x})$, weighted by the inverse of the observation error covariance $\mathbf{R}$ which is the subject of this study. The inverse of the observation errors is used as a weighting, so that an observation with a small error is given more weight than an observation with a larger error. The observation errors are assumed to be uncorrelated, which reduces the observation error covariance matrix to a diagonal matrix, where the entries on the diagonal are simply the variance of the observation. With this knowledge, the

observation term J_o of (5.1) can be written as:

$$\frac{1}{2} \sum_{i=1}^{N} \frac{(y_i - h_i(\mathbf{x}))^2}{\mathbf{R}_{ii}}, \tag{5.2}$$

where N is the number of observations, y_i is the ith observation, $h_i(\mathbf{x})$ is the model estimate of that observation and $\mathbf{R}_{ii}$ the observation variance. It is clear from (5.2) that the magnitude of the difference between an observation and the model's estimate of that observation, $y_i - h_i(\mathbf{x})$, is weighted by the error variance in that observation $\mathbf{R}_{ii}$. An accurate estimate of $\mathbf{R}$ for each observation type is therefore crucial in determining the weighting given to observations during the assimilation process.

In the Met Office data assimilation system, dropsonde data are treated as radiosonde data, and assigned observation error profiles accordingly. The current radiosonde error profiles in operational use at the Met Office are based on observation minus background differences which were calculated using radiosonde and model data over ten years ago (personal communication, Richard Dumelow). The error profiles for radiosonde data from Northern Europe, which were still operational as of March 2007 during the GFDex field campaign, are shown in Figure 5.1. The temperature error (Figure 5.1(a)) is on the order of 1K, with little variation with height. The wind error (Figure 5.1(b)) varies from around 1.5ms^{-1} in the lower atmosphere to a peak of 2.8ms^{-1} at the jet level. The relative humidity (Figure 5.1(c)) has the largest error, of around 12%. This might be expected given that the method for determining the observation errors involves differencing the observation data and model data - if the modelled location of cloud differs to the actual location of cloud then differencing the observation and model data will give a large error.

The observation errors illustrated graphically in Figure 5.1 can be thought of as being composed of two parts: a measurement error and a representivity error. The measurement error is a measure of the accuracy of the instruments used to make the observations. For example, for the Vaisala RS92 dropsondes used during GFDex, Vaisala give a measurement error of 0.2°C for their temperature sensors and 0.5ms^{-1} for the wind sensors (Vaisala, 2006). The representivity error concerns how representative a point measurement is of the conditions in the rest of the grid box that it is used to represent. Since the calculation of the observation error statistics over 10 years ago, the resolution of the forecast models has increased and thus
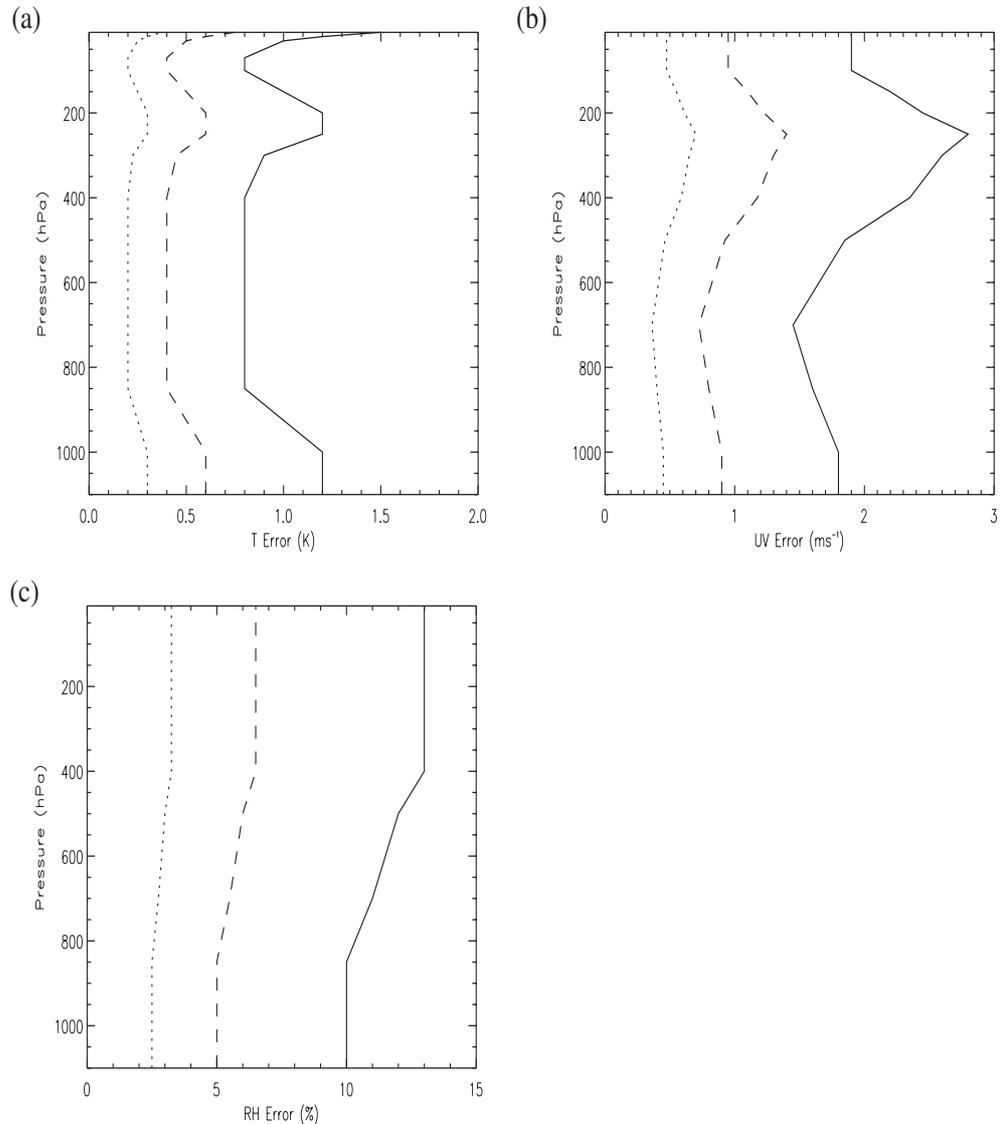
Figure 5.1: *Operational dropsonde observation error profiles (solid lines) of (a) temperature, (b) horizontal wind components and (c) relative humidity, used in the 4D-Var assimilation scheme at the Met Office. These were operational as of March 2007, when the GFDex experiments were conducted. Profiles of half (dashed line) and one quarter (dotted line) of the operational values are also shown.*

the representivity error will have decreased (i.e. a point measurement is likely to be more representative of the average state in a 24km by 24km grid box than a 100km by 100km grid box). Gross error is not accounted for in the observation error covariances; the data assimilation process makes the assumption that any gross error in the observations has been removed prior to assimilation. To satisfy this assumption, the observations undergo a quality control process before being passed to the data assimilation scheme, which uses Bayesian

statistics and removes gross error by assuming that good observations have a Gaussian error distribution and observations with gross error do not (Ingleby, 1993).

Targeted observations are made in sensitive regions, where errors in the initial conditions may grow rapidly and cause a deterioration of the forecast downstream of the initial sensitive region. As these sensitive regions are often remote and data-sparse, there is potential for the forecast background in these regions to contain initial condition error. Targeting observations in these regions has the aim of reducing the initial condition error, and to do this the model state must be adjusted towards the observations rather than the background. There are two ways of doing this. Firstly, one could increase the background error covariance in the sensitive region, so that the background is given less weight here (this is already under trial at the Met Office, personal communication Dale Barker). The other, simpler, method is to reduce the observation error covariance of the targeted observations so that they are given more weight during the assimilation process, and this is the method employed here. A comparison of the operational dropsonde observation error profiles and those calculated from the GFDex data is discussed in Section 5.2, which is used as justification for the reductions in dropsonde observation error variances used in the experiments. Hindcast studies were performed with reduced dropsonde observation error covariances and the set-up of these experiments is discussed in Section 5.3, results presented in Section 5.4 and conclusions in Section 5.5.

5.2 Calculation of Dropsonde Observation Error Profiles Using GFDex Dropsonde Data

To assess the suitability of the error profiles assigned to the dropsonde data, both operationally and in the reduced covariance experiments described in Section 5.3, dropsonde error profiles have been calculated for the GFDex data. The method followed is comparable to the method used by the Met Office to calculate the operational profiles, and compares the GFDex dropsonde and model data differences.

The observation errors ε_0 are the difference between the observation and the true state of the atmosphere at that point. As we do not know the true state we have to use an approximation, so that ε_0 can be expressed as:

$$\varepsilon_0 = \mathbf{y} - \mathbf{H}(\mathbf{x}_t), \tag{5.3}$$

where **y** is the vector of observations and $\mathbf{x}_t$ the model state from a six-hour forecast, which is our best estimate of the true state. $\mathbf{H}(\mathbf{x}_t)$ is the model's prediction of the observations, where $\mathbf{x}_t$ is interpolated to the position and time that the observation was made by the operator **H**. Thus the dropsonde observation errors are calculated by differencing the dropsonde observations with the model's prediction of the dropsonde observations. Assuming that the observation errors are uncorrelated reduces **R** to the variance of the observation errors. However, the variance of ε_0 is due to both observation error and forecast error (as the calculation of ε_0 from (5.3) involves the forecast model). Hollingsworth and Lönnberg (1986) verified short-range forecasts against radiosonde data and found that observation and forecast error have an equal contribution to the variance. Making this assumption, **R** becomes:

$$\mathbf{R} = \frac{1}{2}\langle(\varepsilon_0 - \langle\varepsilon_0\rangle)^2\rangle, \tag{5.4}$$

where the brackets $\langle\ \rangle$ indicate the expected value, i.e. the average. The observation error profiles shown in Figures 5.1 and 5.2 are the standard deviations of the observation error covariance, i.e. $\sqrt{\mathbf{R}}$.

The dropsonde error profiles for the GFDex dropsonde data were calculated using the above method. The GFDex dropsonde data from the 5 targeting cases studies detailed in Chapter 4 were used as **y**, and 6-hour forecasts containing all the dropsonde data (from the ALLDATA run, see Chapter 4) used as $\mathbf{x}_t$. Bilinear interpolation was used as the operator **H** to obtain an estimate of the model profile at the position of the sonde releases. The resulting error profiles for *u*, *v*, *T* and *RH* are shown in Figure 5.2. The error profile is presented only between 950hPa and 350hPa, as outside this range there were an insufficient number of data points to obtain a robust estimate of the error. The error at a particular pressure level was calculated only where there were greater than 50 data points, and a total of 71 out of 74 available dropsonde profiles were used to estimate the error profile (three profiles were excluded where the sonde had malfunctioned and the *u*, *v* or *T* profile was therefore incomplete).

Figure 5.2 shows the observation error profiles calculated for the GFDex dropsonde data. By comparing Figure 5.2(a) with Figure 5.1(a) it is seen that the temperature error profile for the GFDex sonde data has much smaller values than are currently used in the 4D-Var

(a)

(b)

(c)

(d)
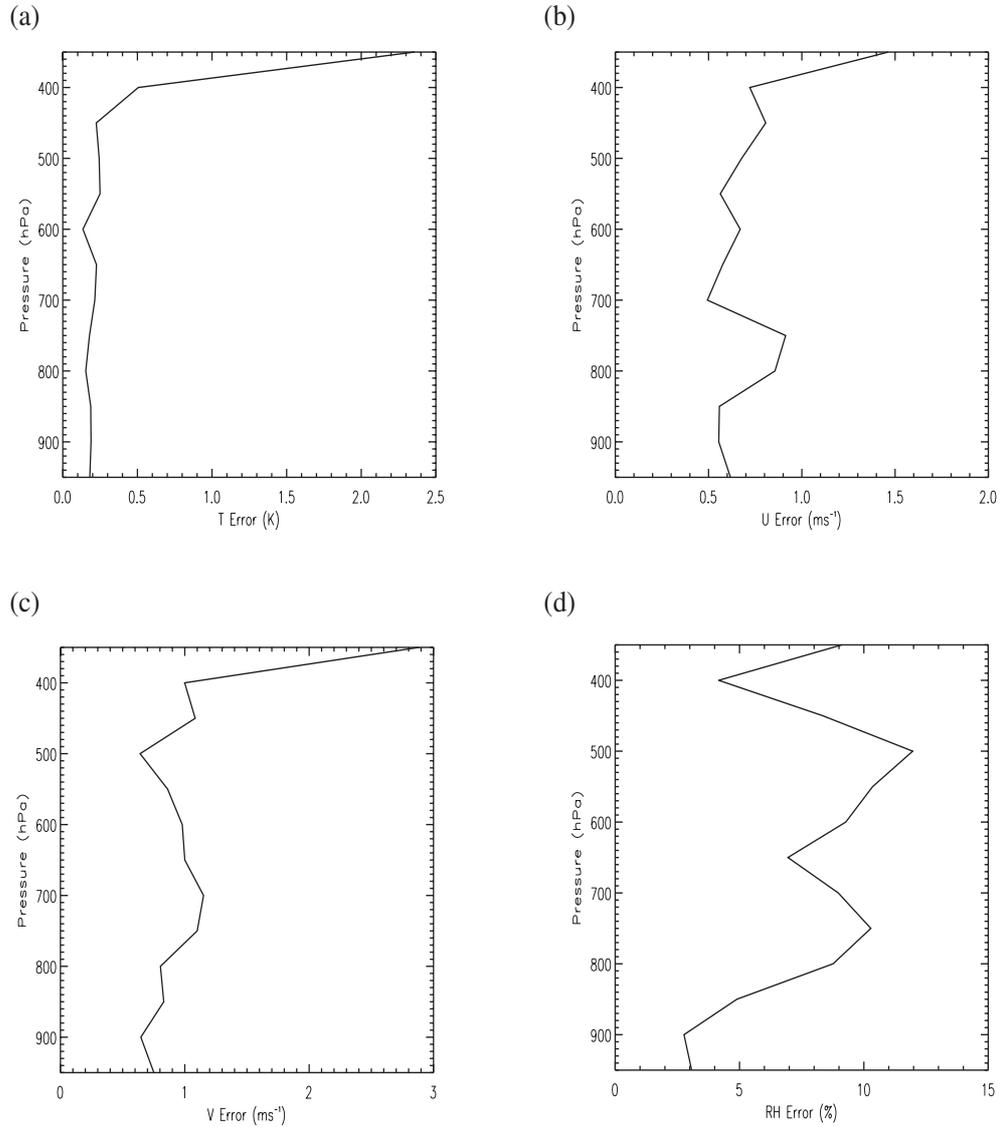
Figure 5.2: *The dropsonde observation error profiles calculated from the dropsonde data for (a) temperature, (b) zonal wind component, (c) meridional wind component and (d) relative humidity.*

data assimilation scheme. The GFDex temperature error profile is constant with height up to 400hPa with a value of 0.2°C, and then increases. Whilst the Met Office temperature error profile is also constant with height from 850hPa to 350hPa, the error is four times the calculated GFDex error at 0.8°C. The GFDex wind error profiles in Figures 5.1(b) and (c) are also smaller than the operational errors, with an average magnitude of less than half that of the operational errors. The GFDex relative humidity errors are most similar to the operational humidity errors, with a difference of only a few percent. There are several reasons why the

calculated GFDex errors are smaller than the operational Met Office errors. Firstly, as was discussed in Section 5.1, the Met Office error profiles were calculated over 10 years ago using much coarser resolution models than the 24km model used for the GFDex hindcasts (note that they were recalculated in summer 2009 and the resulting errors are slightly smaller than the operational errors shown here). The representivity error for the GFDex dropsonde observation error is therefore smaller than for the operational observation error. A second reason is that the Met Office errors were calculated using radiosonde data from several stations over a period of several months in summer and winter (to eliminate seasonal bias), which is a much larger data set than that used to calculate the GFDex errors, therefore the errors calculated here are less robust. The differences between the calculated GFDex dropsonde observation error profiles and the operational profiles suggest that a smaller error should be applied to the dropsonde data when it is assimilated, and that reducing the u, v, T and RH dropsonde observation error profiles by up to one quarter in the experiments performed in this chapter is not unreasonable.

5.3 Methodology

In order to try to increase the impact of the targeted observations, hindcasts were run using the MetUM on the NAE domain. The setup of the hindcasts is the same as is detailed in Chapter 4; the reader is referred to Chapter 4 for details. A brief overview of the hindcast experiments is given here.

The full forecast cycle was re-run for the period of the field campaign, assimilating only routine observations to give a set of forecasts which were not influenced by the targeted observations. These are referred to as the CONTROL forecasts. Hindcasts were then run for the forecast starting at 12 UTC on 01 March 2007, using the background from the CONTROL forecast and assimilating both routine and targeted observations. These are the targeted forecasts.

Table 5.1 describes the nine different hindcasts which were run. The 'TNOMEM' set of experiments assimilated eight dropsondes (set A; Figure 5.3(a)) and are the same set of dropsondes assimilated in the TNOMEM hindcast in Chapter 4. For this set of experiments three different hindcasts were run with the dropsonde observation error profiles reducing from operational (OPS) values (i.e. the TNOMEM experiment from Chapter 4), to half the operational values and lastly to one quarter of the operational values, for u, v, T and RH.

The 'ALL' set of experiments was then run where all available dropsondes were assimilated (set B; Figure 5.3(b)). This experiment increases the spatial resolution of the dropsonde data, from a minimum spacing of 220km to 85km, but does not increase the area sampled by the dropsonde data. The increase in the resolution of the observations should better capture gradients in the Denmark Strait. Again, this experiment was run three times with operational, half operational and one quarter operational dropsonde observation error profiles.

The 'NEW_TNOMEM' set of experiments assimilated the same set of sondes as the TNOMEM hindcast except that it replaced the two sondes on the Greenland coast with two sondes released further away from the coast (set C; Figure 5.3(c)). This experiment was run twice with operational and then half the operational dropsonde observation error covariances. A further experiment 'ALL_NOGL' assimilated all available dropsondes except the two on the Greenland coast (set D; Figure 5.3(d)), with operational error covariances. These two experiments test whether the two Greenland sondes were beneficial to the forecast quality, as they were released close to steep orography and so may contain features not resolved by the model, and may not be representative of the conditions around them.

Table 5.1: *Description of hindcast runs*

| Hindcast	Observations Assimilated	Dropsonde observation errors		
		OPS	$\frac{1}{2}$ OPS	$\frac{1}{4}$ OPS
TNOMEM	A	Y		
TNOMEM_HALF	A		Y	
TNOMEM_QUART	A			Y
ALL	B	Y		
ALL_HALF	B		Y	
ALL_QUART	B			Y
NEW_TNOMEM	C	Y		
NEW_TNOMEM_HALF	C		Y	
ALL_NOGL	D	Y		
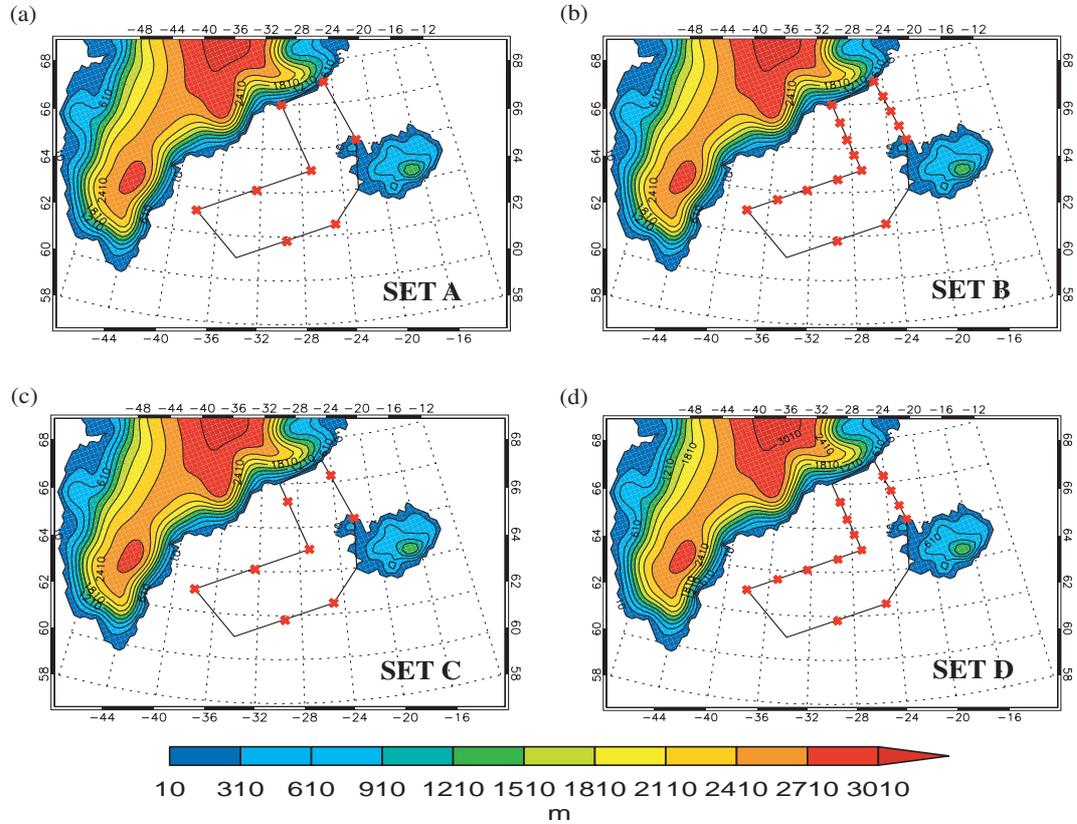
Figure 5.3: *The configurations of observations for the targeting experiment, (a) the original TNOMEM experiment (set A), (b) all sondes (set B), (c) the TNOMEM set with the 2 sondes on the Greenland coast replaced by sondes released further from the coast (set C) and (d) all sondes except the two sondes on the Greenland coast (set D). The flight track for 01 March 2007 flight (solid line), with the model orography (coloured contours). The contour interval is 300m, and the lowest contour plotted is 10m above mean sea-level.*

5.4   Results

5.4.1   Impact of Modifying the Configuration of Observations and Observation Error Specification

To assess the impact of the different dropsonde observation error covariances and set of assimilated dropsondes on the accuracy of the targeted forecast, the forecasts must be compared against a forecast containing no targeted observations (CONTROL forecast) and both forecasts verified. This study used 25km resolution ECMWF analyses as the true state, against which the relative quality of the targeted and CONTROL forecasts could be assessed. Note that this is a change from Chapter 4, where the forecast verification shown is against MetUM

analyses (verification against MetUM analyses produced qualitatively similar results, therefore there are not large differences between the two sets of analyses). As the methodology presented in Section 5.3 changes aspects of the data assimilation system, like the observation error covariances, ECMWF analyses are used for verification as they are produced by an independent data assimilation scheme; the MetUM analyses used to verify the forecasts in Chapter 4 were produced using the same data assimilation system that is modified in the experiments in this chapter.

A measure of how much the assimilation scheme has used the dropsonde data to alter the background profile is given by the *adjustment* (introduced in Chapter 4):

$$Adjustment = \mid CONTROL - SONDE \mid - \mid TNOMEM - SONDE \mid, \qquad (5.5)$$

where TNOMEM could be any targeted forecast from Table 5.1. Dropsonde and model data were interpolated onto common pressure levels, before calculating the *adjustment*. Positive values of *adjustment* indicate that the background profile has been adjusted towards the sonde profile. Figure 5.4 shows profiles of *adjustment* for the temperature field and v-wind component for experiments assimilating observation sets A-C. It can be seen in figures 5.4(a) and (b) that the *adjustment* towards the sonde data has small, mostly positive values for the TNOMEM hindcast, when only a subset of the observations are assimilated with operational error covariances. The effect of decreasing the dropsonde observation error covariances is clearly to cause the model profile to be adjusted more towards the sonde data, especially for the wind components (*u*-component similar to *v*-component; not shown). When all available dropsondes (set B) or a subset of sondes that do not contain sondes adjacent to the steep orography (set C) are assimilated, the *adjustment* is large even without reducing the error covariances from operational values. Figures 5.4(c)-(d) show that the effect of increasing the spatial resolution of the dropsondes is the same as reducing the error covariances, in terms of the resulting *adjustment*. Reducing the error covariances has little effect, which suggests that the model is unable to adjust any further towards the sonde profile - the maximum *adjustment* has been reached. This is also true for a set of sondes excluding the sondes near Greenland (Figures 5.4(e)-(f)), which show much larger *adjustment* than a set of sondes identical except for the two sondes near Greenland.
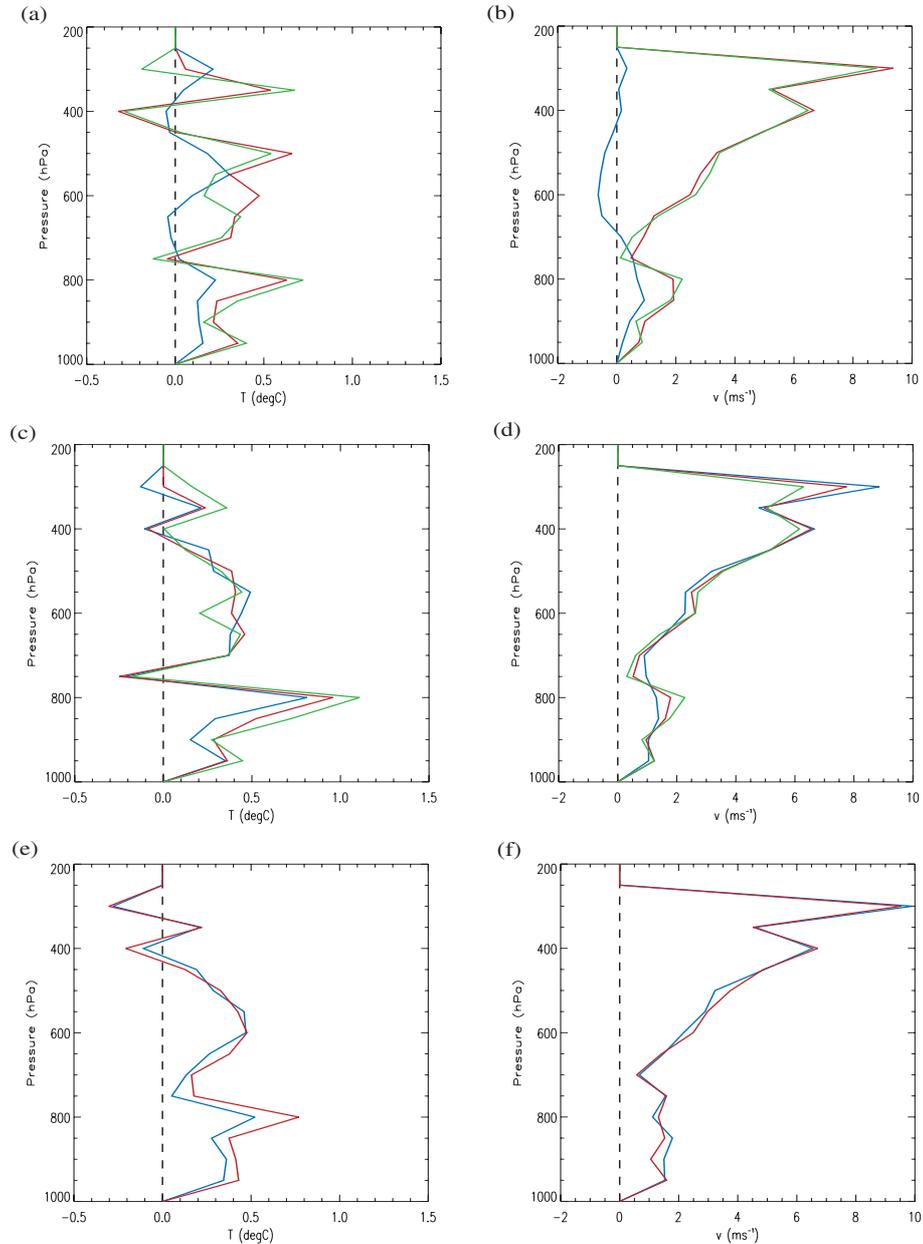
Figure 5.4: *Profiles of average adjustment for (a) and (b) the 'TNOMEM' hindcasts assimilating observation set A, (c) and (d) the 'ALL' hindcasts assimilating observation set B and (e) and (f) the 'NEW_TNOMEM' hindcasts assimilating observation set C. Panels (a), (c) and (e) show the adjustment for the temperature field, and panels (b), (d) and (f) the adjustment for the meridional component of the wind field. The adjustment has been calculated separately for each sonde and then averaged over the total number of sondes at each pressure level. The colour of the line indicates the magnitude of the dropsonde observation error covariances: operational (blue), half operational (red) and one quarter operational (green) values.*
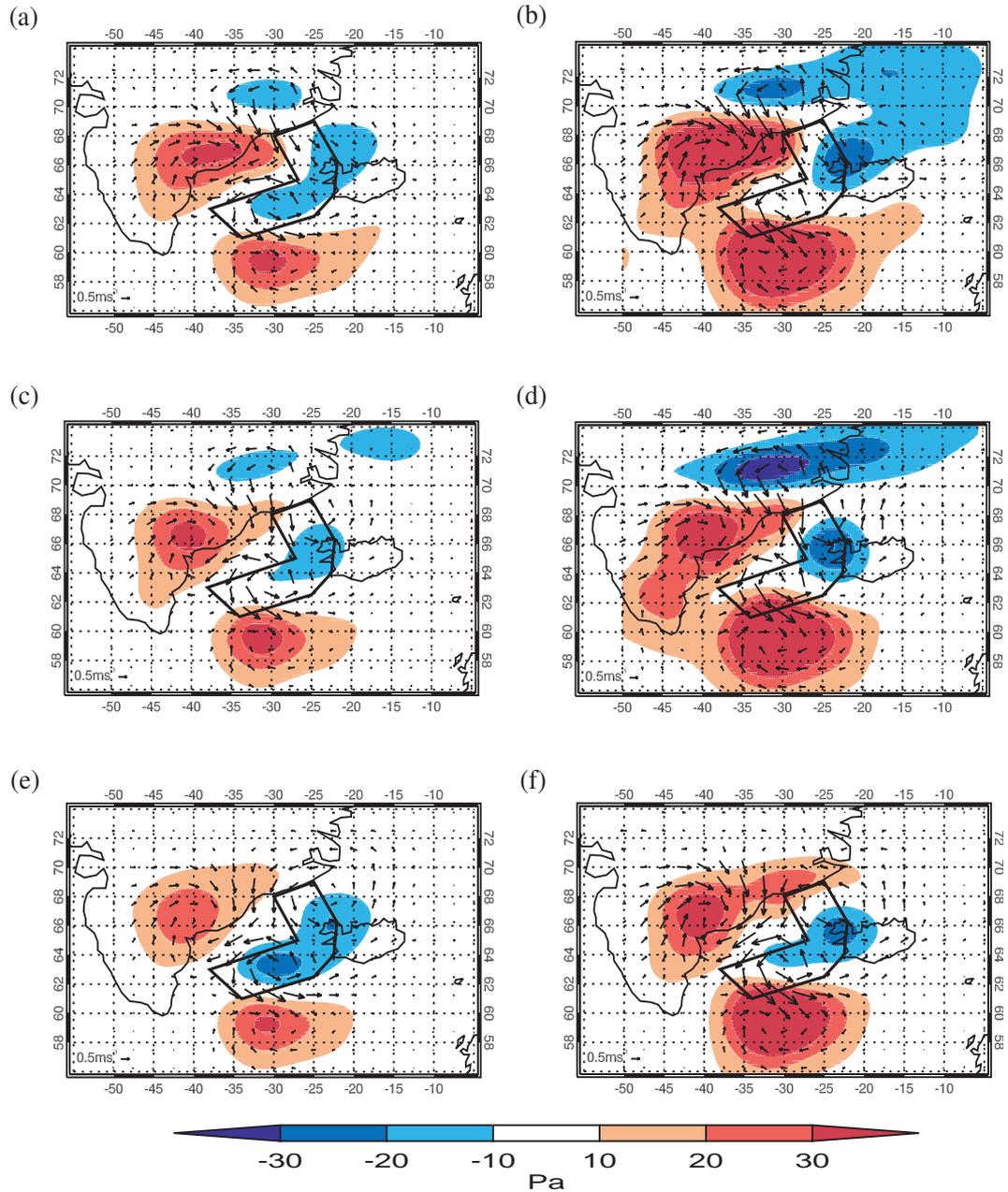
Figure 5.5: *TNOMEM minus CONTROL analysis increment for the 01 March case: pressure (shaded) and wind strength and direction (arrows) for (a) TNOMEM, (b) TNOMEM_HALF, (c) ALL, (d) ALL_HALF, (e) NEW_TNOMEM and (f) NEW_TNOMEM_HALF. The wind differences are a pressure-weighted vertical average over all 38 model levels, and the pressure increment difference is shown at model level 16 (approximately 500hPa). The flight track is overlaid.*

The effect of increasing the weighting of the sondes during the data assimilation is that the impact of the targeted sonde data on the analysis increment is increased. Recall from Fig-

ure 4.5 in chapter 4 (repeated in Figure 5.5(a)) that the targeted observations strengthened the lee cyclone through a stronger cyclonic wind increment and a more negative pressure increment in the TNOMEM analysis increment relative to the CONTROL analysis increment. Figure 5.5(a) to (d) shows the TARGETED - CONTROL analysis increment for the TNOMEM, TNOMEM_HALF, ALL and ALL_HALF hindcasts. From comparison of Figure 5.5(a) with the other figure panels it is clear that increasing the number of sondes and decreasing the dropsonde observation errors leads to a cyclonic wind increment that is stronger than the original TNOMEM experiment, although a more negative pressure increment is not visible at the 500hPa level. Thus the pattern of impact is not greatly changed, rather it is amplified with decreasing error covariances and increasing number of observations. The pattern of impact is modified for the NEW_TNOMEM and NEW_TNOMEM_HALF hindcasts. A comparison of Figures 5.5(e) and (f) with the other figure panels shows that when two sondes in the Denmark Strait are assimilated instead of the two sondes adjacent to Greenland, there is no cyclonic increment induced to the north of the positions of these sondes. The main features discussed earlier, such as the cyclonic wind increment in the Denmark Strait are still visible. This is also true for the *impact*, where forecast improvement and degradation is seen in the same locations, but with differing magnitudes.

The improvement to a forecast was calculated in terms of total energy using:

$$Relative\_Impact = \frac{I}{Error}100, \tag{5.6}$$

where *I* is the impact, calculated using (4.3), and Error is the error in the CONTROL forecast, calculated using (4.4). The multiplication factor of 100 converts the *relative_impact* to a percentage. Here, *relative_impact* is used in place of *I* that was used in Chapter 4 as we are interested in the size of the reduction in forecast error obtained using one configuration of observations and error specification relative to a different configuration and specification.

Figure 5.6 shows *relative_impact*, equivalent to the percentage reduction in forecast error (when the *impact* is positive) for the Scandinavian verification region, calculated using (5.6). A forecast is improved if *relative_impact* is positive and degraded if it is negative. The original TNOMEM experiment (dashed blue line) shows increasing *relative_impact* up to a maximum of 7% after 18 hours optimisation time, after which the *relative_impact* decreases. After
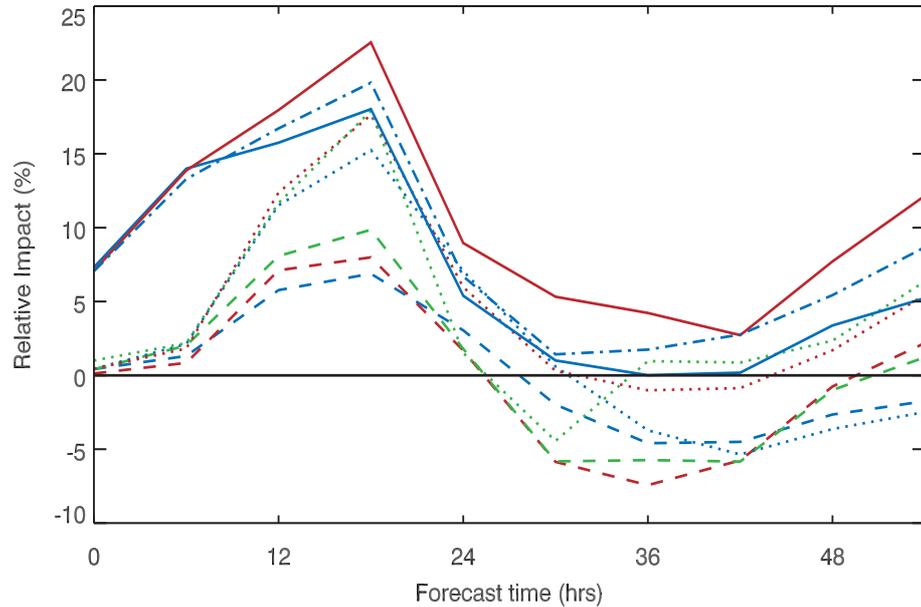
Figure 5.6: *The percentage reduction in forecast error averaged over the Scandinavian verification region, calculated using (5.6), for the 'TNOMEM' hindcasts which assimilated observation set A (dashed lines), the 'ALL' hindcasts assimilating observation set B (dotted lines), the 'NEW_TNOMEM' hindcasts assimilating observation set C (solid lines) and the 'ALL_NOGL' hindcast assimilating observation set D (dash-dot line). The colour of the line indicates the magnitude of the dropsonde observation error covariances: operational (blue), half operational (red) and one quarter operational (green) values.*

around 30 hours optimisation time the *relative_impact* becomes negative, indicating that the effect of the targeted sondes is to deteriorate the forecast in the Scandinavian verification region. Recall from Chapter 4 that the maximum reduction in forecast error when the forecast was verified against MetUM analyses was 5% after 24 hours forecast time. This forecast improvement, seen during the first 30 hours of the forecast, was shown to be a robust result in Section 4.8. Using the ECMWF analyses, forecast improvement is also seen at 24 hours forecast time, although at 3% it is slightly smaller than seen using MetUM analyses, and when calculated as a percentage (i.e. relative to the size of the error in the CONTROL forecast), the peak in impact is seen 6 hours earlier, after only 18 hours forecast time.

It is clear that when the same dropsondes are assimilated with reduced observation error covariances the magnitude of the forecast *relative_impact* is increased. There is a systematic increase in the peak in *relative_impact* as the dropsonde observation error covariances are decreased to half and then one quarter of the operational values, increasing the maximum

relative_impact from 7% to 10%. The timing of the peak improvement does not change; the impact moves with the same features described in Chapter 4, but is larger as the increased weighting given to the targeted sondes during the assimilation has increased their impact on the analysis increment (the forecast background is unchanged by these experiments).

When all available dropsonde data are assimilated, with operational observation error co-variances, the peak *relative_impact* is larger than for any of the experiments assimilating only a subset of the observations. The maximum *relative_impact* increases to 15%, double that seen in the original TNOMEM experiment. When all available dropsonde data is assimi-lated with reduced observation error covariances, the peak in *relative_impact* rises to 17%, but Figure 5.6 shows that there is little difference in the *relative_impact* on the forecast when the dropsonde observation errors are reduced from 1/2 to 1/4 of the operational values. This agrees with the initial model *adjustment*, which varied little as the error covariances were reduced.

The most interesting feature of Figure 5.6 is the *relative_impact* resulting from assimilat-ing a set of sondes that are identical to set A (Figure 5.3(a)) except for replacing two sondes adjacent to Greenland with two sondes in the Denmark Strait (set C; Figure 5.3(c)). The max-imum *relative_impact* is more than doubled from 7% to 17%, when the sondes are assimilated with operational error covariances. As the only difference between the two sets of observa-tions is the proximity of two of the observations to the steep orography of Greenland, it can be inferred that the proximity of the observations leads to less impact, implying that these two observations on the coast of Greenland have had a large negative, rather than positive, im-pact on the forecast. Assimilating set C of observations with reduced error covariances shows an increase in maximum *relative_impact* of around 5%. These results imply that the reason for the increased impact when all observations were assimilated compared to the TNOMEM subset of observations is not purely the increased density of observations, but that the obser-vations in the Denmark Strait may be counteracting the effect of the observations too close to Greenland.

5.4.2 Forecast Deterioration by Sondes Adjacent to Greenland's Orography

It was shown in the previous section that the two sondes located on the coast of Greenland, part of the TNOMEM set of observations analysed here and in Chapter 4, appear to deteriorate

the forecast. To confirm this result, an additional hindcast was run assimilating all sondes except the two sondes on the coast of Greenland (set D; Figure 5.3(d)). This gave a maximum *relative_impact* in the Scandinavian verification region after 18 hours of 20%, compared to 15% when all observations were assimilated (Figure 5.6). This confirms that the inclusion of the sondes on the coast of Greenland deteriorates the forecast for the Scandinavian verification region. This section examines why these observations degrade the forecast.

The first hypothesis that we test is that the two observations close to Greenland degrade the forecast because they contain structure that the model is not capable of resolving. The orography of Greenland has a large effect on the atmospheric circulation both locally and remotely. The effect on the local airflow is to block and distort the flow, creating barrier winds which flow southwards along the coasts, and tip jets in the lee of Greenland. Sondes placed close to steep orography may measure such local effects, which then makes the sonde data unrepresentative of the larger area around it. This data can be easily dealt with during the observation processing, by checking it against the background field, and rejecting the data if it is significantly different from the background field (assumed to be due to measuring local effects). In order to ensure such data is rejected during the observation processing, the Met Office set the initial probability of gross error for the coastal Greenland radiosonde stations such as Tasiilaq (Figure 1.1) to 0.2-0.3 instead of 0.01 so that the data is more easily rejected.

Figure 5.7 shows model profiles at 3 locations: (1) on the Greenland plateau, (2) at the location of one of the coastal dropsondes, and (3) in the Denmark Strait. Each point is separated from the next by approximately 200km, this is the separation distance used in the GFDex targeting experiments, and equal to the horizontal correlation length-scale for temperature assumed by the data assimilation scheme. The first thing to note is that the profiles, particularly at upper-levels, do not look significantly different. The temperature in the lowest part of the atmosphere at point (2) (Figure 5.7(b)) is around 10°C colder than at point (3) (Figure 5.7(c)) due to the fact that point (2) is over ice and point (3) is over open ocean, but the wind profiles are similar (Figure 5.7(e) and (f)). The largest difference in the profiles at the three locations is that at point (1) the profile starts at 750hPa (Figure 5.7(a) and (d)), as it is located on the Greenland plateau which at that location is 2700m above sea level. The dropsonde observation at point (2) has been overlaid on Figures 5.7(b) and (e); the dropsonde and model profiles have a similar structure although the modelled temperature is up to 5°C too warm and

the modelled windspeed up to 5ms^{-1} too slow throughout the depth of the troposphere. This shows that the dropsonde observation does not contain structure that cannot be resolved by the model (remember we are comparing the dropsonde observation to a model profile that does not contain the observation). Therefore the forecast degradation has not occurred because the model background has been adjusted towards observational data that contains structure that the model is not capable of resolving.
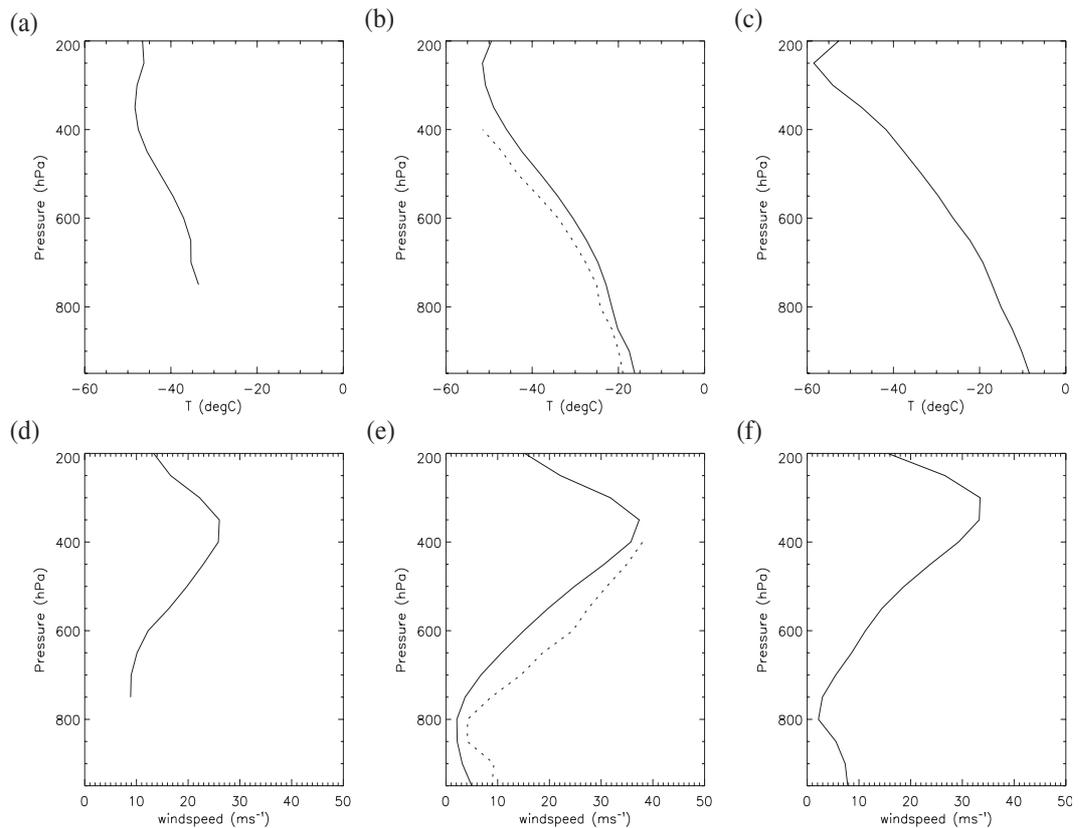

Figure 5.7: *Model profiles from the CONTROL forecast for (a) and (d) over the Greenland plateau (70°N 32°W), (b) and (e) adjacent to the Greenland coast (68°N 30°W) and (c) and (f) in the Denmark Strait (66°N 28°W), for temperature (top row) and wind speed (bottom row). The dropsonde observations, interpolated to the same pressure levels as the model data, are overlaid on (b) and (e).*

Having disproved our first hypothesis, we now test the hypothesis that it is the spreading of the observational data from the sondes on the Greenland coast that causes the forecast degradation. This is motivated by the fact that the conditions at low levels (here meaning near the surface) differ significantly between locations (1) (Figure 5.7(a) and (d)) and (2) (Figure 5.7(b) and (d)). Figure 5.3(a) showed the positions of the sondes released during the

flight, relative to the orography of Greenland. This shows two targeted sondes (marked by red crosses) in the northern part of the flight track positioned on the south-eastern coast of Greenland. From the figure it can be seen that within 100km of these sondes the height of the orography rises from sea-level to over 1000m. As was discussed in Chapter 2, during the data assimilation process **B** acts to spread out observational information with a maximum radius equal to a correlation length-scale defined by **B**. For temperature this is approximately 200km, or the distance between the profiles in Figure 5.7. It is clear from Figure 5.7 that the dropsonde profile at (2), in particular the lowest part of the atmosphere, is representative of conditions in the Denmark Strait, not of conditions over the Greenland plateau.

To test this hypothesis, it is necessary to calculate the analysis increments resulting from the assimilation of the two sondes on the coast of Greenland, recalling from Chapter 2 that the size of the analysis increment is proportional to the spreading of the observational data by the background errors. Figure 5.8 shows the ALL analysis increment minus the ALL_NOGL analysis increment generated by assimilating all sonde data except the two Greenland sondes (observation set D). The difference therefore shows the analysis increment due to these two sondes. The cross-section cuts through the position of one of these sondes, at 30°W. Neglecting the influence of the other coastal sonde, which is several hundred kilometres to the north-east of this sonde (which was released at 68°N, 30°W), we can consider the difference in the analysis increment seen here to be due to the assimilation of this sonde. It is clear from the cross-section that **B** has acted to spread out the information contained in this observation both to the east and to the west, over a distance of several hundred kilometres (the cross-section is at 68°N where 10° longitude is approximately 400km). The MetUM and 4D-Var systems use a vertical co-ordinate based on height rather than pressure, and this is terrain following near the ground. This means that when **B** has spread the information from this sonde observation to locations around it, it has spread some of the information up along the steeply sloping orography. This is evident by the tilt seen in Figure 5.8, starting from the location of the observation at 30°W and following the slope of Greenland to the west, which also matches the slope of the model levels. This suggests that it is the spreading of observational data along terrain-following model levels, which in reality is up a steep slope, that has caused a degradation of the initial state and therefore the forecast.

It is important to find a way to utilise observations near steeply sloping orography in a
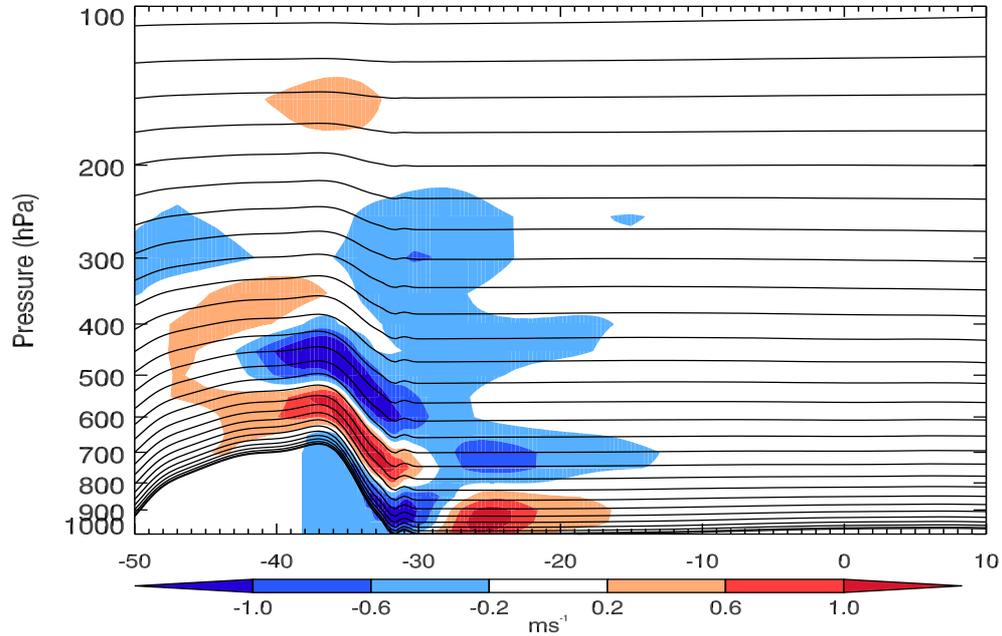
Figure 5.8: *Cross-section through the Denmark Strait at 68°N from 50°W to 10°E showing the ALL minus ALL_NOGL analysis increment (coloured contours) for the v-wind component and pressure on model-levels (black contours). The pressure levels are terrain-following near the surface, therefore the lowest pressure level indicates the orography; the cross-section cuts through Greenland on the left. Note the log scale on the y-axis.*

manner that does not result in the degradation of the forecast downstream. Potential solutions could be not to site observations next to steep slopes, or to automatically reject data below the orography height. An approximate solution following these lines has been attempted here, by re-running the TNOMEM hindcast, but excluding data below 850hPa from the two son-des on the Greenland coast. The choice of the cut-off of 850hPa was a compromise based on the fact that the model orography reaches 750hPa only 200km from the observation, but that excluding too much of the profile could itself be detrimental to the forecast, as data that is spread away from the slope is likely to be beneficial at all forecast levels. Figure 5.9 shows the *relative_impact*, calculated using (5.6) for the TNOMEM forecast and the new TNOMEM forecast that does not contain Greenland sonde data below 850hPa. The differ-ence in *relative_impact* is striking; the maximum forecast improvement is doubled using this method. It gives weight to the conclusion that the spreading of observational data from the

lower model-levels upslope is the problem here, as removing this part of the profile increases the impact this set of sondes has on the forecast. A better solution to this problem would be to improve the representation of the background error covariances to stop the spreading of observational data upslope; investigating this solution was beyond the scope of this thesis.
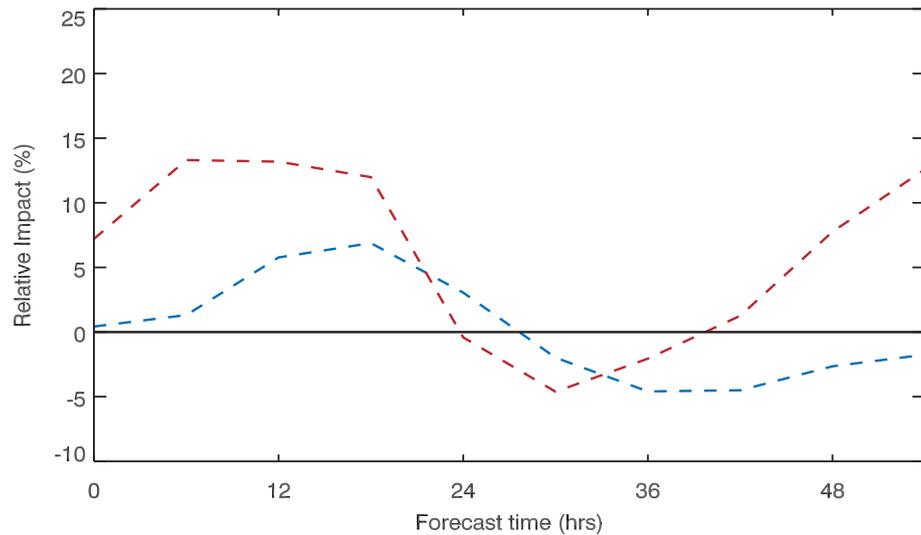

Figure 5.9: *The percentage reduction in forecast error averaged over the Scandinavian verification region, calculated using (5.6), for the TNOMEM hindcasts which assimilated observation set A (blue line) and observation set A with no data assimilated below 850hPa for the two sondes adjacent to Greenland (red line). Both experiments used the operational dropsonde observation error covariances.*

5.5 Conclusions

This chapter focussed on the 01 March 2007 targeting case. The original hindcasts, the results of which were presented in Chapter 4, showed relatively small impact. Here the observation error covariances and set of sondes assimilated is varied to examine the potential to increase the impact of the targeted observations.

The GFDex dropsonde data and model data containing all available observations were used to calculate the dropsonde observation errors. The calculated values were smaller than the operational values used in the assimilation process for temperature and wind, but of similar magnitude for relative humidity. It should be noted that the calculated errors from the GFDex data are likely to be lower than would be found in general due to the small sample size and period used to determine the covariances. The difference in magnitude between the operational and calculated dropsonde errors motivated the reduction in the dropsonde observation

errors in the hindcast experiments.

When the original TNOMEM set of dropsonde data was assimilated, reducing the error covariances caused the model to adjust further towards the sonde data, as it increased the weighting given to the dropsonde data as it was assimilated. Increasing the spatial resolution of the dropsonde data also increased the adjustment of the model profile towards the sonde data, so that further reducing the dropsonde observation error covariances then had little impact. It would seem that reducing the error covariances and increasing the spatial resolution of the dropsonde data have the same effect in terms of the model adjustment towards the sonde data. Reducing the dropsonde observation error covariances and increasing the number of observations increases the impact of the observations on the analysis increment, and on the forecast, without changing the pattern of the impact.

When the TNOMEM set of sondes was modified so that no sondes were close to the steep orography, the adjustment was large when the data were assimilated with operational error covariances, and reducing the covariances had little effect. The main effect of the targeted sondes was to deepen the area of low pressure in the Denmark strait, but a noticeable difference to the hindcasts which assimilated the two sondes on the coast of Greenland was that the cyclonic increment over Greenland was not present. The maximum improvement was also increased to 17% (relative to the other hindcasts), rising to 22.5% when the same set of sondes were assimilated with half the operational error covariances.

It was shown that two dropsondes located on the coast of Greenland caused a degradation of the forecast. The degradation was not caused by the sondes containing structure that the model was not able to represent, rather by the observational information being spread up steep slopes, to locations where the observational data was not representative of the atmospheric conditions. Removing the lowest part of the dropsonde profile (below 850hPa) increased the impact of the set of targeted sondes, however this solution is not ideal because the data below 850hPa is representative of conditions in the Denmark Strait and should improve the initial conditions in that region. A better solution would be to not assimilate data below the orography height or to re-formulate the background error so that the observational data is not spread up steep slopes. This result has implications for routine radiosondes released from stations close to steep orography, such as other stations on Greenland, or the Rocky mountains in the USA for example.

<center>**CHAPTER 6**</center>

<center>**Singular Vector Experiments**</center>

This chapter examines the ability of the data assimilation scheme to extract important dynamical structures from the targeted observations and use it to correct initial condition error. We test the hypothesis that targeted observations show little impact on the forecast because the 4D-Var scheme used to assimilate the targeted observations cannot make full use of the baroclinic information which may be contained within the targeted observations. To test this hypothesis, an experiment is performed that assimilates sets of synthetic targeted observations which capture an initial singular vector structure. The structure is not present in the background field or the routine observations, thus assuming that the initial condition error has a tilted, growing structure.

6.1 Background

Singular vectors were introduced in Chapter 3 as perturbations which are optimal for baroclinic growth. They are created by finding the eigenvectors of a forecast model which is tangent linear to the full model. The growth of singular vectors over an optimisation time period is measured using a norm; total energy has been shown to be an appropriate norm for targeting purposes (Palmer et al., 1998) and total-energy singular vectors were therefore used to construct sensitive area predictions for the GFDex targeting experiments.

To illustrate the structure of a typical singular vector perturbation, Figure 6.1 shows the structure of the leading total-energy singular vector that was used to construct the TESV SAP for the 01 March 2007 targeting case. The optimisation time for the perturbation was 24 hours. The initial wind, temperature and PV fields show a horizontally confined perturbation tilted against the vertical wind shear (the background u-wind component along this cross-section is shown later in Figure 6.12(c)). This is consistent with Hartmann et al. (1995) who showed that at initial time singular vectors are confined to wave packets tilted against the vertical wind shear, with maximum amplitude in the mid-troposphere at short wavelengths. The normalised

total energy of the singular vector perturbation shows a peak at 600hPa (Figure 6.2). After 24 hours evolution, the perturbations have untilted and have grown in amplitude. The maximum energy is now around the tropopause level, which may be due to the short optimisation time; Hartmann et al. (1995) showed that the final perturbation has maximum amplitude at both the tropopause and surface only if the optimisation time was long enough. Figure 6.2 shows a peak in normalised total energy at 300hPa, with a much smaller secondary peak at 1000hPa. The wavelength of the perturbations has increased, and they are located downstream of the initial perturbation.

The growth mechanism of singular vectors from initial to final time was described by Hoskins et al. (2000) in terms of 'PV unshielding'. The background zonal wind is assumed to have a positive shear with height, so that the wind speed has a minimum at the surface and maximum near the jet level. The structure of the initial perturbation in PV (Figure 6.1(e)) shows a positive PV anomaly that is tilted against the direction of the wind shear, with negative PV anomalies above and below it. These negative PV anomalies are acting to confine the singular vector perturbation in the vertical. The action of the sheared zonal flow is to untilt these perturbations, thus 'unshielding' the positive PV anomaly. The perturbations propagate upwards as a Rossby wave, leading to the maximum in the final energy of the perturbation being located near the tropopause. If the initial perturbations are located lower in the troposphere, coupling of Rossby waves at the lower and upper boundaries leads to sustained growth and final maxima at both the surface and tropopause (Badger and Hoskins, 2001). Badger and Hoskins (2001) related the growth rate of singular vectors to the growth rate of the fastest growing normal mode, superimposing singular vector perturbations onto an Eady basic state. Buizza and Palmer (1995) showed that singular vectors of an atmospheric state can amplify by a factor of ten in a 36-hour period.

Buizza and Palmer (1995) calculated the singular vectors of the global atmospheric circulation. The locations of the singular vectors they identified were consistent with the main regions of baroclinic instability in the northern hemisphere, such as the regions of cyclogenesis for the Pacific and Atlantic storm tracks. The structure of these singular vectors was comparable to those identified by Hartmann et al. (1995), with the dominant singular vectors exhibiting an against-shear tilt with maximum energy in the lower troposphere at initial time and near the tropopause at final time. They also found an upscale transfer of energy, from
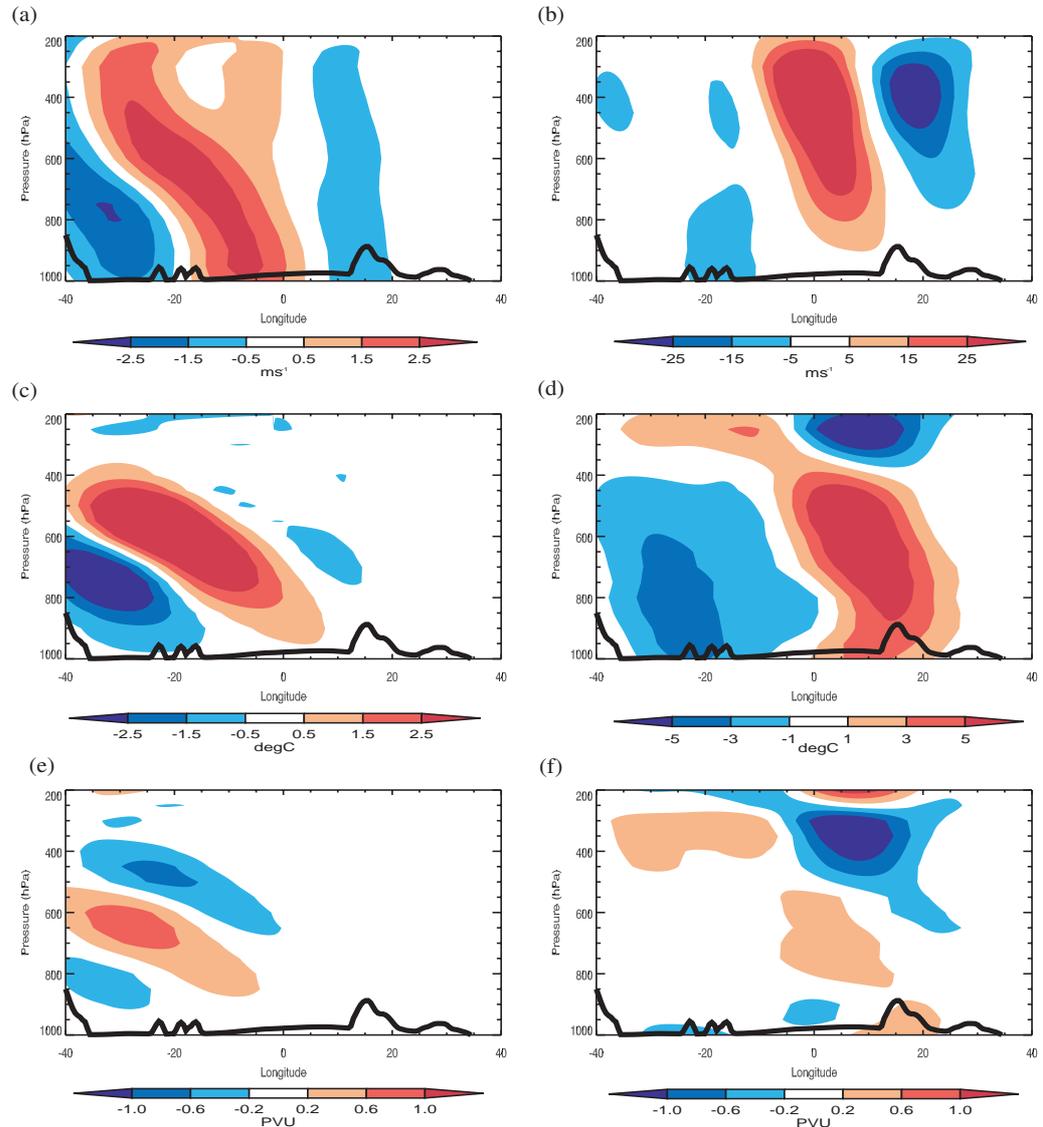
Figure 6.1: *Cross-sections at 66°N from 40°W to 40°E through the initial (left column) and evolved (right column) singular vector for (a) and (b) meridional wind, (c) and (d) temperature and (e) and (f) PV. The singular vector is the leading singular vector from the TESV SAP for 01 March 2007, which has an optimisation time of 24 hours. The perturbations have been multiplied by a factor of fifty before plotting.*

subsynoptic scales at initial time to synoptic scales at final time.

It was shown in Chapter 4 that the impact of the GFDex targeted observations on the forecast quality in the verification regions was small. For two cases, 24 and 26 February, it was shown that the lack of impact occurred because the targeted observations gave little new information to the assimilation. For the remaining two cases, 01 and 10 March, the targeted
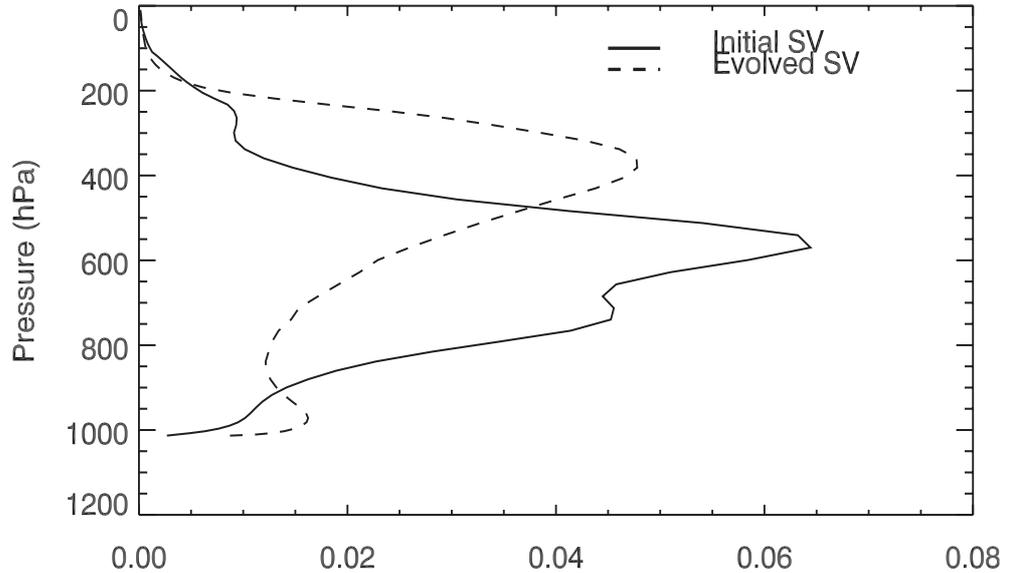
Figure 6.2: *Variation with height of the total energy of the initial (solid line) and evolved (dashed line) singular vector, each normalised by the total energy of the singular vector at that time ($2\times10^6 m^2 s^{-2}$ at T+0 and $4\times10^7 m^2 s^{-2}$ at T+24). The optimisation time is 24 hours.*

observations did provide additional information to the assimilation (Figure 4.2(c) and (e)), but the data assimilation scheme did not appear to be able to make full use of this additional data to improve the analysis (Figure 4.2(h) and (j)). Three of these four targeting cases (24 and 26 February, 01 March) released targeted observations along flight-tracks into a TESV-defined sensitive area. Figure 6.3 shows that these sensitive areas (the flight tracks for these cases indicate where the sensitive region is) coincided with regions of baroclinic instability. Here the Eady index (3.7) is used as a measure of the baroclinicity of an atmospheric state, where the vertical wind shear is calculated over the 1000 to 300 hPa layer. Therefore it is plausible that the targeted observations from these cases contained baroclinic information that was not utilised by the assimilation scheme. Here we test the hypothesis that the lack of impact relates to the inability of the data assimilation scheme to extract information from targeted observations and use it to correct initial condition error of a baroclinic structure. Section 6.2 describes the methodology used to test this, Section 6.3 shows the results and the conclusions are detailed in Section 6.4.
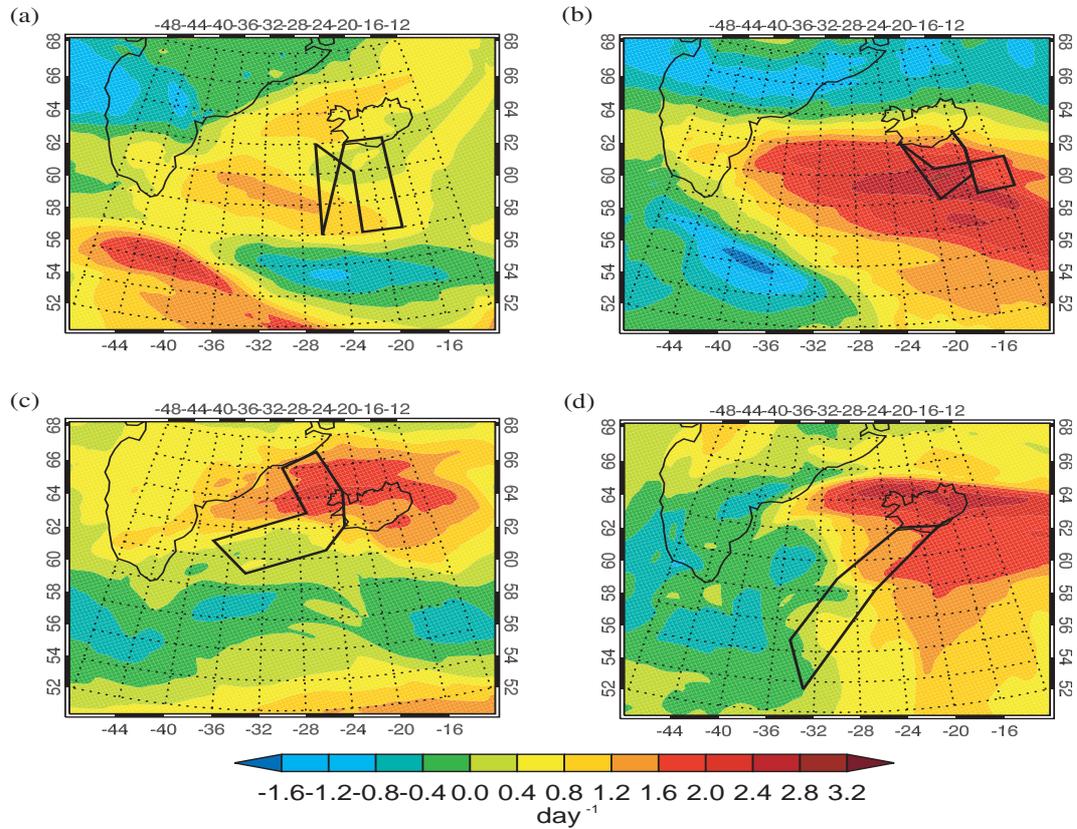
Figure 6.3: *Eady Index (3.7) for the 1000-300hPa layer for the GFDex targeting cases of (a) 1200 UTC 24 February, (b) 1200 UTC 26 February, (c) 1200 UTC 01 March and (d) 1200 UTC 10 March. The flight track corresponding to each case is shown by a thick black line.*

6.2   Methodology

6.2.1   Synthetic Targeted Observation Profiles

Sets of synthetic targeted observations containing a singular vector structure 'SSV', were created by taking profiles through a perturbed analysis. The perturbed analysis was constructed by adding wind and temperature perturbations from the leading ECMWF TESV from the 01 March 2007 targeting case on to the CONTROL 1200 UTC analysis for the same day. A control set of observations 'NOSV' was also created whereby profiles of u, v and T were taken from the unperturbed analysis (1200 UTC CONTROL analysis). As the background field used in the Met Office 4D-Var system is valid at 0900 UTC and the singular vector perturbations are valid at 1200 UTC an analysis was used in place of the background field. This required the assimilation of profiles from the unperturbed analysis as well as from the per-

turbed analysis, as the unperturbed analysis is not the same as the background field that the observations are compared against in the assimilation. The SSV and NOSV synthetic observations consisted of profiles of *u*, *v* and *T*, from 1000hPa to 200hPa with data every 50hPa. Some profiles intersected Greenland, and here the nearest pressure level above the surface was used as the lowest data point, e.g. 650hPa over the Greenland plateau.

The lead and optimisation times for the singular vector used in this experiment were both 24 hours. This singular vector was chosen because at targeting time at 1200 UTC on 01 March, the initial SV was located in the region of the flight track (Figure 6.4(a)), and after 24 hours the evolved SV was located within the verification region (Figure 6.4(b)). The leading SV amplifies by a factor of ten during the 24 hour optimisation time. The TESVs are described in detail in Chapter 3, it is necessary here to recall that these are dry TESVs, meaning that moisture is not taken into account in their calculation. Therefore no relative humidity observations were included in the synthetic observations that were assimilated, to ensure that the relative humidity was not perturbed by the synthetic observations.
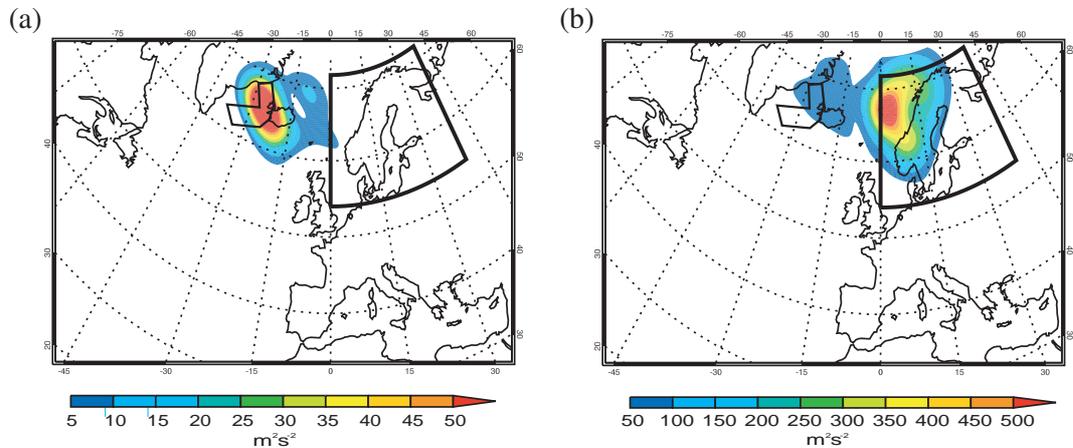

Figure 6.4: *Vertically integrated total energy of fifty times the SV perturbations at (a) initial time, and (b) after 24 hours evolution at final time. Note the difference of a factor of ten between the scales for (a) and (b). The total energy is summed over 850, 500 and 250 hPa levels. The flight track from the 01 March 2007 GFDex targeting case, and Scandinavian verification regions are marked by thick black lines.*

The structure of the SV perturbations used in this experiment was shown in Figure 6.1. Note that Figure 6.1 showed the SV perturbations multiplied by a factor of fifty. The size of the unmultiplied SV perturbations is very small, therefore if these values were to be added

to the analysis then the observation minus background difference would be extremely small which would result in almost no adjustment of the background towards the observations. Therefore the TESV perturbations were multiplied by a factor of 50 before being added to the analysis. The size of the resulting perturbations is shown in Figure 6.5. The average perturbation in v is 1.5ms^{-1} with a minimum of 0 and a maximum of 3.5ms^{-1}. The average observation minus background error in v for the GFDex experiments is between 2 and 5ms^{-1} (Figure 4.2) which is slightly larger than the average SV perturbations used in this experiment. The observation minus background error in T (not shown) is slightly smaller than the average SV perturbations shown in Figure 6.5(b). The size of the perturbations resulted in up to 10 temperature observations being rejected during the observation processing stage, this is only 1% of the total temperature observations processed. No wind observations were rejected.
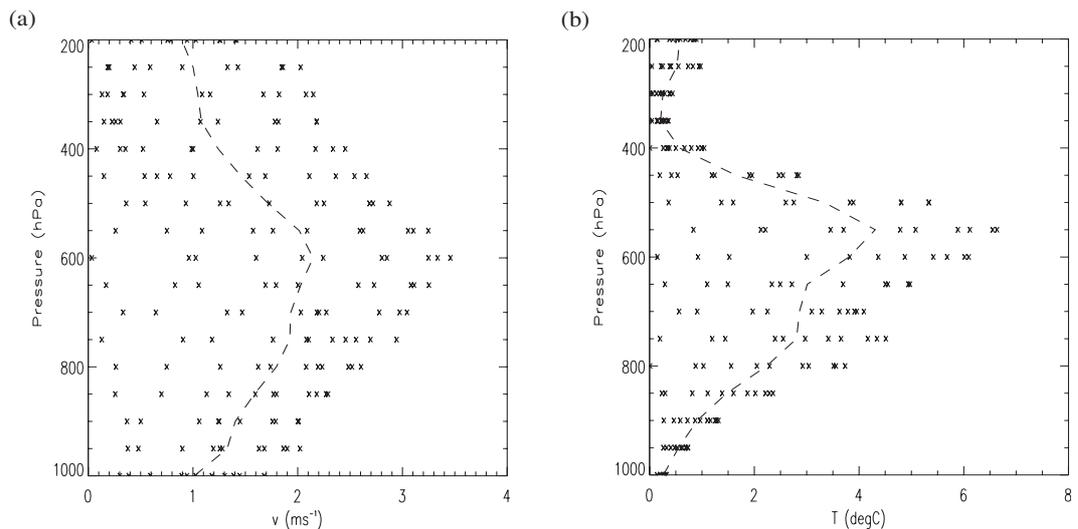

Figure 6.5: *The magnitude of the SSV perturbations used in the N55_grid experiment, for the synthetic observation profiles at 66°N. The size of each individual perturbation (crosses) and the layer-average perturbation (dashed line) for (a) the v-component of the wind and (b) temperature.*

6.2.2 Experimental Setup

The synthetic observations were arranged in a grid in order that the singular vector structure could be fully captured by the synthetic observations. Five pairs of hindcasts, equivalent to the TNOMEM runs described in Chapter 4, were run for 1200 UTC 01 March 2007. Each pair assimilated the SSV and NOSV observation profiles. The number of synthetic observation profiles was varied from a regular grid of 55 'N55_grid' (Figure 6.6(a)) to 30 'N30_grid'

(Figure 6.6(b)) to 8 'N08_grid' (Figure 6.6(c)). A further experiment was run with 8 synthetic observation profiles located along the 01 March flight track, at the same points as the dropsondes used in the TNOMEM experiment 'N08_nogrid' (Figure 6.6(d)) . An additional experiment assimilated the 55 synthetic observation profiles on a grid, but reduced the observation error covariances for the synthetic observations only to half the operational values, 'N55_halfcov'. The resulting observation error profiles for *u*, *v* and *T* are shown in Figure 5.1. The synthetic observations were treated as radiosonde observations in the observation processing and assimilation. Radiosonde soundings from Keflavik, Tasiilaq and Ittoqqortoormiit were denied from the assimilation, as they were located within 150km of a synthetic observation. This avoided having two different observations in approximately the same location.

The maximum area covered by synthetic targeted observations is approximately $1 \times 10^6 \text{km}^2$, i.e. ten times larger than the area covered by dropsonde observations during GFDex. The N30_grid and N55_grid experiments should therefore include enough observations over a large-enough area that an impact should be seen in the forecast (Leutbecher et al., 2002). As the number of synthetic observations decreases in the experiments, the longitudinal spacing is increased and so the resolution of the SV structure is decreased. Therefore in theory the N55_grid experiment should best capture the SV as this experiment has the highest resolution of data available to the data assimilation scheme.

The SSV and NOSV profiles are compared in Figure 6.7. The addition of the 50xSV perturbation results in only small changes in the profiles assimilated. There are minor changes to the strength of the low-level jet seen in some of the *v*-wind profiles (Figure 6.7(e)-(h)) with the addition of the perturbations generally resulting in a strengthening of the low-level northerly winds, for example Figure 6.7(f) shows a strengthening of 2ms^{-1} at 800hPa. The perturbation to the temperature field is largest at around 700hPa and results in a shallower lapse rate, with a weak inversion added in the more easterly profiles. A cross-section through the SSV and NOSV initial potential temperature field (Figure 6.8) confirms that the addition of the 50xSV perturbations results in a field that looks similar to the analysed field.
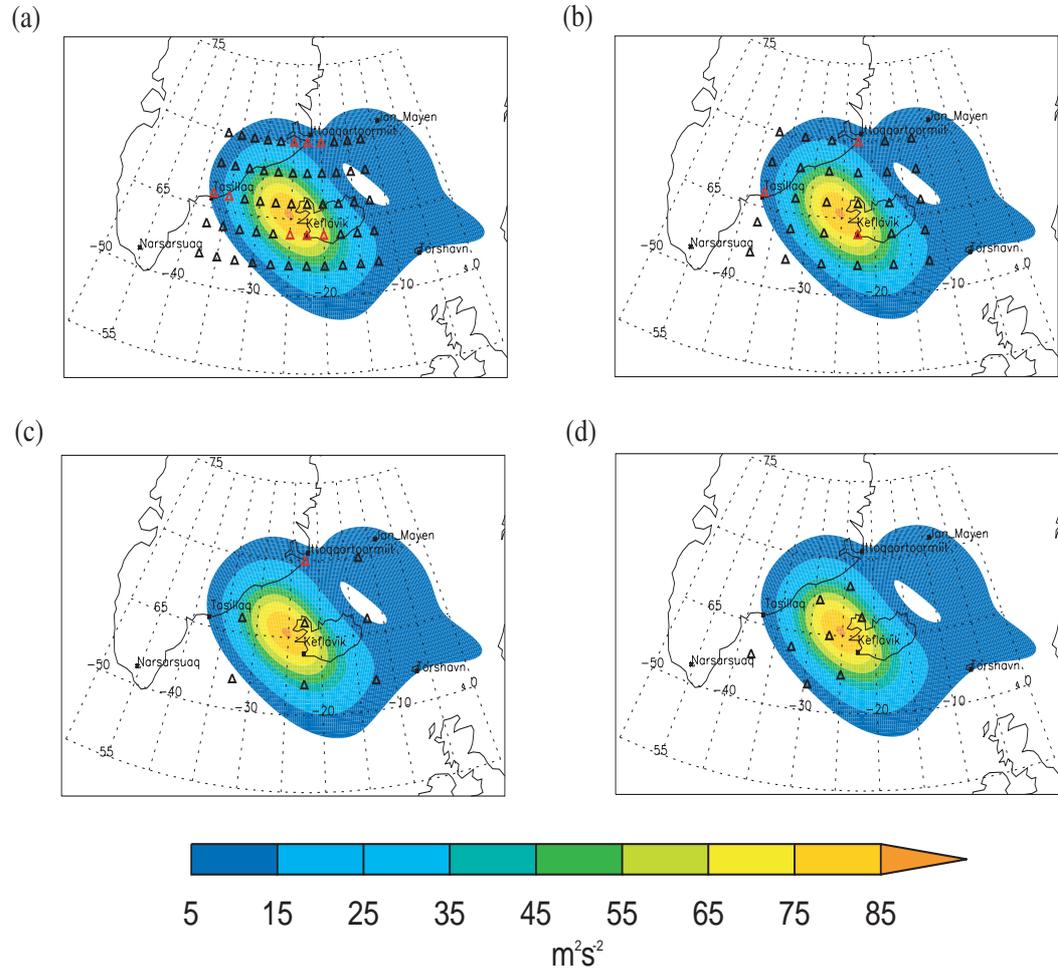
Figure 6.6: *The locations of the synthetic targeted observations (triangles) for the experiments: (a) a grid of 55 sondes with a spacing of 2 degrees latitude and 2.5 degrees longitude, (b) a grid of 30 sondes with a spacing of 2 degrees latitude and 5 degrees longitude, (c) a grid of 8 sondes with a spacing of 4 degrees latitude and 10 degrees longitude and (d) 8 sondes along the flight track from the 01 March targeting case. The locations of nearby radiosonde stations are marked with asterisks, and synthetic observations lying within 150km of a radiosonde observation are marked with red triangles; these radiosondes were denied from the assimilation. The total energy of 50 times the initial singular vector perturbations (coloured contours) marks the location of the SV.*

6.3   Results

6.3.1   Structure of Analysis Increments

To assess the ability of the data assimilation scheme to produce a tilted increment that captures the singular vector structure, the vertical structure of the SSV-NOSV analysis incre-
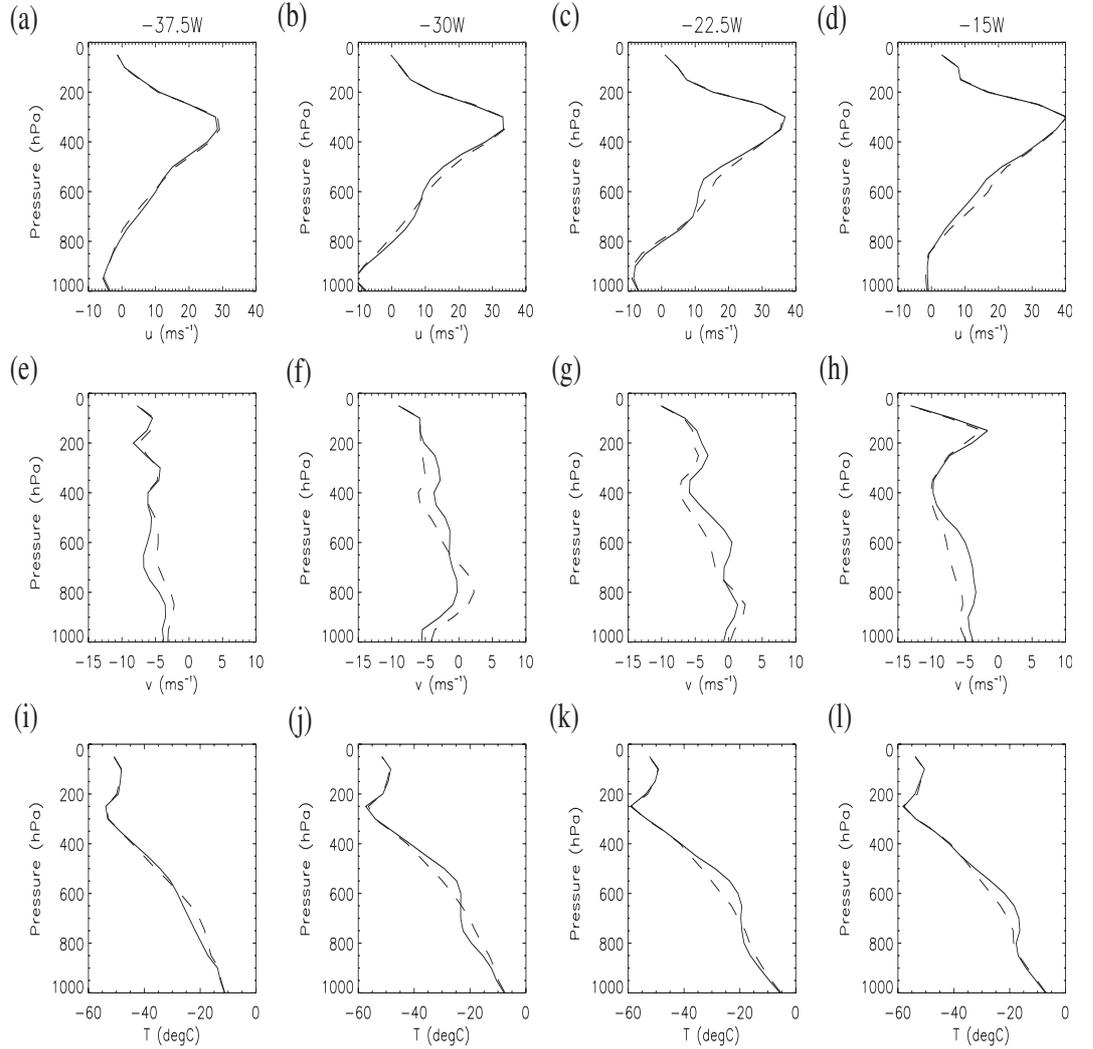
Figure 6.7: *Synthetic observation profiles from the N55_grid experiment, at 66°N for the SSV (solid lines) and NOSV (dashed lines) experiments for (a) - (d) u-component of the wind, (e) - (h) the v-component of the wind and (i) - (l) temperature. The longitude of each profile is labelled at the top of the top row of plots.*

ments was compared to that of the 50xSV perturbations. As the analysis increments do not contain PV as a variable, the PV of the analysis increments was calculated from the u, v and θ fields using the formula for Ertel's PV, q:

$$q = \frac{1}{\rho} \underline{\zeta} \cdot \underline{\nabla}\theta, \tag{6.1}$$

(a)   (b)
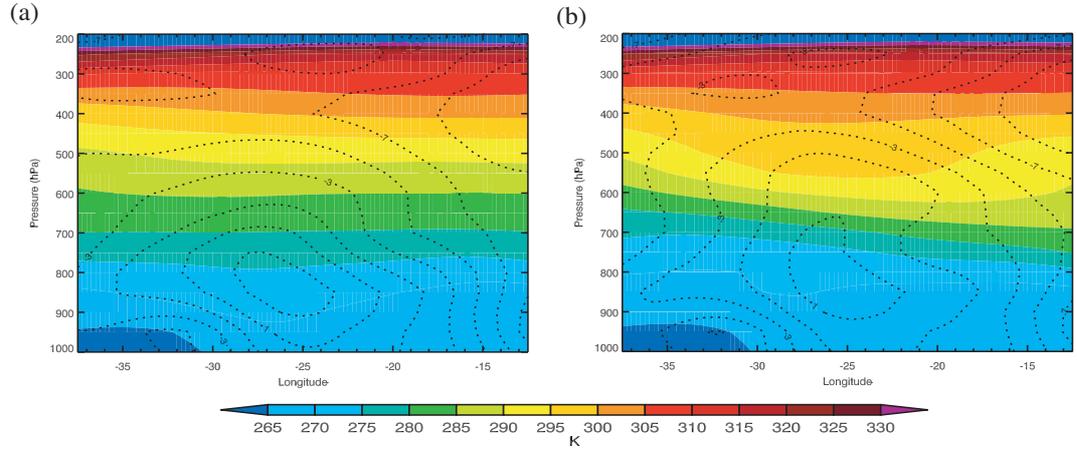
**Figure 6.8:** *Cross-sections at 66°N from -37.5°W to -12.5°W of potential temperature (coloured contours) and v-component of the wind (black contours) for (a) the 1200 UTC CONTROL analysis used to make the NOSV observation profiles and (b) 50×SV plus the analysis, used to make the SSV profiles.*

where $\underline{\zeta}$ is the absolute vorticity, and $\underline{\nabla}$ is defined in terms of u, v and p. In isobaric coordinates, assuming hydrostatic balance, this transforms to

$$q = -g\left(-\frac{\partial v}{\partial p}\frac{\partial \theta}{\partial x} + \frac{\partial u}{\partial p}\frac{\partial \theta}{\partial y} + \zeta_z \frac{\partial \theta}{\partial p}\right), \tag{6.2}$$

neglecting horizontal gradients in the vertical wind component. ζ_z is the vertical component of the absolute vorticity, calculated as

$$\zeta_z = (\underline{k} \cdot \underline{\nabla} \times \underline{u}) + f = \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + f. \tag{6.3}$$

Figure 6.9 shows the SSV-NOSV analysis increments for the N55_grid experiment compared to the 50xSV perturbations. Note that the right column of Figure 6.9 is equivalent to the left column of Figure 6.1, plotted on a different scale. The analysis increment clearly captures the singular vector structure, showing a perturbation that is tilted against the shear, with a maximum in the mid-troposphere that is in phase with the SV perturbation. The amplitude of the SSV-NOSV perturbation in *v* is larger than for the 50xSV perturbation and comparable for the *T* and *PV* fields. Two noticable differences are that the perturbation is more confined than than SV perturbation and the tilt of the SSV-NOSV analysis increments is reduced compared to the 50xSV perturbation. That the *v* and *T* fields exhibit a similar tilt is unsurprising as geostrophic and hydrostatic balance constraints are imposed during the assimilation process,

therefore the resulting analysis increments are in thermal wind balance. Following Holton (2004) thermal wind balance is satisfied if

$$\frac{\partial v_g}{\partial lnp} = -\frac{R}{f}\left(\frac{\partial T}{\partial x}\right)_p,$$
(6.4)

where v_g is the geostrophic meridional wind component (here we consider only one component of thermal wind balance), p is pressure, and the gradient of T with x is computed on an isobaric surface. From this it can be seen that if the horizontal gradient in temperature is small, then the geostrophic wind will vary little with height. Therefore a shallow gradient in the SSV-NOSV temperature perturbation is associated with a flatter wind perturbation, through the thermal wind constraint. The reduced tilt is likely to result from the barotropic nature of the structure functions used in the assimilation process; the spreading of data is controlled by the background error covariances, which have little variation in height. This would result in a reduction of the tilt of the strongly tilted singular-vector perturbation that was present in the observations.

With a decreasing number of synthetic observations assimilated, the SV perturbation structure becomes less well-represented. Figure 6.10 shows the SSV-NOSV analysis increments for v and PV for each of the experiments (with the exception of the N55_grid experiment which is shown in Figure 6.9). The SV perturbation is captured to some extent by all experiments, with the v and PV perturbations exhibiting a tilt against the wind shear which is flatter than for the SV perturbations. The amplitude of the SSV-NOSV perturbation decreases as the number of synthetic sondes assimilated decreases, and the perturbations become more confined. When only 8 sondes are assimilated along the flight track the perturbation is most confined, as the area covered by the sondes is decreased, and the phase of the perturbation is shifted upstream by a quarter of a wavelength (Figure 6.10(e)). When the N55_grid experiment is repeated with the error covariances halved (N55_halfcov) the increased weighting given to the observations during assimilation has made the amplitude of the initial perturbation larger than for the SV perturbations and spurious structures have appeared (Figure 6.10(g)).

6.3.2 Forecast Evolution

The evolution of the tilted structure seen in the analysis increments can be seen by running 24-hour forecasts, and examining the SSV-NOSV forecast differences. Note that the
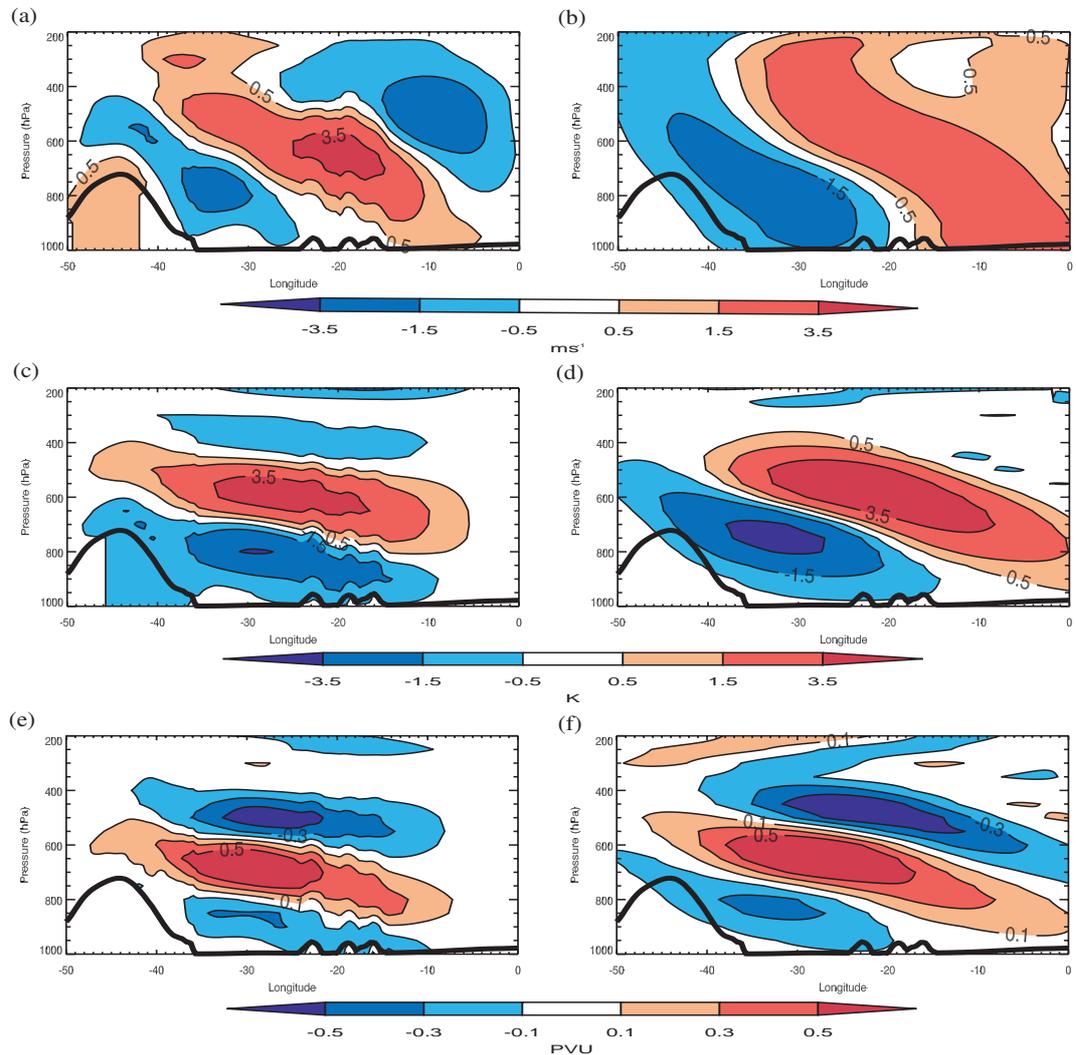
Figure 6.9: *Cross-sections at 66°N from 50°W to 0°W showing the SSV-NOSV analysis increment for the N55_grid experiment (left column) and 50xSV perturbation (right column) for (a) and (b) v-component of wind, (c) and (d) potential temperature, (e) and (f) PV. The surface pressure (thick black line) is used to show the orography.*

50xSV perturbations from the TESV SAP were only available for the initial and evolved singular vector, not at intermediate times, therefore the SSV-NOSV forecast differences can be compared to Figure 6.1(b, f) only. The *v* and *PV* SSV-NOSV forecast differences are shown in Figure 6.11 for the N55_grid experiment. Note that the cross-section plotted has been expanded from the cross-sections plotted in Figure 6.9 (and match that plotted in Figure 6.1) and the contour interval changed. When comparing the SV and SSV-NOSV forecast it is important to remember that the SV was linearly evolved at T42 resolution (approximately 300km),
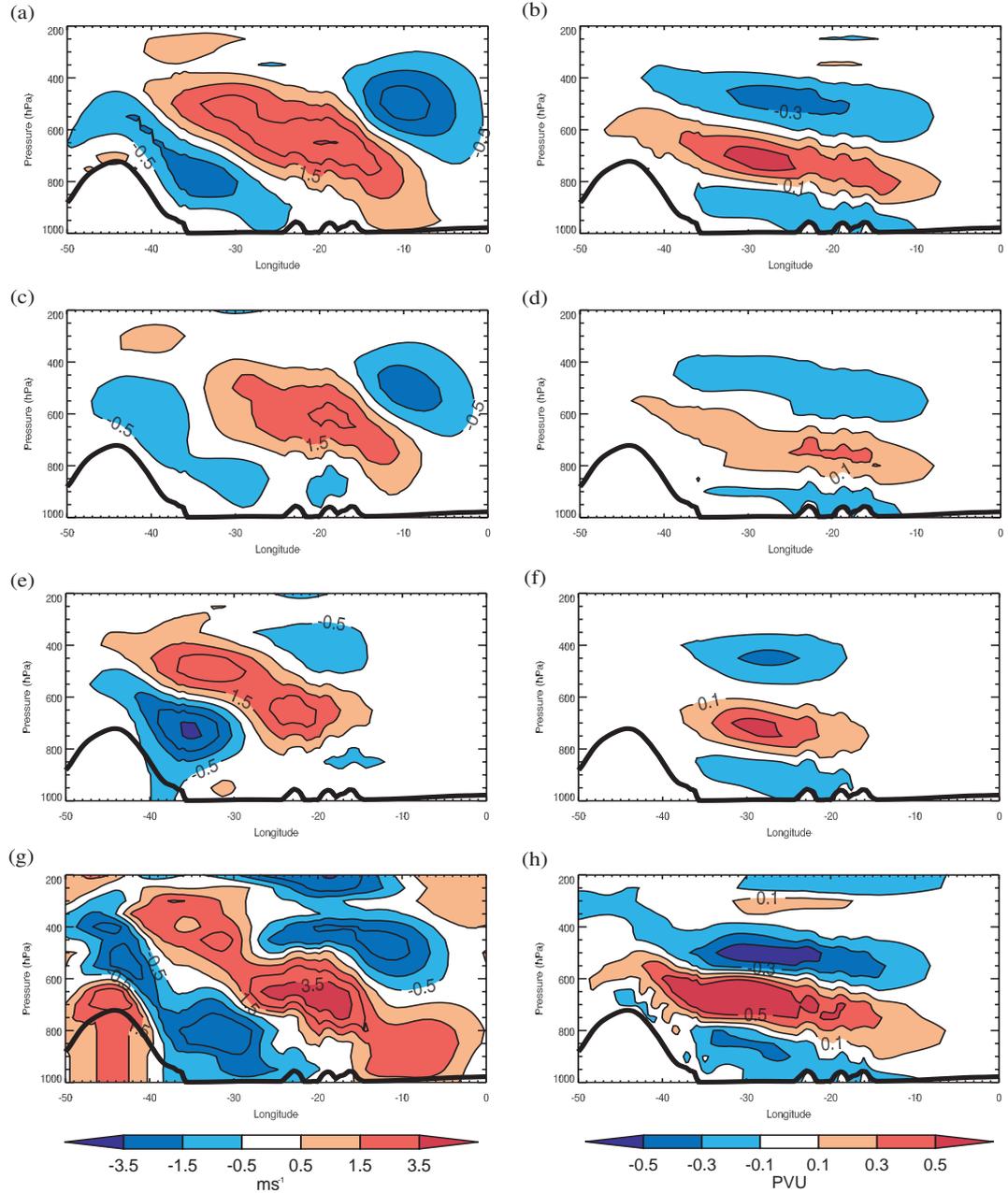
Figure 6.10: *Cross-sections at 66°N from 50°W to 0°W showing the SSV-NOSV analysis increment for the v-wind component (left column) and PV (right column) for the (a) and (b) N30_grid, (c) and (d) N08_grid, (e) and (f) N08_nogrid and (g) and (h) N55_halfcov experiments. The surface pressure (thick black line) is used to show the orography.*

whereas the forecasts were nonlinearly evolved at 24km resolution. Therefore some of the differences between the SV and SSV-NOSV forecasts could be due to the differing resolution (the SSV-NOSV perturbations may contain finer-scale structure than the SV) and nonlineari-

ties. The perturbation appears to be more tilted in Figure 6.11(a); this is due to the change of scale only. The SSV-NOSV perturbation is initially tilted, and untilts as it evolves during the forecast. This untilting can be seen in both the *v* and *PV* components but is clearest in the *v* component, as there is a lot of noise in the PV component, due to the advection of finescale PV anomalies in three dimensions (the perturbation is not planar as in idealised experiments). This is because the background state is not a parallel flow and there are gradients in the PV field (Figure 6.12(b)). As the meridional wind is also non-uniform, it is advecting PV into and out of the cross-section (Figure 6.12(a)). The growth of the SSV-NOSV perturbation is consistent with growth through PV-unshielding (Hoskins et al., 2000). The perturbation becomes less confined as it moves downstream and untilts, and moves up towards the tropopause. The evolved *v* field shows a dipole in the *v* perturbation, which has significant amplitude from 200-800 hPa and is barotropic in structure (Figure 6.11(i)). This is qualitatively similar to the SV perturbation but with reduced amplitude (Figure 6.1(b)). There is also a dipole in the evolved PV perturbation at 200hPa (Figure 6.11(j)), and a positive perturbation below it. The positive part of the dipole structure is greater in amplitude than for the evolved SV perturbation (Figure 6.1(f)); this could be due to the advection of PV as discussed earlier. The evolved forecast perturbation for the N55 experiment is slightly out of phase with the evolved 50xSV perturbation- the downstream propagation of the SSV-NOSV forecast perturbation is faster than for the 50xSV perturbation (following the leading negative PV anomaly at 500hPa). In the linear evolution of the SV, the tropopause cannot deform, which makes the PV dipole seen in the evolved SV higher than for the nonlinear perturbation. There is an upper-level jet at 300hPa visible in the *u*-wind field (Figure 6.12(c)). This controls the speed with which the SV-like perturbations propagate downstream. The position of the leading edge of the wave packet is consistent with the leading edge of the zonal jet. The PV-unshielding and downstream development of the SSV-NOSV forecast is seen for all experiments (not shown).

The total energy contained in the initial and evolved SSV-NOSV perturbations is shown in Figure 6.13. For all experiments the shape and area of the initial energy is a reasonable match to the 50xSV perturbation (Figure 6.4(a)). However the magnitude of the initial energy decreases with a decreasing number of synthetic observations. This agrees with the conclusions from examining the SSV-NOSV analysis increments that the SV perturbation is better approximated with a greater number of sondes. For the N55_halfcov experiment there is spu-
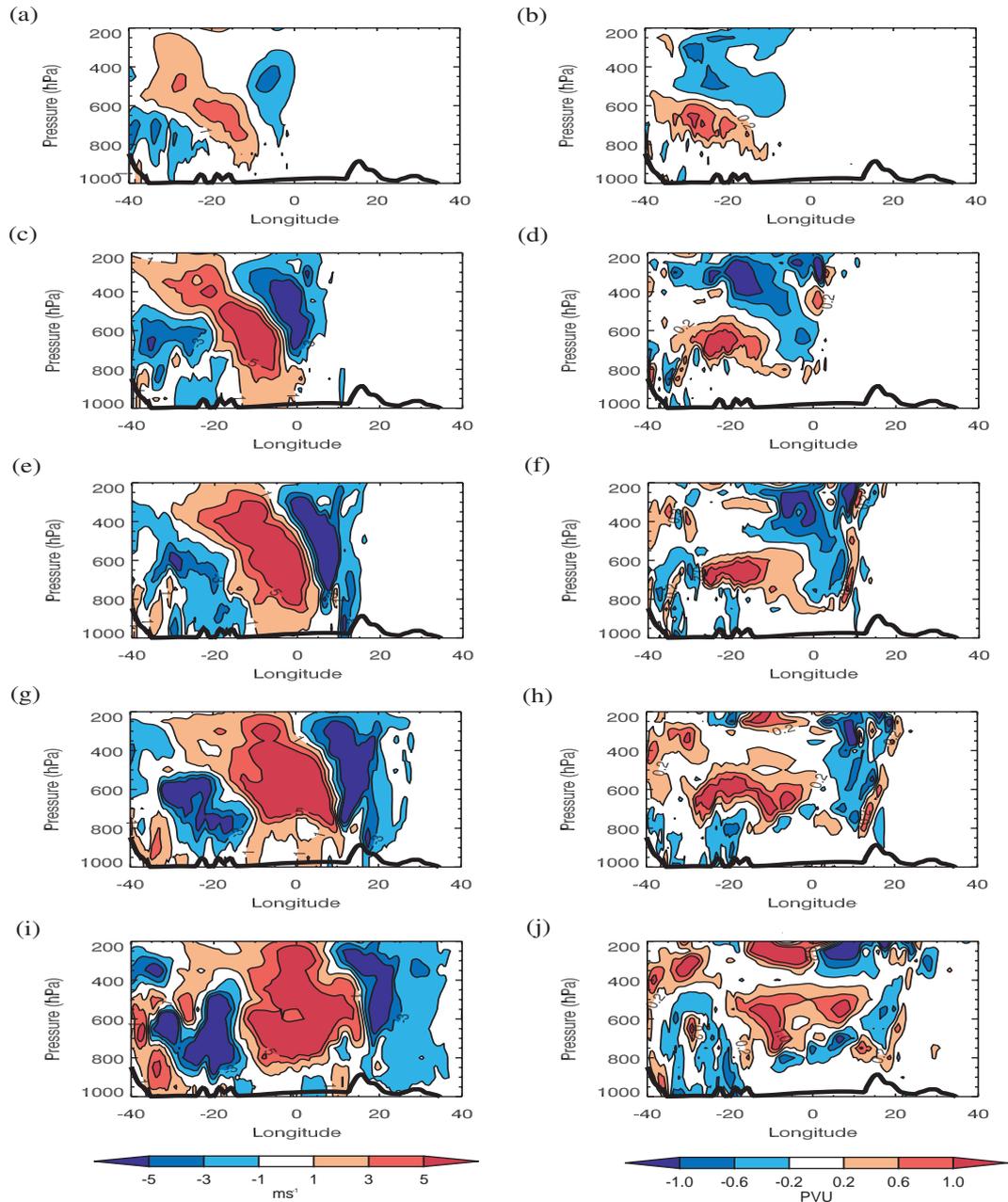
Figure 6.11: *Cross-sections at 66°N from 40°W to 40°E showing the SSV-NOSV forecast difference for v (left column) and PV (right column) for the N55_grid experiment for forecast times of (a) and (b) 0 hours, (c) and (d) 6 hours, (e) and (f) 12 hours, (g) and (h) 18 hours and (i) and (j) 24 hours. The orography is represented using the surface pressure (thick black line).*

rious structure to the initial energy, seen by an extension of the perturbation to the north-west over Greenland, where a second maximum is visible. The initial singular-vector like structure moves downstream to the same location as the SV perturbation for all experiments (compare
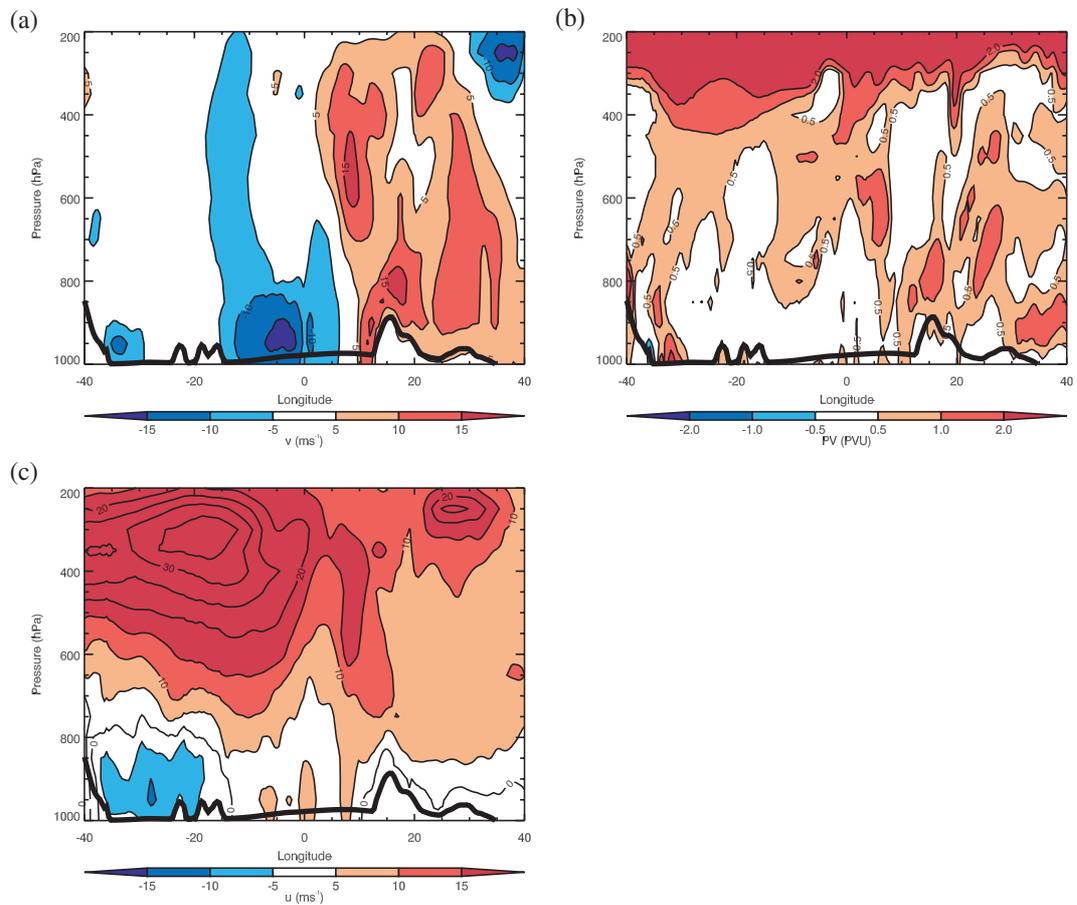
Figure 6.12: *Cross-sections at 66°N from 40°W to 40°E showing the NOSV forecast at T+0 for the (a) v-component of wind, (b) PV and (c) u-component of wind. The surface pressure (thick black line) is used to show the orography.*

the right column of Figure 6.13 with Figure 6.4). However, consistent with the amplitude of the initial energy, the amplitude of the evolved perturbation is also reduced compared to the SV perturbation, and decreases for decreasing number of synthetic observations assimilated. This is also due to the initial SSV-NOSV perturbation covering a smaller area when fewer sondes are assimilated, as the perturbations are more confined. The N08_nogrid experiment is not shown; the total energy of the initial and evolved perturbation are qualitatively similar to the N08_grid experiment. For the N55_halfcov experiment the evolved energy looks similar to the N55_grid experiment, therefore it appears that the spurious structures added by the 4D-Var scheme were decaying structures, as they are not present in the evolved perturbation.
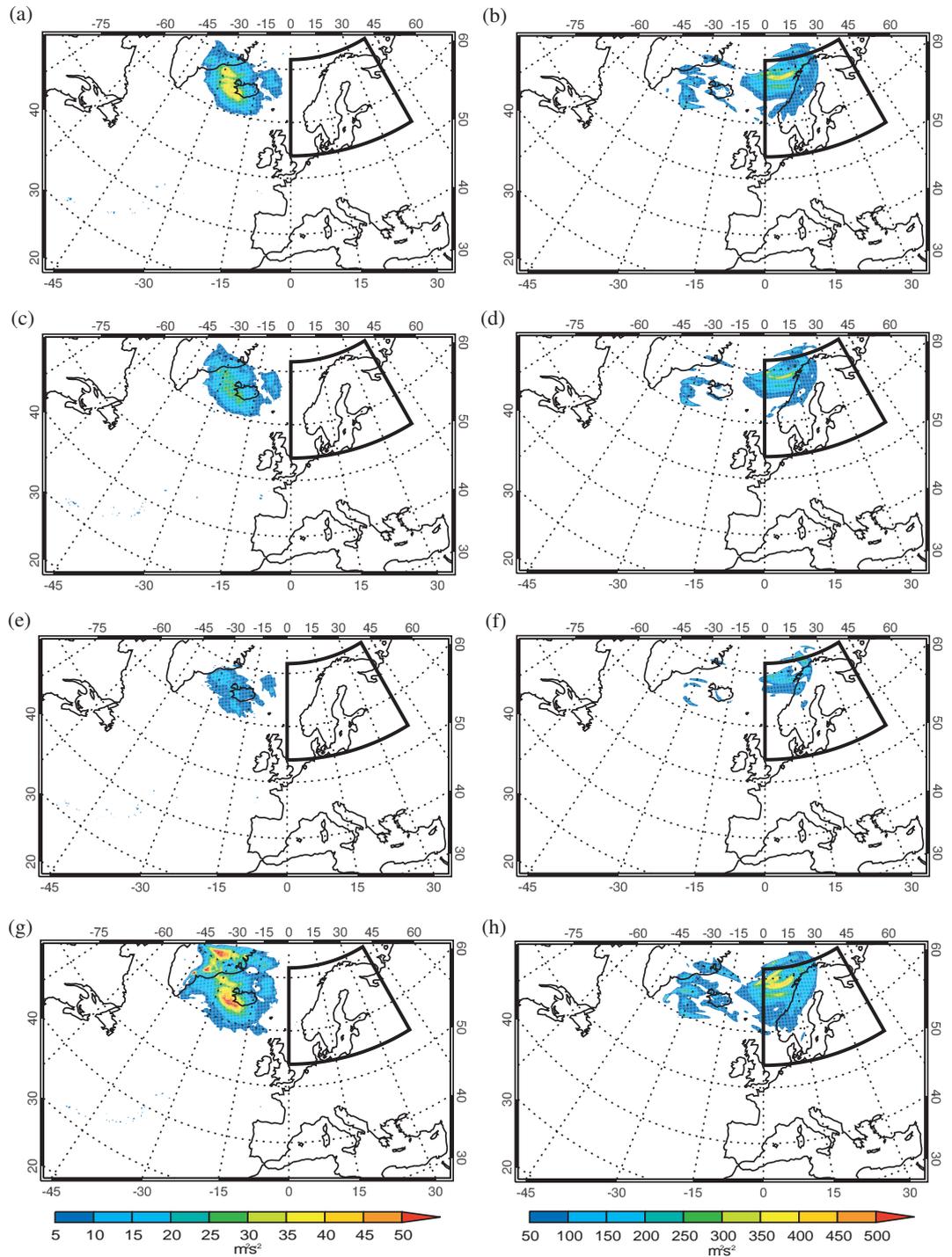
**Figure 6.13:** *Integrated total energy of the SSV-NOSV forecasts at initial time (left column), and after 24 hours evolution (right column) for (a), (b) N55_grid, (c) and (d) N30_grid, (e) and (f) N08_grid and (g) and (h) N55_halfcov. Note the difference of a factor of ten between the scales for the left and right columns. The Scandinavian verification regions are marked by thick black lines.*

6.3.3   Growth Rate of Forecast Perturbation

In order to compare the growth rate of the SSV-NOSV perturbations with that of the 50xSV perturbations, the total energy plotted in Figure 6.13 is spatially integrated. As there is noise present in the SSV-NOSV forecast, particularly at longer forecast times, the total energy is thresholded at each forecast time to exclude noise from the calculation of integrated total energy. As the SV data is only available for the initial and final evolved state, exponential growth is assumed at intermediate times in order to compare the growth of the SSV-NOSV perturbations with those from the SV. Figure 6.14 shows the growth rate of the 50xSV perturbation, and SSV-NOSV perturbations for all experiments, plotted on a log scale on the y-axis. The initial energy for the SSV-NOSV experiments typically less than for the 50xSV perturbation. This gap can be at least partly explained by the observation processing, as assimilating the 55 sondes with reduced error covariances (i.e. giving them greater weighting) increases the initial energy and eliminates this gap. The only experiment to show different behaviour to the 50xSV perturbation is the N08_nogrid, where the total energy reduces in the first 6 hours and then increases again. It is clear that the gradient of the lines is similar for all experiments except the N08_nogrid experiment, meaning that the exponential growth rates are similar. Table 6.1 shows the average exponential growth rate for all experiments and 50xSV perturbations. The exponential growth rate is reduced by 25% for the N55 and N30 experiments, and for the N08_nogrid is slightly larger than the 3.09^{-1} growth rate for the 50xSV. Therefore despite the initial perturbations not fully capturing the 50xSV structure, the growth of the perturbations does closely resemble that of a singular-vector like perturbation.

6.4   Conclusions

This experiment tested the hypothesis that targeted observations have little forecast impact because the assimilation schemes used to assimilate the data are not able to extract baroclinic information from the observations and use it to correct an initial condition error that has a baroclinic structure. Sets of targeted synthetic observations were assimilated, which contained a singular vector structure that was not present in either the background field or the routine observations assimilated. The singular vector structure was provided by the leading TESV from the TESV SAP for 01 March 2007.

It is clear that the data assimilation scheme is capable of extracting baroclinic informa-
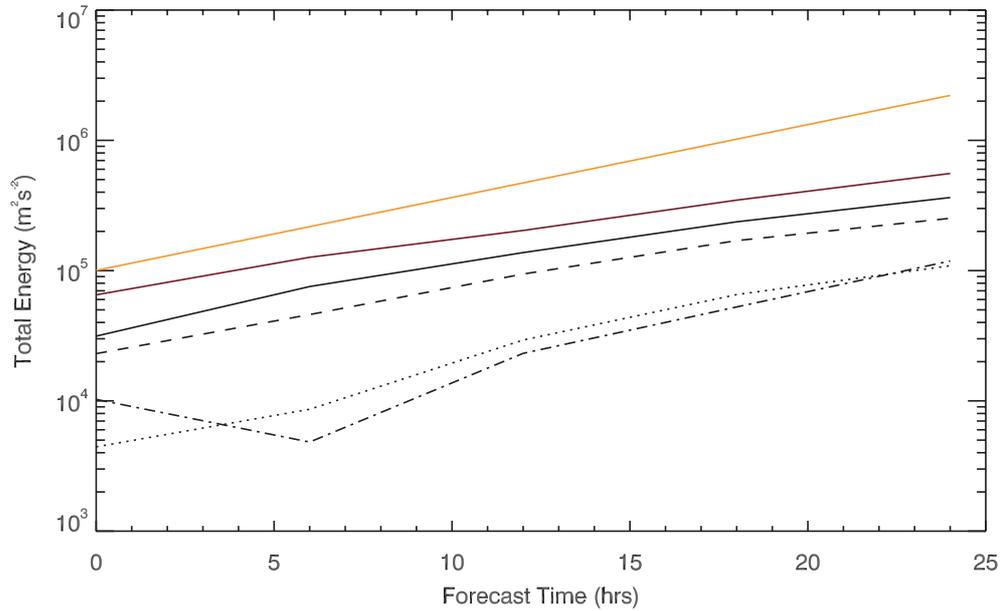
Figure 6.14: *Vertically integrated total energy above a threshold value for the 50xSV perturbations (orange line) and SV-NOSV forecasts: N55_grid (solid black line), N30_grid (dashed line), N08_grid (dotted line), N08_nogrid (dashed-dotted line) and N55_halfcov (solid purple line). The threshold value increases linearly from $10m^2s^{-2}$ to $50m^2s^{-2}$ during the 24-hour period. Values between 0 and 24 hours forecast time for the 50xSV perturbations were obtained by assuming an exponential error growth between the two times. Note the log scale on the y-axis.*

tion from a grid of observations, disproving the hypothesis. The extent to which the data assimilation scheme was able to reproduce the singular vector structure that was present in the observations is dependent on how well the singular vector structure is captured by the targeted observations that are assimilated, and weighting given to the synthetic observations (the observation error variance) during the assimilation process.

The analysis increments show a baroclinic structure most closely representing the singular vector perturbations for a grid of 55 observations (covering an area of approximately $1x10^6km^2$). The structure of the singular vector is reproduced in the analysis increments even with a grid of 8 observations, however both the amplitude and tilt are reduced compared to the singular vector as the number and density of observations are decreased, and the perturbations also become more confined. The reduced tilt in both T and v fields is consistent with thermal wind balance, and can be explained by a smearing-out of the tilted structure by the background error covariance in the assimilation process. For 8 synthetic observations along a

Table 6.1: *Perturbation Growth Rates*

Experiment	Average Growth Rate (day^{-1})
50xSV	3.09
n55_grid	2.45
n30_grid	2.39
n08_grid	3.20
n08_nogrid	2.45
n55_halfcov	2.15

flight track, which captured less of the singular vector structure than the other experiments, a singular-vector structure was seen in the increments, but it was phase-shifted upstream of the SV structure. This result has implications for the targeting experiments detailed in Chapter 4. The number of sondes released in each of the experiments was on the same order of magnitude as the N08 experiments here, but the area covered by the sondes was most similar to the N08_nogrid. This implies that had a tilted structure been present, the observations would have failed to sufficiently capture it, and this would have impaired the ability of the data assimilation scheme to produce a tilted increment. These results are also consistent with Cardinali and Buizza (2003) who found small forecast impact for 10 NORPEX targeting cases due to a mis-match between the location of the targeted dropsondes and the 10 leading SVs, which explained 40% of the forecast error.

The structure and growth of the forecast perturbations most closely resemble the SV for the experiment assimilating 55 sondes. When 30 sondes are assimilated the reduced longitudinal spacing from 2.5 to 5 degrees compared to the 55 sonde experiment makes only a small difference; the assimilation scheme is still able to extract the baroclinic structure from the synthetic targeted observations. All experiments show an untilting of the initial tilted PV and v perturbations, and an upward propagation towards the tropopause that is consistent with growth by 'PV-unshielding' (Hoskins et al., 2000), and the evolved SV structure.

The initial energy contained in the forecast perturbations, as measured by total energy, is smaller than for the SV perturbation for all experiments. The initial energy is decreased as

the number of sondes assimilated is decreased. This gap in initial energy is because the data assimilation scheme does not adjust the background fully towards the observations, as it takes account of the errors in the observations. When the weighting given to the observations is increased by halving the synthetic observation errors, then this initial energy gap is eliminated. This leads to spurious structures in the initial forecast field, which are decaying and therefore do not degrade the downstream forecast. The average growth rate of the total energy of the perturbations over 24 hours is approximately equal for all experiments, and the same or less than the growth rate of the SV itself.

That a 4D-Var scheme can correct a growing error is itself not a new result. Johnson et al. (2006) showed that 4D-Var schemes were capable of adding increments that corrected such growing errors. However, the 4D-Var scheme also added growing errors as it attempted to correct for errors which were in fact decaying in amplitude. This experiment proves that a 4D-Var scheme can extract baroclinic information from a set of as few as eight observations, and use this information to correct the initial condition error by producing a tilted increment.

The position of the observations in the data assimilation time window and the length of the time window can affect the ability of the data assimilation scheme to use baroclinic information contained in the observation. In this experiment observations were placed in the middle of the time window, as this is the time that the SV perturbations were valid for. It would be interesting to compare the results from an identical set of experiments where the observations were placed at 1500 UTC rather than 1200 UTC, i.e. at the end of the data assimilation window rather than the middle. The length of the data assimilation window is a key factor in the ability of the data assimilation scheme to produce a tilted increment. Using single-observation experiments with a 4D-Var scheme, Thépaut et al. (1996) found that a 24 hour window worked best, and that 4D-Var struggled to produce a tilted increment when the data assimilation window was only 6 hours as in this experiment. This is because the structure functions (columns of **B**, i.e. the variation of the background error with height at a particular location) in 4D-Var have a fairly barotropic structure at initial time. Unlike for the Kalman Filter, in 4D-Var **B** is not calculated at each time step within the 4D-Var window, however it is allowed to evolve during the assimilation window. Therefore given a sufficiently long time window the vertical structure of **B** will begin to develop baroclinic structures, and so will be more able to make use of baroclinic information contained in an observation. Therefore

repeating this experiment with a longer data assimilation window may improve the ability of the data assimilation scheme to extract the SV structure from the observations, when only a few observations capture this structure.

CHAPTER 7

Targeting of a Polar Low in the Norwegian Sea

In this chapter the impact of targeted observations on the forecast of a polar low in the Norwegian sea is examined. In a setup that differs from the conventional targeting method, only the fraction of the sensitive area covered by the polar low was targeted, with a dense network of sondes, to try to improve the forecast of the polar low landfall. This is the first time that a small-scale system such as a polar low has been targeted.

7.1 Background

Observational studies of polar lows have attempted to classify polar lows according to their cloud structure, or the conditions in which they form. One sub-type of polar low is the reverse-shear polar low. In this chapter other types of polar low are neglected to focus only on reverse-shear polar lows in the Norwegian sea, in order to provide background to the targeting case study of such a polar low which occurred in March 2008. Reverse-shear flow was first defined by Duncan (1978) as an airflow which is in the opposite sense to the thermal wind in that layer. In the northern hemisphere, the direction of thermal wind is such that there is cold air to its left (the jetstream is one example of a thermal wind), thus a reverse-shear flow has warm air to its left and cold to its right. The term is somewhat confusing, as the direction of shear itself is not reversed, rather the low-level wind is in the opposite direction to the direction of shear (Figure 7.1); the vertical wind shear is in thermal-wind balance as normal. The reverse-shear flow at low levels in the Norwegian Sea is often set-up by the passage of a synoptic-scale low-pressure system, drawing southwards cold dry Arctic air which has been trapped over the ice sheet, as the system moves eastwards across the Norwegian Sea. At the same time the cyclonic circulation associated with the low-pressure system advects northwards relatively warm air from the south. Reverse-shear polar lows are therefore commonly observed to form in the Norwegian sea during cold air outbreaks which follow the passage of a synoptic-scale system. A climatology of reverse-shear conditions constructed using 40-year ECMWF Re-analysis (ERA-40) data by Kolstad (2006) showed that in the Norwegian

sea conditions are favourable for reverse-shear development 10-15% of the winter months. Reverse-shear polar lows are strongest in the lowest part of the atmosphere. The structure of a reverse-shear polar low compared to a typical baroclinic mid-latitude cyclone is shown in Figure 7.2. Relative to the direction of travel, the surface low-pressure centre of the baroclinic mid-latitude cyclone lies ahead of the upper-level trough (Figure 7.2(a) and (b)), whereas the surface low-pressure centre of the polar low lies to the rear of the upper-level trough (Figure 7.2(c) and (d)). In the classical baroclinic case, warm-air advection leads to ascent ahead of the surface low and cold-air advection to descent to the rear of the surface low. This situation is reversed for reverse-shear, with warm-air advection and ascent to the rear of the surface low and cold-air advection and descent ahead of the surface low. The typical length and height scales of both a baroclinic cyclone and reverse-shear polar low are indicated on Figure 7.2. The scale-height H is related to the Rossby radius of deformation L by:

$$H = \frac{fL}{N},\tag{7.1}$$

where f is the Coriolis parameter and N the Brunt-Vaisala frequency, taken to be $10^{-2}\mathrm{s}^{-1}$. The latitude of the baroclinic cyclone was taken to be 50°N, and 70°N for the polar low. Using typical values for L, the values of H obtained show that the baroclinic system has a scale height of 5km, much larger than the 1.5km for the polar low. N was taken to be the same for both systems; H therefore decreases as the Rossby radius of deformation decreases and the latitude at which the system forms increases. However during cold-air outbreaks, N may become very small and therefore H can become large. This is especially true if the polar low forms in a region where the synoptic-scale flow is cyclonic. In these conditions (7.1) is modified to:

$$H^2_{mod} = \frac{f(f + \xi_{av})L^2}{N^2},\tag{7.2}$$

where H^2_{mod} is the modified scale-height and ξ_{av} is the relative vorticity associated with the synoptic-scale flow (Rasmussen and Turner, 2003). Taking ξ_{av} to be equal to $2f$, and substituting into (7.2), $H_{mod} \approx 2H$.

Many polar lows are baroclinic, and form when an upper-level trough with positive vorticity moves over a low-level baroclinic zone where there is a gradient in the temperature field,
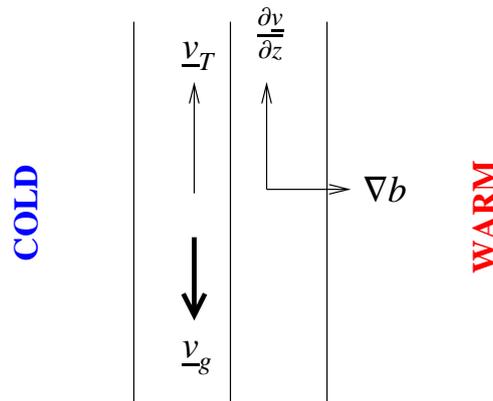
Figure 7.1: *Reverse-shear flow. The arrows indicate the direction of the thermal wind ($\underline{v}_T$), the wind shear ($\partial \underline{v}/\partial z$), the gradient in buoyancy (∇b) and the direction of the low-level 'reverse-shear' flow $\underline{v}_g$. The black lines are isotherms.*

i.e. a front. This type of baroclinic development is similar to the Type B cyclogenesis for mid-latitude cyclones put forward by Pettersen and Smebye (1971). Nordeng (1990) studied two polar lows which formed in the Norwegian sea, via this mechanism. It was suggested that air-sea interaction could have been responsible for the further development of one of the polar lows studied. Air-sea interaction has been investigated as another development mechanism for polar lows. The two mechanisms most studied are conditional instability of the second kind (CISK; Charney and Eliassen (1964)) and wind-induced surface heat exchange (WISHE; Emanuel and Rotunno (1989)). Rasmussen (1979) proposed that polar lows developed via the CISK mechanism, where convection is triggered by the lifting of air parcels in a region of low-level convergence (and cyclonic vorticity) to their level of free convection. The latent heat release associated with the formation of cumulus cloud will then create divergence, which will enhance the low-level convergence. This mechanism requires an initial disturbance and an atmosphere which is conditionally unstable (i.e. there is convective available potential energy (CAPE) present). Emanuel and Rotunno (1989) argued that WISHE, which does not require a pre-existing reservoir of CAPE, may better explain polar low development. In the WISHE mechanism, low-level convergence leads to increased surface fluxes of sensible and latent heat, and convection. Polar lows are most commonly observed in the winter months (Wilhelmsen, 1985) when there is potential for large air-sea temperature differences giving rise to low static stability, and potential for large air-sea fluxes. Thus polar lows often form near the boundary between the sea ice and open ocean, where the air-sea

(a)

(b)

(c)

(d)
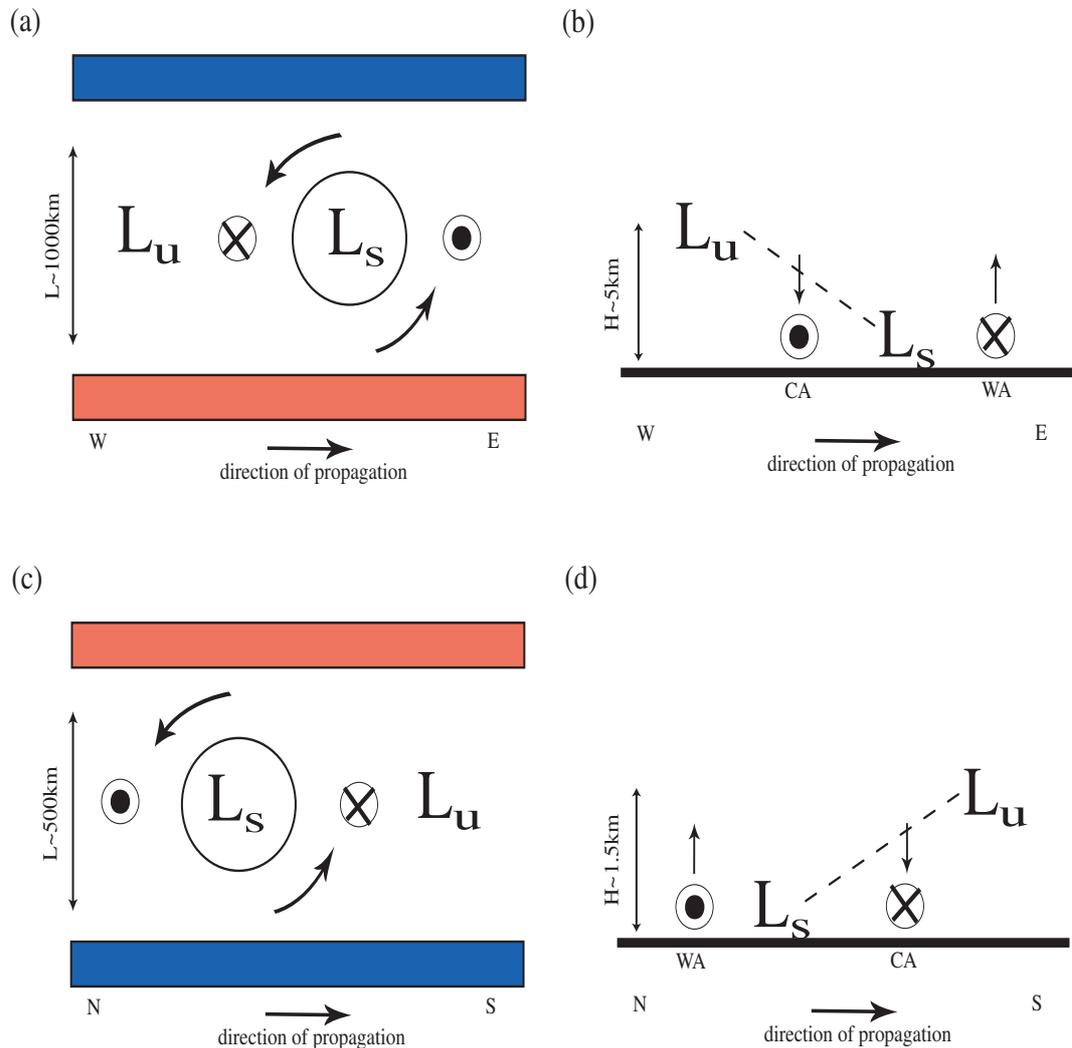
Figure 7.2: *A comparison of the structure of a normal baroclinic cyclone (a) and (b) and a reverse-shear polar low (c) and (d). Panels (a) and (c) show a plan view looking down on the cyclone, where L_s is the centre of the surface low and L_u the position of the upper-level trough. WA (CA) is warm (cold) air advection, and the red (relatively warm air) and blue (relatively cold air) colours indicate the direction of the thermal gradient. The circle with a cross through it indicates descent resulting from cold-air advection, and the circle with a dot in it indicates ascent resulting from warm-air advection. Panels (b) and (d) show a vertical cross-section through the cyclone, where the dashed line indicates the direction of tilt of the pressure field with height. Here the circle with a cross through it indicates wind directed into the page, and the circle with a dot in it indicates wind directed out of the page, due to the cyclonic flow around the cyclone. L indicates the Rossby radius of deformation, H is the scale-height relating to L (see text). The co-ordinates have been rotated on all figure panels so that the direction of propagation of the system is always from left to right.*

temperature difference suddenly increases, and static stability decreases. During cold-air out-breaks, where cold dry Arctic air is advected southwards over open ocean, the cold air can be advected far from the ice edge, and polar lows form much further south. In assessing the likelihood of polar low formation on any given day, the Norwegian Meteorological Service (Met.No) look for regions where the temperature difference between the sea surface and air at 500hPa exceeds 43°C (personal communication, Gunnar Noer, Met.No). The Norwegian Atlantic current brings warm salty water north from the northern Atlantic into the Norwegian sea and results in warm sea-surface temperatures (SSTs) near the Norwegian coast, giving potential for large sea surface temperature to air temperature differences.

In an attempt to better understand polar lows, many climatologies have been constructed using a variety of methods, from objective identification using mslp charts to subjective iden-tification using infra-red satellite images. Wilhelmsen (1985) identified polar lows in the Nordic (Norwegian and Barents) seas using mslp analyses, from 1972-1982. Of the total of 71 polar lows identified in the 10 year period, only two occurred in the period May - August, with a peak in the winter months. There was also a distinct nadir in February. Harold et al. (1999a) also identified a maximum in polar mesocyclones in winter months, following sub-jective identification from infra-red satellite images over a 2-year period in the mid-1990s. Harold found that the larger mesocyclones, 200-600km in diameter, tended to form further from the ice-edge than smaller mesocyclones with a diameter less than 200km (Harold et al., 1999b). The north Norwegian sea was one of the areas that saw the greatest density of polar mesocyclone formation, the Greenland Sea being the other region.

Harold's study was not able to distinguish between polar mesocyclones and polar lows, as no wind speed data was available to check for winds greater than $15ms^{-1}$ which differ-entiates polar mesocyclones and polar lows. A recent study by Blechschmidt (2008) used thermal infra-red satellite imagery in conjunction with satellite-derived winds speeds from the Special Sensor Microwave Imager (SSM/I) to identify polar lows occurring in the Nordic seas in 2004 and 2005. The majority of polar lows identified in this study formed in the Nor-wegian sea, which supports the results of Harold et al. (1999b). In addition to this, 22% of polar lows identified formed in conditions of reverse-shear during this 2-year period. Neither Blechschmidt (2008) or Bracegirdle and Gray (2008) who objectively identified polar lows from a database of cyclonic features extracted from Met Office global analyses, identified the

February minimum first reported by Wilhelmsen (1985). Bracegirdle and Gray (2008) limited their study to polar lows which formed during cold-air outbreaks in the Nordic seas; such conditions suggest that a substantial proportion of polar lows they identified could be classified as reverse-shear. Bracegirdle and Gray (2008) found a difference in structure between polar lows forming in the north and south of the Nordic seas. Polar lows forming in the north of the region, like the polar low targeted during IPY-THORPEX, tended to have stronger baroclinicity at low levels and weaker convection, a situation reversed for polar lows which formed further south further from the ice-edge, where the atmospheric static stability was lower.

Despite improvements to numerical weather prediction models and the increased use of satellite observations, producing accurate forecasts of polar lows is still a challenge (Condron et al., 2006). One of the challenges is the lack of routine meteorological observations in the Arctic with which to constrain weather forecast models. The European Arctic can be considered data-sparse, particularly with respect to high-resolution profile observations. Section 2.1 discussed the shortage of radiosonde and AMDAR data in this region, and the fact that the satellite radiance profiles (which have a good data coverage over this area) are only available above the cloud tops. As polar lows form over the ocean in poorly-observed regions, initial condition errors may contribute significantly to the forecast error. Targeted observations attempt to reduce the initial condition error in data-sparse regions and may represent a way to improve the forecasts of these systems.

Targeted observations were made of a reverse-shear polar low in the Norwegian sea during the Norwegian IPY-THORPEX field campaign. This took place over three weeks in February and March 2008. The main observational platform for the campaign was the DLR aircraft, which was deployed from Andenes (69N, 16E, see Figure 1.1) in Norway, into the data sparse regions to the north and west. In this chapter the set-up of the targeting experiment is described in Section 7.2, and the hindcast experiments in Section 7.3. A synoptic overview of the polar low case is given in Section 7.4 and the results described in Sections 7.5 to 7.7.

7.2 Set-up of Targeting Experiments

7.2.1 Sensitive Area Predictions

During the field campaign, sensitive area predictions were provided by the ECMWF Data Targeting Suite (DTS) which was under trial at ECMWF during this period. The DTS is part of the Eurorisk-PREVIEW programme. The DTS allows the user to view ensemble forecast products to determine if there is a likelihood of a severe weather event, and then to request SAPS to be run for a verification region and optimisation time of choice. For the purposes of our field campaign, both TESV and ETKF SAPs were received for a lead time of 36 hours and optimisation times of 12, 24 and 36 hours, for a fixed verification region over Scandinavia (Figure 7.3), although we were able to request additional sensitive area predictions as well. Note that the position and size of the verification region, although similar to the Scandinavian verification region used in GFDex, could not be made exactly the same as the size of verification region used in the DTS was fixed. The verification region was positioned to catch polar lows moving south and affecting the Norwegian coast.

The sensitive area predictions were run at too low a resolution to resolve the polar low itself (the ETKF was run on a 2.5 degree grid and the TESVs at T42 resolution), but could resolve the cold-air outbreak associated with the occurrence of the polar low. During the field campaign, there were two periods of cold air outbreaks, during which time the ETKF SAP consistently identified a sensitive region over the Norwegian and Barents seas, to the north of Norway, irrespective of optimisation time. An example ETKF SAP for this synoptic situation is shown in Figure 7.3(a). It is unsurprising that with low-level northerly flow the ETKF should identify the data-sparse oceans that are the origin of the northerly flow, as the ETKF includes an approximation to the radiosonde network, and thus 'knows' about the lack of observations over the Arctic oceans. The TESV SAPs showed little sensitivity in the region of the Norwegian sea during either the cold air outbreaks or during periods of westerly or south-westerly flow. It was observed during GFDex that the location of the TESV sensitive areas was often upstream of the verification region, with respect to the direction of the 500hPa flow, and the distance from the verification region proportional to the strength of this upper-level flow. During the cold-air outbreaks the upper-level flow was westerly, and the TESV sensitive areas were west of the verification region (Figure 7.3(b)).

(a)

(b)
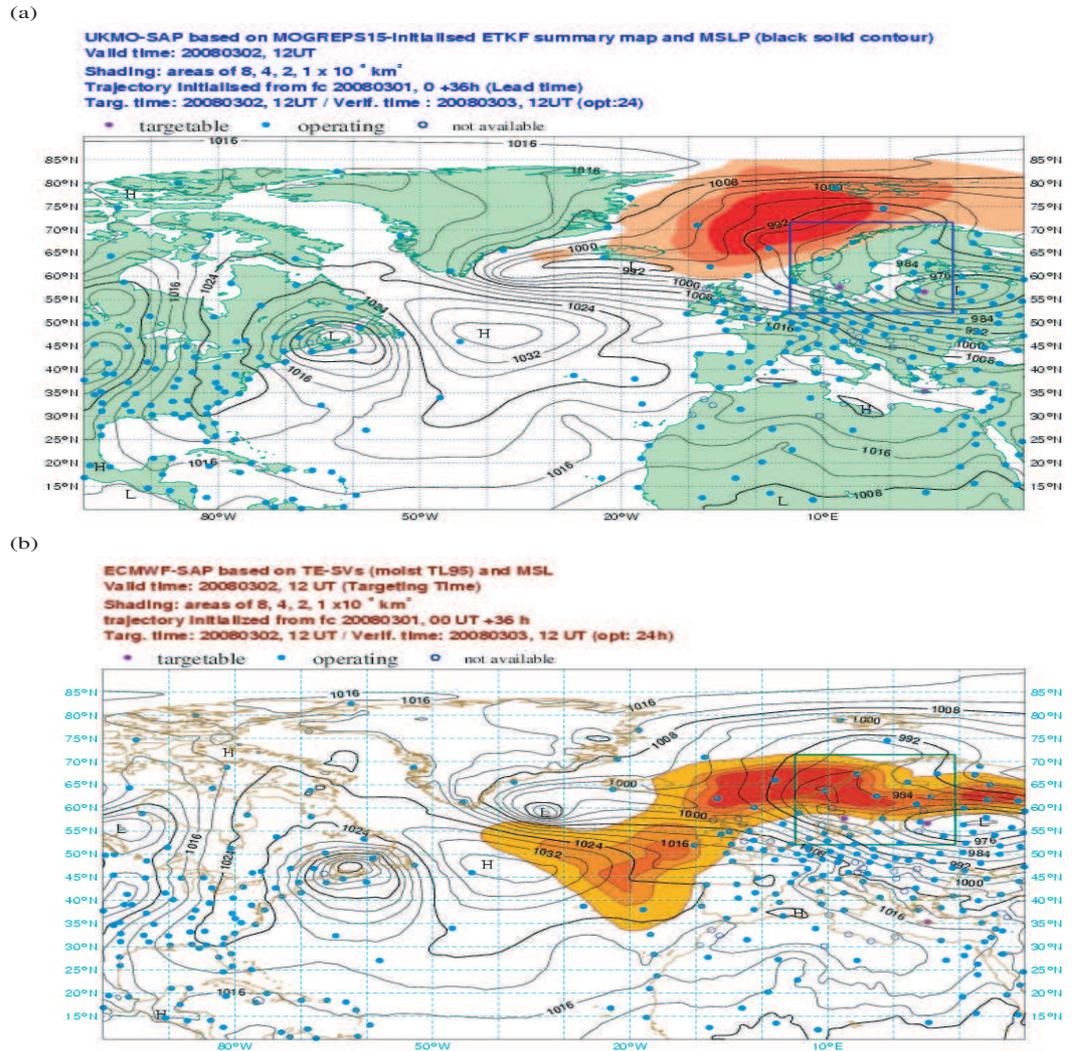
Figure 7.3: *(a) ETKF SAP and (b) TESV SAP during a cold air outbreak, for a 36hr lead and 24hr optimisation time. The shading indicates the sensitive regions, with the most sensitive 1, 2, 4 and 8x10^{6}km^2 of the forecast domain shaded. The darkest shading indicates the region of maximum sensitivity. The black contours are a 24-hour forecast of the mean sea-level pressure valid at targeting time. The verification region over Scandinavia is outlined.*

7.2.2   Targeted Flight Details

For dropsonde targeting experiments the target region is limited by the range of the aircraft used. The DLR aircraft that was the observational platform for the IPY-THORPEX field campaign has a maximum range of 3148km at 31,000ft. The main target area was the Norwegian Sea, to the north and west of Andenes where the aircraft was based during the campaign. The ability to use Longyearbyen (78°N, 15.5°E, see Figure 1.1) in Spitsbergen as

a re-fuelling stop increased the north and west extent of the area that it was possible to target, into the north Norwegian and south Greenland seas. The eastward extent of the target region was confined to 30°E, for political reasons, therefore it was not possible to target polar lows which developed in the Barents Sea.

Previous field campaigns, (see Langland (2005) for an overview) including GFDex, have all followed a similar methodology for targeting observations. Firstly, the verification region is defined. This is the region that we wish to improve the forecast for. Next, sensitivity methods are used to determine where in the forecast domain the forecast for the verification region is sensitive to initial condition errors, i.e. where an initial error may result in a large forecast error downstream. Lastly, the flight is designed and implemented, which tries to cover as much of the target region as possible with additional, usually dropsonde, observations. The dropsondes are spaced according to the horizontal correlation length-scales assumed by the data assimilation system, which are on the order of 200km (discussed in Section 4.2.3). However, this targeting methodology may not be suitable when the the key dynamical feature that will affect the forecast in the verification region has itself a scale of only a few hundred kilometres.

The targeting methodology followed during IPY-THORPEX differed from the above by identifying the sensitive region, and then exclusively targeting the dynamical feature within the verification region - the polar low - with a dense network of dropsondes. Polar lows are small-scale features with a diameter of 200-1000km, so this approach leaves much of the sensitive area un-sampled by targeted observations. The spacing of the dropsondes was smaller than in GFDex, with a minimum spacing of around 50km, i.e. larger than the 48km grid-box used in the data assimilation scheme, but smaller than the assumed length-scales of around 200km (see Section 4.2.3) used by the 3D-Var scheme used to assimilate the data for the hindcast experiments described in Section 7.3.

The polar low was analysed with in-situ observations made during three flights at different stages of its life-cycle: the first flight at 1200 UTC on 03 March 2008 was made into the northerly airstream across the Arctic front, as a polar low was forming (Figure 7.4(a)), the second flight six hours later at 1800 UTC released sondes into the newly-formed polar low (Figure 7.4(b)), and the third flight at 1200 UTC the following day was into the mature polar low (Figure 7.4(c)), six hours before it made landfall. For the purposes of this study, the
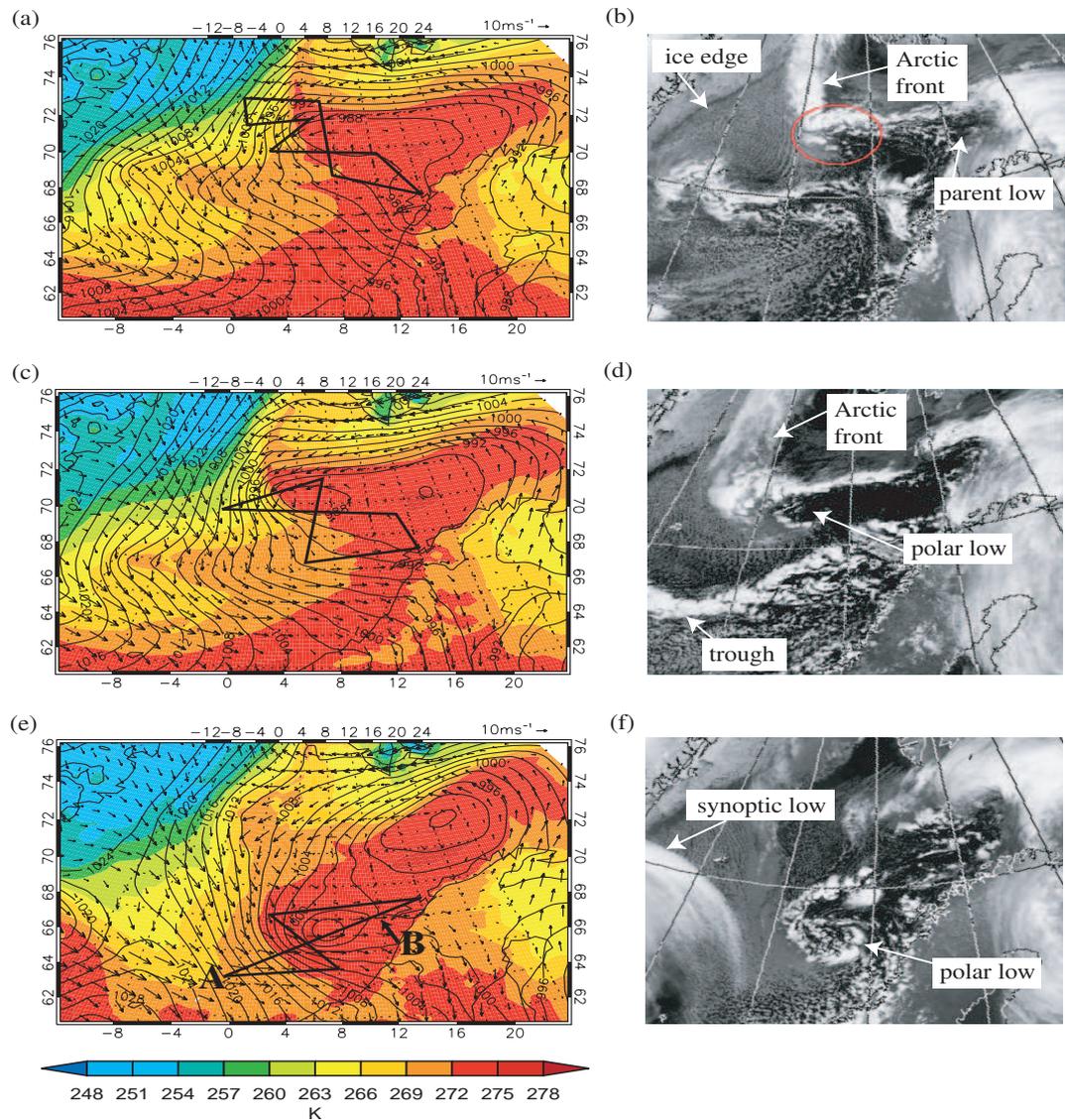
Figure 7.4: *The flight tracks (thick solid line) for the targeting flights at (a) 1200 UTC 03 March, (c) 1800 UTC 03 March and (e) 1200 UTC 04 March. The synoptic situation at targeting time is shown by the ECMWF analysed mean sea-level pressure (black contours), potential temperature at 950hPa (coloured shading) and vector wind at 950hPa (arrows). Infra-red satellite images from the Advanced Very High Resolution Radiometer (AVHRR) instrument onboard the NOAA-18 satellite, showing the cloud cover over the north Norwegian sea at times corresponding to the flights through the polar low: (b) 1145 UTC 03 March, (d) 1921 UTC 03 March and (f) 1135 UTC 04 March. The red circle on panel (a) indicates the location of the developing polar low. The letters A and B on panel (e) correspond to the start and end points for cross-sections used later in the chapter.*

first and second flights are treated as targeting flights and data from the third flight is used to check the quality of the ECMWF analyses used to verify the data (Section 7.6.1). Data from all three flights was sent to the GTS; however, only data from the second and third flights were operationally assimilated as data from the first flight missed the cut-off time for the operational forecast run at the Met Office (although it was used in the update forecast run which is used to provide the background for the subsequent forecast). The flight tracks were designed to be able to study the structure of the polar low, and to capture the gradients in low-level wind to the west of the polar low. The first flight also overflew a decaying polar mesocyclonic feature at 71°N,10°E, which is visible as a mass of convective cloud on the infra-red satellite image to the north-west of Andenes (Figure 7.4(b)).

7.3 Design of Hindcast Studies

7.3.1 Model Description

To analyse the impact of the targeted observations, the forecasts were re-run using a new limited-area domain set-up using the Met Office CAMM. For this experiment the CAMM was set-up to use a 3D-Var data assimilation scheme on a 48km grid, and the Unified Model on a 24km grid. The forecast model therefore has the same set-up as for the GFDex experiments, albeit on a slightly modified area (shifted north and east relative to the NAE domain). The NAE domain used to analyse the GFDex data was deemed unsuitable to use for this experiment, as it does not extend far enough north or east to cover the area that sondes were released in during the IPY-THORPEX field campaign. The biggest difference between the experiments detailed in this chapter and those detailed in Chapters 4, 5 and 6, is the use of the 3D-Var assimilation system, which does not take into account the time that the observations were made, and also has static forecast error covariances. It would have been preferable to use a 4D-Var data assimilation scheme to be consistent with the GFDex experiments, but this was not set-up to work with the CAMM at the time of the field campaign. Kelly et al. (2007) showed that using 4D-Var can give a smaller impact than 3D-Var, by the propagation of information from data-rich areas to data-poor areas. However the lack of observations in the Arctic region means that this is not likely to be an important factor in this case, therefore it is speculated that the use of time-varying covariances in the 4D-Var scheme would result in a larger impact from using 4D-Var in this case, as was seen in Bergot (2001) and Liu and Zou

(2001). The lateral boundary conditions came from the Met Office operational global forecast run.

7.3.2   Description of Hindcast Runs

The hindcast runs performed for the Norwegian targeting experiments are similar to the GFDex hindcast runs detailed in Chapter 4 and therefore the same naming convention has been used. The different experiments are detailed in Table 7.1. A CONTROL forecast cycle was run (Table 7.1), starting from the 1200 UTC forecast cycle on 03 March and finishing with the 1800 UTC forecast cycle on 03 March, assimilating only routine observations. The forecasts were run out to 30 hours beyond the analysis time. Note that the dropsondes released at 1200 UTC on 03 March were the first dropsondes from the IPY-THORPEX field campaign to be operationally assimilated into the forecast, therefore there was no need to run the CONTROL cycle for a longer period of time to eliminate the effects of previous targeted observations, as was done for the GFDex experiments.

A series of TARGETED hindcasts were then run (Table 7.1). Note that the numbers attached to the end of the hindcast name refer to the analysis time for each hindcast, e.g. '12' refers to an analysis time of 1200 UTC. The first two runs are equivalent to the TNOMEM runs in GFDex, and the last to the TMEM run in GFDex. The TNOMEM_12 run took the CONTROL background for 1200 UTC and assimilated routine and targeted observations for the 1200 UTC forecast window (i.e. assimilated targeted observations from the first flight) only. The TNOMEM_18 run took the CONTROL background for 1800 UTC and assimilated routine and targeted observations for the 1800 UTC forecast window (i.e. assimilated targeted observations from the second flight) only. The TMEM_18 run took the CONTROL background and assimilated routine plus targeted observations for the 1200 UTC forecast window, and then used this as the background for the 1800 UTC forecast, where routine plus targeted observations were assimilated. The TMEM_18 targeted forecast was then run from 1800 UTC, therefore the TMEM_18 forecast has memory of the targeted observations from the previous (1200 UTC) cycle.

During the field campaign the radiosonde network in the Arctic region was supplemented by additional radiosonde launches from Bjørnøya, Ny Alesund and Murmansk (see locations on Figure 1.1), and the data from these sondes assimilated operationally into the forecast

Table 7.1: *Description of hindcast runs*

Hindcast	Targeted Sondes		Analysis Time
	1200 UTC	1800 UTC	
CONTROL	N	N	12, 18
TNOMEM_12	Y	N	12
TNOMEM_18	N	Y	18
TMEM_18	Y	Y	18

models. There were also radiosonde ascents from two Norwegian coastguard ships operating in the Norwegian sea. Unfortunately the coastguard ships were unwilling to disclose their position in real-time, and so these ascents were not assimilated into the operational forecasts. Radiosonde launches from Franz-Josef Land were also planned, but were prevented due to technical difficulties. The additional radiosonde ascents that were assimilated operationally were not removed from the hindcasts and were treated as routine observations in the hindcast runs, regardless of their position.

7.4   Synoptic Overview of Case Study

At targeting time at 1200 UTC on 03 March 2008, the polar low was in the process of forming, visible as a mass of convective cloud in the infra-red satellite image (circled on Figure 7.4(b)). On this image the shape of the convective cloud indicated that there was vorticity present, but there was no closed mslp contour on the analysed mslp chart at this time (Figure 7.4(a)), just a large area of low pressure situated north of Norway. This is intersected by the southern end of an Arctic front, visible as a sharp gradient in the 950hPa potential temperature contours and as a north-south oriented cloud band at $4°$E on the infra-red satellite image. At this time another cyclonic-cloud feature is visible to the south-west of the developing polar low ($70°$N,$10°$W). There is also evidence of a cold-air outbreak, with low-level cloud streets oriented parallel to the low-level wind, that begin just south of the ice edge, where the air-surface temperature difference suddenly increases.

At targeting time at 1800 UTC on 03 March 2008, a polar low is visible both as a closed contour circulation on the mslp analyses (Figure 7.4(c)) and on the infra-red satellite image

(Figure 7.4(d)). The polar low is embedded in an area of low pressure on the flank of a synoptic-scale low, which has its centre further east. The polar low has formed in flow which has reverse-shear at low levels; there is northerly flow in an area with a low-level east-west temperature gradient. The polar low at this time is baroclinic, with an upper-level trough situated ahead of the polar low (not shown). The forward (i.e. in the direction of movement) tilt with height is the opposite to normal baroclinic flow, and is due to the reverse-shear. The system at this time has warm air advection behind (to the north) of the polar low and cold air advection ahead (to the south) of the polar low (Figure 7.2(c)). The satellite image at the time shows streets of fairly shallow convective clouds in the cold northerly airstream coming off the ice by the coast of Greenland, to the west of the polar low. There are bands of deeper convective cloud associated with fronts: one aligned north-south along the low-level Arctic front, one to the north of the low pressure system, and one trough also associated with the parent low-pressure system, that shows some curvature on the western edge of the cloud band. There is some curvature seen in the cloud pattern where the polar low has formed, but at this time there is no clear comma or spiraliform structure.

In the period from targeting until the polar low makes landfall, the track of the polar low is southerly, towards the coast of Norway. At the time of the third flight at 1200 UTC on 04 March, the polar low is still some distance north of the Norwegian coastline. The satellite image at the time of the third flight through the polar low shows convective cloud in a spiraliform shape (Figure 7.4(f)). The polar low is now a mature system, with strong near-surface winds (950hPa) to the west of the centre, in the northerly airstream (Figure 7.4(e)). The polar low continued to move towards the south in the following six hours and made landfall near Trondheim on the Norwegian coast at around 1800 UTC on 04 March. The analysed 10m winds at the time of landfall were on the order of 21ms^{-1}, and the radar images (not shown) show heavy precipitation which would have fallen as snow due to the sub-zero temperatures. The polar low dissipated quickly once it made landfall, although some remnant cloud structures could be seen moving south-east over Norway during the 6 hours following landfall (not shown).

7.5   Initial Impact of Targeted Sondes

To determine the initial impact of the targeted sondes on the forecast, the analysis increments for the targeted forecasts can be compared with the analysis increments from the CONTROL forecasts. Figure 7.5 shows the vertically averaged difference in wind increment, and difference in pressure increment at 850hPa for the TNOMEM_12-CONTROL and TNMEM_18-CONTROL runs. The TMEM_18-CONTROL analysis increment difference is not shown, as it is similar to the TNOMEM_18-CONTROL analysis increment difference. The initial impact of the targeted sondes at 1200 UTC is to deepen the area of low pressure that the polar low is forming in. They also strengthen the north-south pressure gradient, resulting in a strengthening of the northerly winds within the northerly airstream to the west of the area of low pressure. The changes to the winds along the northernmost part of the flight track have resulted in a strong anticyclonic increment north of the flight track, and a compensating cyclonic increment over Svalbard. The targeted sondes at 1800 UTC appear to shift the circulation associated with the developing polar low to the east. This shift in position is seen throughout the troposphere, and has a barotropic structure. There is also a marked weakening of the northerly winds to the west of the polar low. This is in contrast to the effect of the targeted sondes in the previous forecast cycle at 1200 UTC, which strengthened this flow.
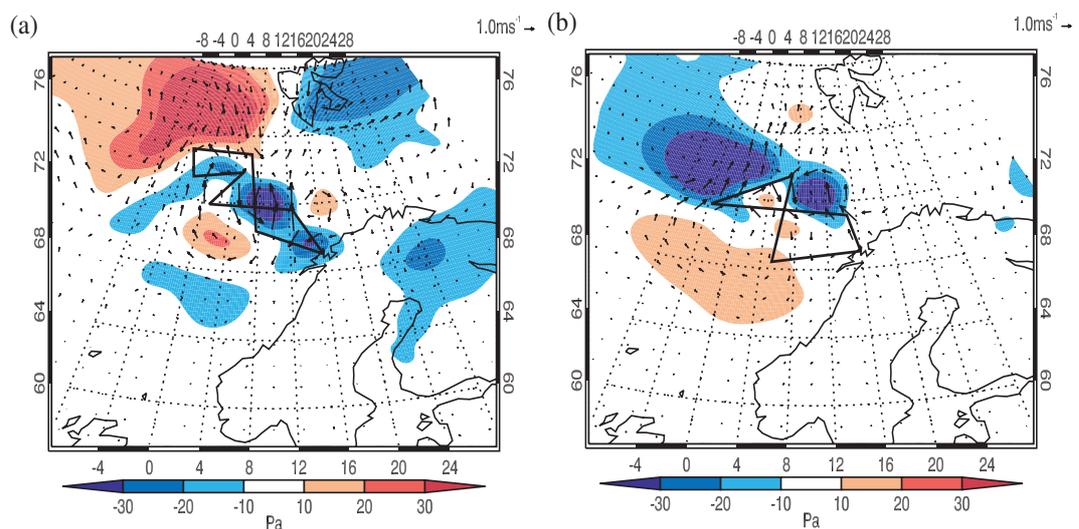

Figure 7.5: *Targeted minus CONTROL analysis increment for the wind strength and direction (arrows) and pressure (contours) for (a) TNOMEM_12 and (b) TNOMEM_18 increments. The wind differences are a pressure-weighted vertical average over all 38 model levels, and the pressure is shown at model level 9 (approximately 850hPa). The flight track is overlaid (thick black line).*

7.6 Impact of Targeted Sondes on the Forecast for the Verification Region

7.6.1 Validation of Verifying Analyses with Observational Data

To assess the impact of the targeted sondes on the forecast for the verification region over Norway, the forecasts were verified against ECMWF global 25km analyses, consistent with the forecast verification presented in Chapter 5. Global ECMWF analyses were used as they were the highest resolution analyses available that covered the area of interest. It was not possible to use MetUM analyses as the global analyses are available only on a 40km grid, which is a coarser grid than the 24km grid used in this study, and the NAE analyses do not extend sufficiently northwards to include the locations where the sondes were released.

It was discussed in Section 7.1 that the Arctic regions can be considered data sparse, particularly with respect to high-resolution profile measurements of the atmosphere. This lack of observational data available to update the background field could impact on the quality of analyses produced for this region. It was preferential to use analyses to verify the forecasts for this case due to the spatial and temporal resolution of the dataset, and the number of variables available. To ensure that the analyses were of good enough quality to use to verify the forecast for both the verification region on the larger-scale and the smaller-scale polar low, they were checked first checked against observations.

Figure 7.6 shows a comparison between the ECMWF analysed surface winds and the surface winds recorded by the targeted dropsondes during flights 2 and 3 through the polar low. When comparing the analysis and dropsonde data it is important to remember that the dropsonde data was used in the assimilation to constrain the background field. Therefore a bad match between the dropsonde and analysis data would suggest that the assimilation scheme has not been able to adjust the background towards the sonde data. It is also important to note that the observation errors have already been taken into account in the production of the analysis, whereas although the dropsonde data plotted in Figure 7.6 has been quality-controlled this does not remove errors resulting from the accuracy of the instruments used, for example. Therefore we would not expect to see an exact match between the analysed and observed winds. Figure 7.6(a) shows reasonable agreement between the analysed and observed winds. There is also good agreement in wind speed and direction between the
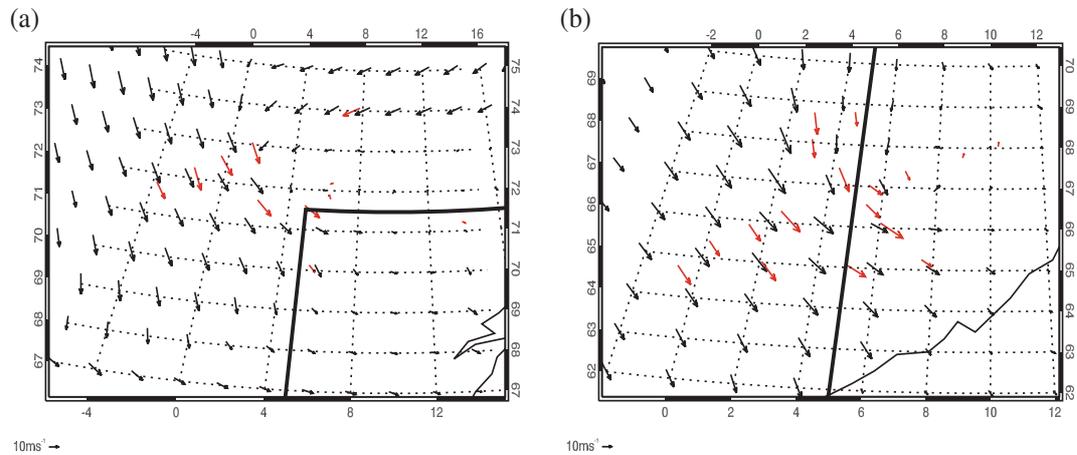
Figure 7.6: *Surface wind speed from ECMWF analyses (black arrows) and targeted dropsondes (red arrows) for (a) targeting flight 2 at 1800 UTC 03 March and (b) flight 3 at 1200 UTC 04 March. The edge of the verification region is marked by a thick black line, and the coast of Norway by a thinner black line.*

dropsonde and ECMWF analysed winds in Figure 7.6(b). The wind speed and direction are in good agreement through the depth of the atmosphere (not shown).

Figure 7.7 shows a comparison between ECMWF analysed 10m winds and the 10m winds derived from the satellite-borne QuikSCAT (Quick Scatterometer) instrument, at the time of the polar low landfall in the analyses. The general pattern of windspeed in the Norwegian Sea corresponds well in Figures 7.7(a) and (b). Figure 7.7(c) shows that differences in 10m windspeed between ECMWF and QuikSCAT are small in the region of the polar low but there are differences of over 10ms^{-1} to the north of Scotland, where a synoptic-scale system is approaching. Chelton et al. (2006) found large differences between ECMWF analyses and QuikSCAT winds for a high-wind speed case, and attributed the difference to the lack of spatial variability in the 10m winds in the ECMWF model, and smoothing out of small high wind speed features. Renfrew et al. (2009) compared QuikSCAT and ECMWF winds with low-level aircraft wind observations made during GFDex and also found that the highest wind speeds were underestimated by the ECMWF model. However, Renfrew et al. (2009) also found that compared to the aircraft observations, the QuikSCAT winds above 20ms^{-1} were generally over-estimates. This bias in the QuikSCAT winds was also found by Moore et al. (2008) when comparing QuikSCAT winds for high-winds events around Cape Farewell with buoy observations. The discrepancy between the ECMWF analyses and QuikSCAT winds

(a)

(b)

(c)
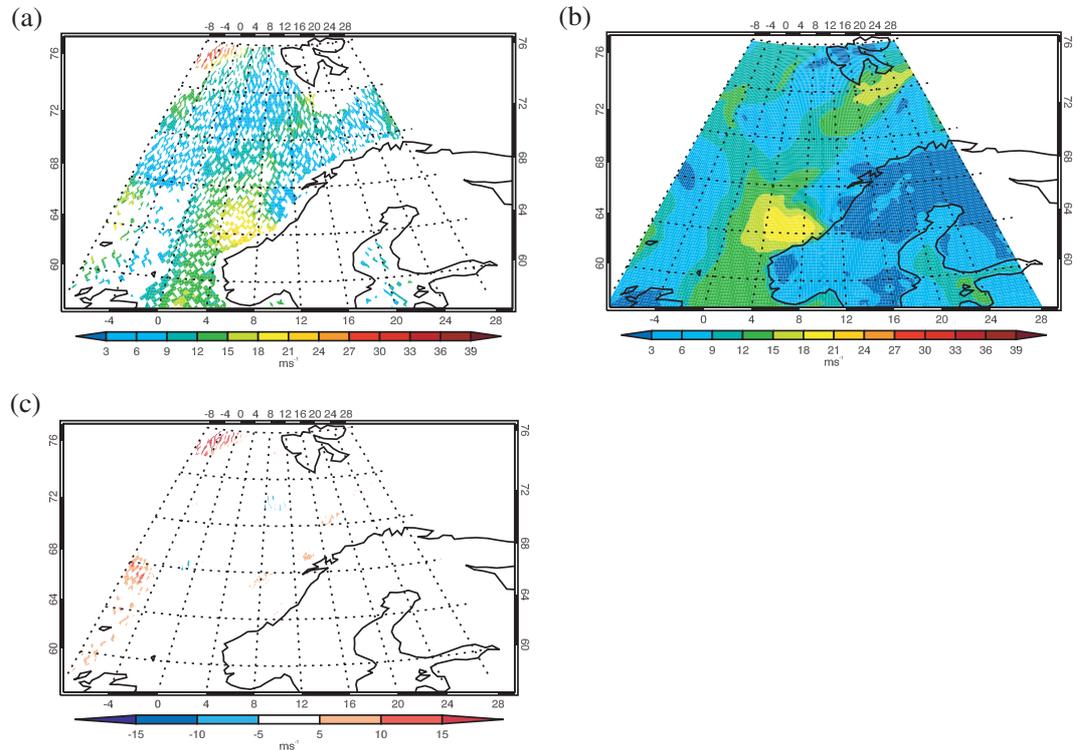
Figure 7.7: *10m wind speed corresponding to the time of the polar low landfall on 04 March, from (a) QuikSCAT descending pass at 1800 UTC, (b) ECMWF analysis at 1800 UTC and (c) QuikSCAT minus ECMWF.*

seen in Figure 7.7 is therefore likely to be a combination of model error and shortcomings in the QuikSCAT wind retrieval.

To compare the position and intensity of the polar low in the analyses and observations, the motion of the polar low was removed from the wind fields displayed in Figure 7.7(a) and (b). The direction and speed of the polar low in the six-hour period preceeding landfall was calculated using AVHRR and Moderate Resolution Imaging Spectrometer (MODIS) satellite images, and this was removed from the QuikSCAT winds to give Figure 7.8(a). The direction and speed of the polar low during the same period was calculated using the ECMWF analyses, and this removed from the ECMWF analysis at 1800 UTC 04 March, to give Figure 7.8(b). The speed of the polar low was in good agreement between the satellite images and ECMWF analyses, although the satellite track had a slightly stronger easterly and weaker southerly component to it. The location of the centre of the polar low, as defined by the centre of the circulation of the 10m system-relative winds is in good agreement between Figures 7.8(a)

and (b). The low-level circulation is of greater intensity in the ECMWF analysis than the QuikSCAT winds, as shown by the westerly winds to the north of the polar low centre in Figure 7.8(c). Overall, the ECMWF analyses compare well with the observations of the polar low, and therefore are suitable to use to verify the forecast of the polar low.
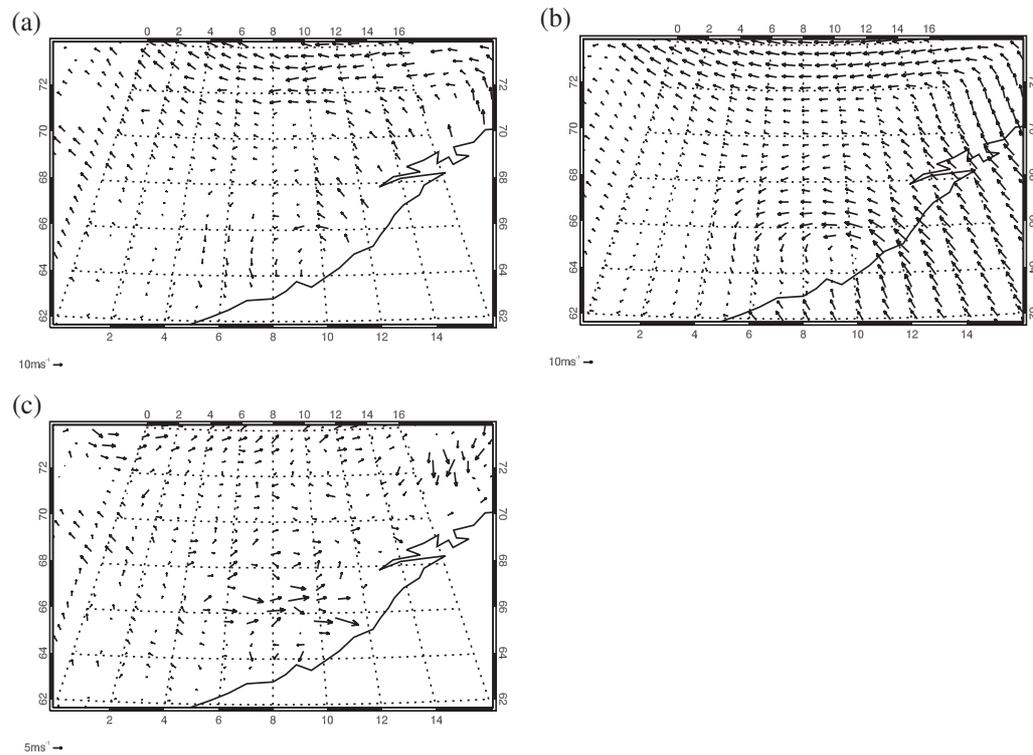
Figure 7.8: *System-relative 10m winds at the time of the polar low landfall from (a) QuikSCAT, (b) ECWMF analysis and (c) QuikSCAT minus ECMWF. The times are the same as for Figure 7.7. To obtain the system-relative winds, the system velocity was calculated from AVHRR and MODIS for (a) and from the ECMWF analyses for (b), and removed from the wind fields shown in Figure 7.7(a) and (b).*

7.6.2   Impact of Targeted Sondes on the Forecast

Consistent with the analysis of the GFDex targeted sondes in chapters 4 and 5, the impact of the targeted sondes was assessed by calculating the impact, I (4.3) in terms of total energy (using the same vertical levels as in Chapter 4 of 850, 500 and 250hPa). Figure 7.9 shows I for the TNOMEM_18 and TMEM_18 forecasts started from 1800 UTC on 03 March. The initial impact for the TMEM_18 run (Figure 7.9(b)) is more spread out than for the TNOMEM_18 run (Figure 7.9(a)), as it has memory of the 1200 UTC targeted sondes from the previous

forecast run 6 hours earlier. In both the TNOMEM_18 and TMEM_18 runs the initial impact is located around the centre of the polar low, with an increase in total energy to the east of the polar low centre. This is consistent with the eastward shift in the system seen in the analysis increments. Both forecasts show impact propagating with the developing polar low from the target region into the verification region; however, the TNOMEM_18 forecast shows greater improvement associated with the centre of the polar low, with only small areas of forecast degradation. The TMEM_18 forecast also shows forecast improvement associated with the synoptic-scale system approaching from the west. Note that the magnitude of I for this polar low case is larger than for equivalent forecast times for the GFDex cases.

A large contribution to I at the time the polar low makes landfall comes from the kinetic energy of the forecast difference at the 850hPa level (not shown). To better understand the positive and negative impact seen in Figure 7.9(e) as a result of this, Figure 7.10 shows the strength of the 10m winds at the time of the polar low landfall. The 10m winds are shown instead of the 850hPa winds as these show a better correspondence with the mean sea level pressure. Figure 7.10(a) shows the strength of the 10m winds in the ECMWF analyses, these are on the order of 21ms^{-1} in a region to the southwest of the polar low centre, shown by the closed mslp contours. In the CONTROL forecast from 1800 UTC valid at the same time as the analysis (1800 UTC 04 March) the polar low has already made landfall further south, and therefore the region of strong winds is too far south and too small. The TMEM_18 forecast looks similar to the CONTROL forecast, but without a defined mslp centre. The TNOMEM_18 run is a better forecast of the polar low position and strength, as also seen in Figure 7.12, although the region of strong winds extends too far south. The CONTROL forecast from 1200 UTC looks similar to the TMEM_18 forecast, with no defined central pressure as the polar low has already made landfall and begun to move across Norway. The region of strong winds accompanying the polar low is displaced southwards, and is weaker than the analysed winds. There is also a secondary cyclonic feature, overdeveloped compared to the analysis, that is south-west of the parent low, and can be identified by the elongated MSLP contours and region of strong winds. The TNOMEM_12 forecast shows little improvement on the 1200 UTC CONTROL forecast in terms of the winds associated with the polar low, however it is improved in terms showing less development of the secondary cyclonic feature. The region of strong winds accompanying the polar low is stronger than in the CONTROL fore-
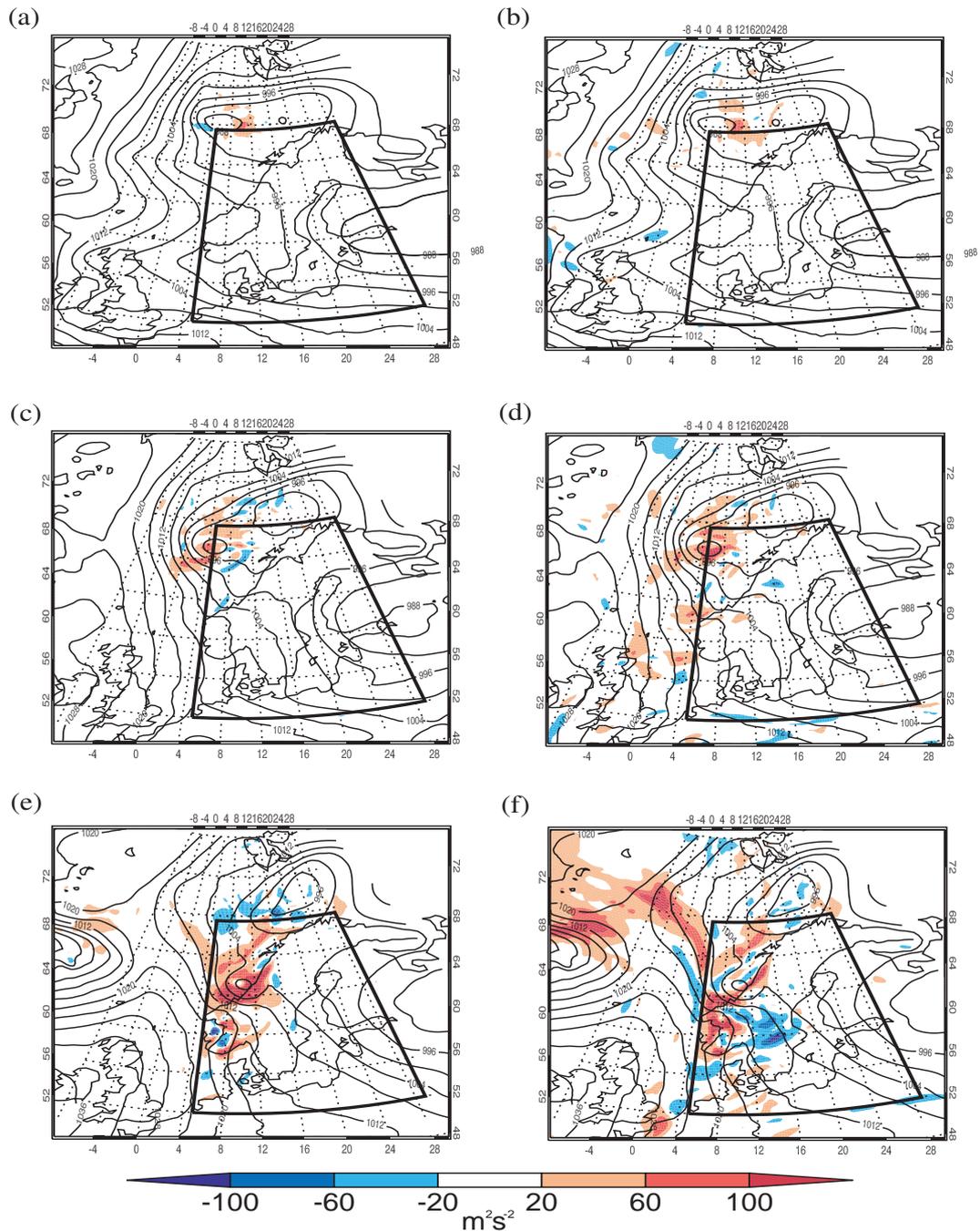
Figure 7.9: *Impact, I, calculated using (4.3) of the targeted sondes for the TNOMEM_18 (left column) and TMEM_18 (right column) 1800 UTC forecasts at (a) and (b) 0, (c) and (d) 12, and (e) and (f) 24-hours forecast time. The analysed mean sea-level pressure is contoured. The verification region over Scandinavia is outlined.*

cast, but displaced south and narrower than appears in the analysis. It is clear that the greatest impact of the targeted sondes is on the forecast of the polar low and not of the larger-scale

synoptic features within the domain.
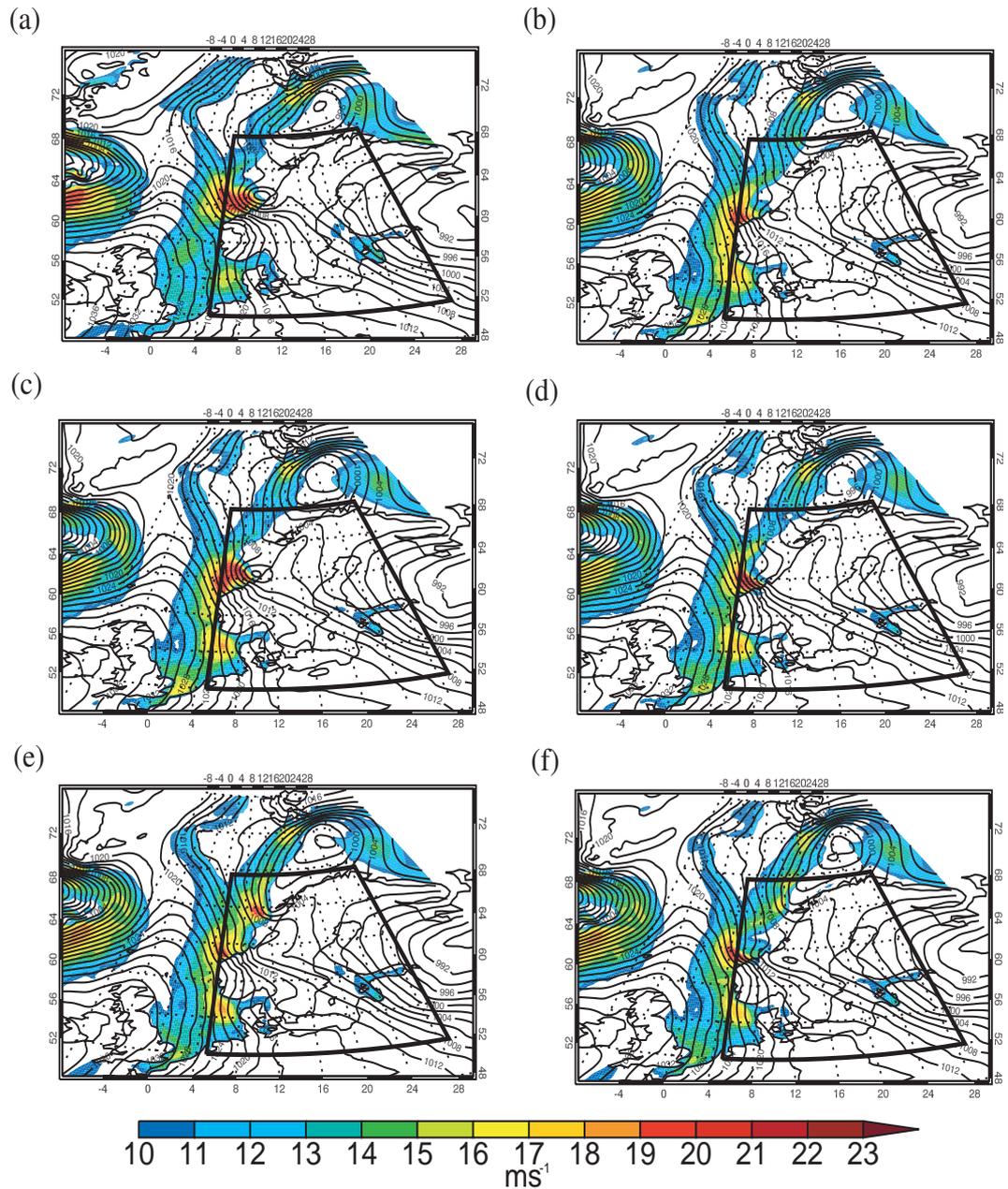
Figure 7.10: *MSLP (black contours) and 10m wind strength exceeding 10ms^{-1} (coloured contours) for 1800 UTC 04 March, from (a) ECMWF analysis, 24-hour forecasts from (b) CONTROL 1800 UTC, (c) TNOMEM_18 and (d) TMEM_18, and 30-hour forecasts from (e) CONTROL 1200 UTC and (f) TNOMEM_12.*

Figure 7.11 shows the *Relative_Impact* (5.6) of the targeted sonde data on the forecast for the verification region. The *Relative_Impact* is non-zero at the initial forecast time because

the target region is on the edge of the verification region, therefore some of the initial impact is within the verification region. For the forecast started at 1200 UTC, the TNOMEM_12 forecast, the *Relative_Impact* increases to a maximum after 18 hours, but at 30 hours forecast time, when the polar low makes landfall the impact is small and negative. The positive impact occurs not because the forecast of the polar low is improved but because the TNOMEM_12 forecast better represents the secondary cyclonic feature which was overdeveloped in the 1200 UTC CONTROL forecast. For the forecasts started at 1800 UTC, the *Relative_Impact* is above zero throughout the forecast. The *Relative_Impact* for the TNOMEM_18 forecast increases with increasing forecast time to a maximum forecast improvement of 17% relative to the CONTROL forecast error after 24 hours forecast. The *Relative_Impact* for the TMEM_18 forecast peaks 12 hours earlier than for the TNOMEM_18 forecast. After 24 hours forecast time when the polar low makes landfall, the improvement is 10% therefore the forecast is improved relative to the CONTROL, but not by as much as the TNOMEM_18 forecast, which does a better job of capturing the extent of the region of high wind speed associated with the polar low. It is also interesting to note that the forecast error (in the CONTROL forecast) after 30 hours was $100 \text{m}^2 \text{s}^{-2}$ (not shown); this is equivalent to the average forecast error over the equivalent verification region for the GFDex cases after 60 hours (Figure 4.3(d)). Thus, a reduction in forecast error of 17% is a greater achievement in this case.

7.7   Impact of Targeted Sondes on the Forecast of the Polar Low

The impact of the targeted sondes tracks with the polar low from the target to verification region, and the polar low is the most important dynamical feature that is affecting the weather in the verification region. Therefore the targeted sondes have improved the forecast for the verification region by improving the forecast of the polar low. It is important to examine the forecast of the polar low itself to see how the targeted sondes have improved the forecast of this system. The emphasis here is on the forecast of the polar low landfall.

A comparison of the central pressure of the polar low in the analyses and CONTROL forecast shows that without the targeted sondes in the forecast, the polar low is consistently too weak in the CONTROL forecast, particularly after 12 hours when strength of the polar low weakens much faster than is seen in the analyses (Figure 7.12(a)). When targeted sondes are included only at 1800 UTC, the initial impact of the targeted sondes is to strengthen the
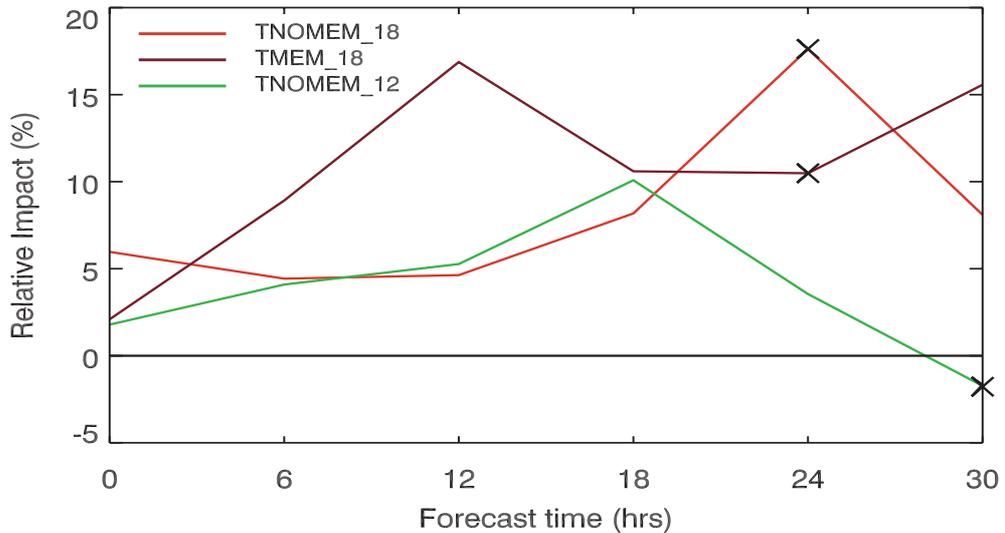
Figure 7.11: *Relative_Impact, calculated using (5.6) for the TNOMEM_18 (red line) and TMEM_18 (brown) and TNOMEM_12 (green) forecasts, averaged over the Scandinavian verification region. The time that the polar low makes landfall is marked by crosses; the polar low makes landfall 30 hours into the TNOMEM_12 forecast, and 24 hours into the TNOMEM_18 and TMEM_18 forecasts.*

polar low, and the polar low remains deeper than the CONTROL forecast, but shallower than the analysis (Figure 7.12(b)). When targeted sondes are included at 1200 UTC and 1800 UTC and the forecast run from 1800 UTC (the TMEM_18 run) the initial strength of the polar low is similar to the TNOMEM_18 forecast, but weakens too quickly and does not improve the intensity of the system after 12 hours forecast. There is no closed contour present in the 24-hour forecast for the TMEM_18 run, so the final estimate of central pressure and position are not accurate. The targeted sondes from 1200 UTC degrade the forecast of the intensity of the polar low, as measured by the minimum mslp, as the central pressure is weaker than the CONTROL forecast throughout the forecast (Figure 7.12(a)). Both the CONTROL and TNOMEM_12 forecasts have a central pressure weaker than the analysed central pressure, particularly for the last 12 hours of the forecast.

As polar lows are relatively small features compared to mid-latitude weather systems, with a scale of only a few hundred kilometres, it is important to have an accurate forecast of the track. A track error of only 100km could mean that a warning of severe weather is given to the wrong place, and a region further up the coast is unexpectedly hit by severe weather. The track of the polar low was calculated using the position of the central pressure at 6-hourly intervals. The track produced by the ECMWF analyses (Figure 7.13) were compared with a
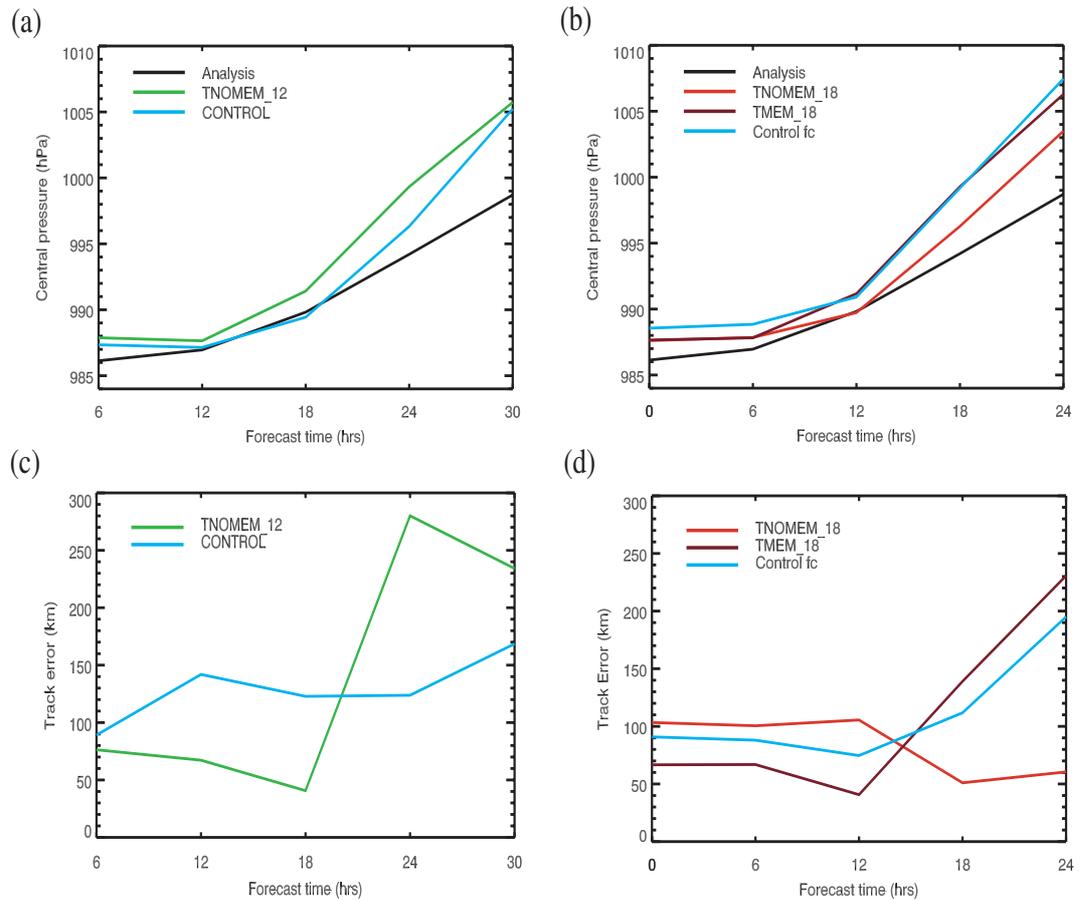
Figure 7.12: *The central pressure of the polar low (top panels) as measured by the ECMWF analyses (black line), and CONTROL forecast (blue), and (a) TNOMEM_12 (green) and (b) TNOMEM_18 (red) and TMEM_18 (brown) forecasts. The error in the track of the polar low (bottom panels), calculated as the difference in position of the centre of the polar low between the forecast and analysis at each forecast time. The x-axis has been adjusted so that the same period of time is displayed for both forecasts, with the polar low making landfall at the last forecast time shown in the figure.*

track calculated by eye from MODIS and AVHRR satellite images (not shown). Due to the small scale of the polar low, and the difficulty in precisely locating the centre of the polar low from the satellite images and assigning accurate latitude and longitude values to its position, this track has a large error and is not shown. The position of the polar low at each analysis time corresponded well to the satellite track, given the limitations of producing the satellite track. Therefore we can have some confidence in the accuracy of the ECMWF analyses with regards to the track of the polar low.

The tracks of the polar low in both the forecasts and ECMWF analyses are shown in
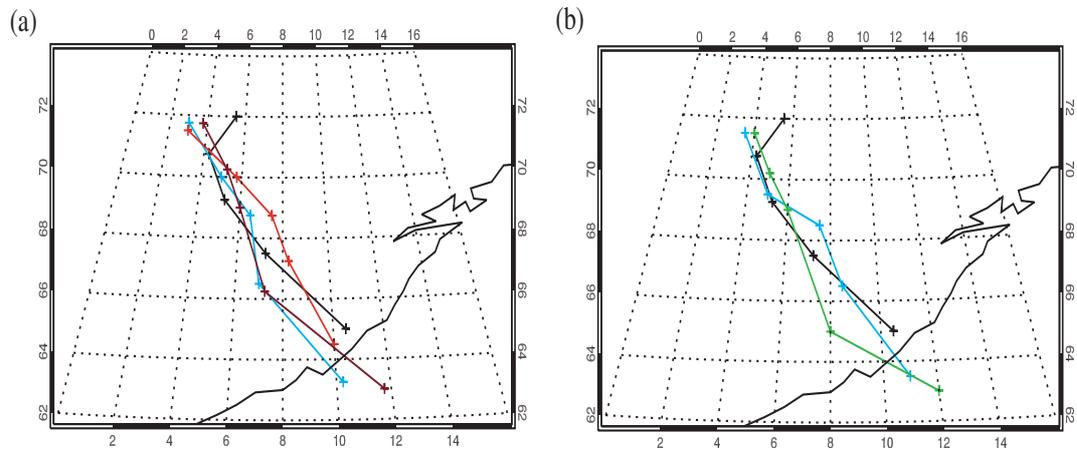
**Figure 7.13:** *Position of polar low plotted every 6 hours for ECMWF analysis (black line) and (a) CON-TROL 1800 UTC forecast (blue line), TNOMEM_18 (red line) and TMEM_18 (purple line) forecasts and (b) CONTROL 1200 UTC forecast (blue line) and TNOMEM_12 forecast (green line).*

Figure 7.13. For forecasts starting at 1200 UTC (Figure 7.13(b)) and at 1800 UTC (Figure 7.13(a)) the initial position of the polar low in the forecasts is to the east of the position in the analyses. This could be due to the fact that the polar low is only just forming, and is situated within a large area of low pressure. Therefore it is difficult to accurately identify the centre of the polar low at this time, using either mslp or 850hPa relative vorticity. To facilitate the comparison of the position of the polar low at each forecast time with the analysed position, the track error of the forecasts was computed from the tracks in Figure 7.13. The track error of the forecasts of the polar low are shown in Figure 7.12(c) and (d), where the error is defined as the great circle distance between the position of the centre of the polar low (defined by the central pressure) in the analysis and forecast. For the first 12 hours of the forecast the track error of the TNOMEM_18 and TMEM_18 forecasts are similar to the CONTROL forecast, and the error remains similar. After 12 hours forecast time, the track error in the TMEM_18 forecast increases rapidly, coinciding with the period where the strength of the polar low decreases. The error in the TNOMEM_18 forecast decreases after 12 hours forecast. After 24 hour forecast, as the polar low makes landfall, the track error in the TNOMEM_18 forecast has been reduced by 60% to 60km, compared with 195km in the CONTROL forecast (Figure 7.12(d)). The track error in the TNOMEM_12 forecast is smaller than the CONTROL forecast for the first 18 hours of the forecast, after which it becomes larger; therefore, the targeted sondes at 1200 UTC do not improve the position of the polar low landfall relative to

a CONTROL forecast (Figure 7.12(c)).
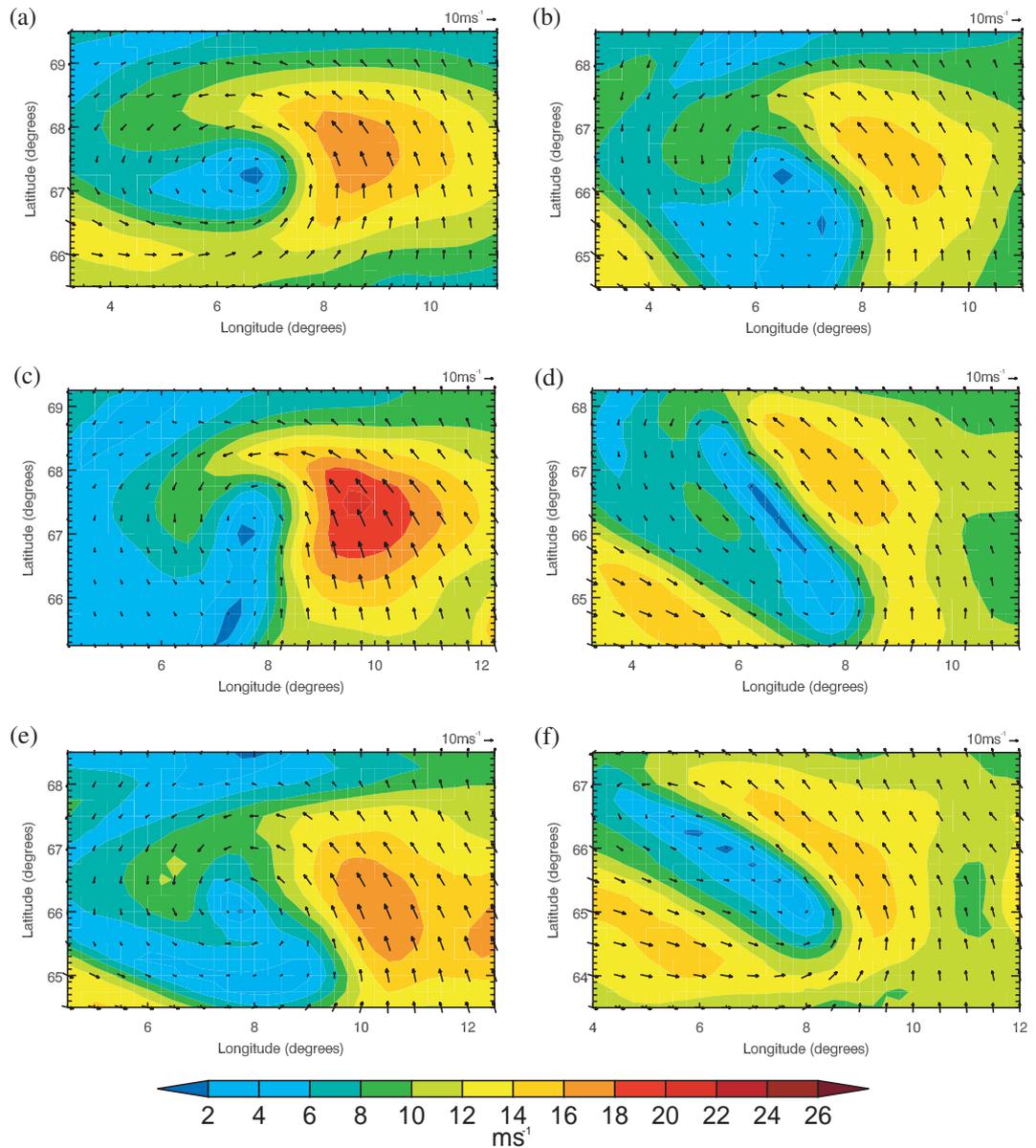
Figure 7.14: *System-relative wind strength (coloured contours) and direction (arrows) at 10m for a box of radius two degrees latitude and four degrees longitude, where the centre of each box is the centre of the polar low, as defined by the minimum MSLP. For (a) ECMWF analysis, (b) CONTROL 1800 UTC forecast, (c) TNOMEM_18 forecast, (d) TMEM_18 forecast, (e) CONTROL 1200 UTC forecast and (f) TNOMEM_12 forecast at 1200 UTC 04 March, six hours before the polar low makes landfall.*

Figure 7.10 shows both the position and intensity of the polar low at the time of landfall in the analyses. The positions of the polar low in the CONTROL and targeted forecasts are

different, and both are also different to the analyses. Thus if the mslp or wind field were to be differenced, the difference field would tell you only that the polar low is displaced in the targeted forecast, relative to the CONTROL forecast and analyses, not if there is any difference in the strength in the targeted forecast. To compare the intensity of the polar low in the targeted and CONTROL forecasts, the centre of the polar low was found and data within 2 degrees latitude and 4 degrees longitude of the centre extracted. To look at the intensity of the polar low, the direction and speed of travel in the 12-hour period centred on 1200 UTC 04 March was computed from the ECMWF analyses and removed from the wind field to leave the system-relative winds. This analysis was performed for the analysed fields and forecast fields corresponding to 1200 UTC on 04 March, 6 hours before landfall (subtracting the ECMWF analysed system velocity from all forecast fields). This period is used instead of landfall, due to the interaction of the polar low with the orography as the polar low makes landfall on the Norwegian coast. Figure 7.14 shows the system-relative winds of the polar low. The analysed system-relative winds (Figure 7.14(a)) show a clear cyclonic circulation, with calm conditions in the centre and maximum wind strength to the east of the polar low centre. Recall that the ECMWF analyses assimilated the dropsonde data from the third flight through the polar low, and that in Section 7.6.1 it was shown that the analyses and dropsonde data were in good agreement. Therefore the ECMWF analyses should well-represent the actual windspeeds in the polar low at this time. Both the 1800 UTC CONTROL and TNOMEM_18 forecasts (Figure 7.14(b) and (c)) show a calm centre, and strongest winds to the east of the centre, however this region of strong winds is too weak in the CONTROL forecast and too strong in the targeted forecast. The calm centre is also not as well-defined. Neither the TMEM_18 or TNOMEM_12 forecasts (Figure 7.14(d) and (f)) improve on their respective CONTROL forecasts; The TMEM_18 and TNOMEM_12 forecasts look similar, with an elongated circulation, and weaker winds around the calm centre.

To assess whether the forecast improvement or degradation illustrated so far for the surface fields corresponds to an improvement or degradation of the vertical structure of the polar low, cross-sections were taken through the centre of the polar low. Figure 7.15(a) and (b) shows cross-sections through the polar low along the flight track, from points A to B marked on Figure 7.4(e). For figure panels (c) to (g) the start and end latitudes (but not longitudes) were adjusted to ensure that the cross-section passed through the centre of the polar low for all
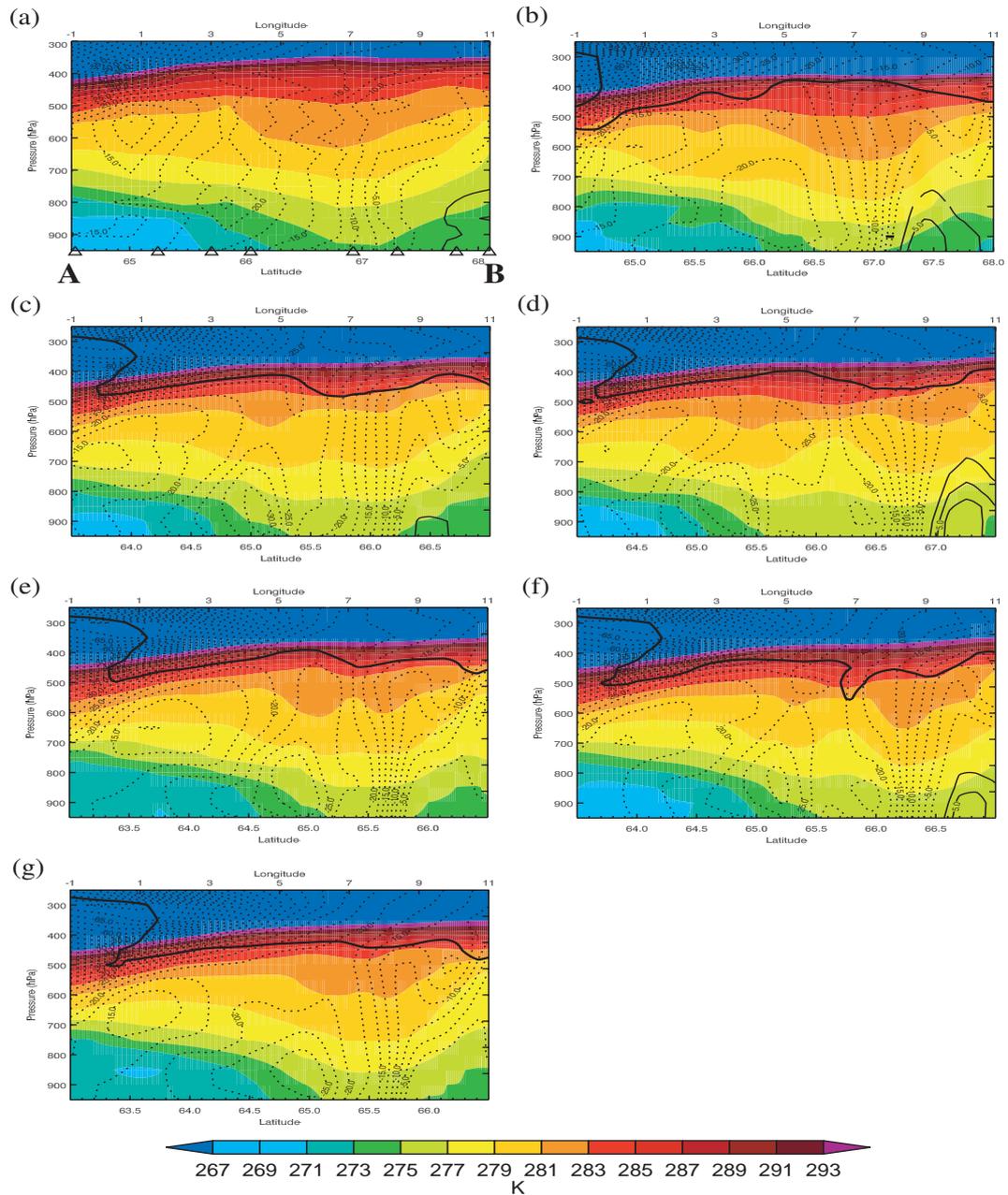
Figure 7.15: *Vertical cross-section from 1°W to 11°E shifted north or south so that each section passes through the polar low centre at 1200 UTC 04 March, six hours before the polar low makes landfall, for (a) dropsonde data, (b) ECMWF analysis, (c) CONTROL forecast from 1800 UTC, (d) TNOMEM_18 forecast, (e) TMEM_18 forecast, (f) CONTROL forecast from 1200 UTC and (g) TNOMEM_12 forecast. Profiles of theta (coloured contours), v-wind component (dotted contours) and PV=2 contour (thick contour). The location of points A and B are shown in Figure 7.4(e). The triangles in (a) mark the dropsonde locations.*

forecasts, so that the resulting vertical structures could be easily compared with the observed structure and analysed fields. Note that as the start and end longitudes were not adjusted, the centre of the polar low is not in the same position on each cross-section. The vertical structure obtained from the dropsonde data (Figure 7.15(a)) shows a good agreement with the ECMWF analyses (Figure 7.15(b)) in the position of the polar low, and of the main features. The observed and analysed *v*-wind field shows the cold-air outbreak to the west of the polar low centre, which is strongest below 700hPa, as is expected for a reverse-shear system. There is a strong gradient in meridional wind speed from strong northerly winds to the west of the polar low to weak southerly winds to its east. This gradient is well-defined up to the tropopause, with almost no tilt in the vertical. There is also an upper-level jet at 300mb with northerly winds reaching $60\mathrm{ms}^{-1}$, to the west of the polar low.

The 1800 UTC CONTROL forecast has northerly winds $10\mathrm{ms}^{-1}$ stronger than observed or analysed to the west of the polar low. The upper-level jet in the west of the cross-sections is associated with a synoptic-scale system that is moving eastwards towards the Norwegian coast. This jet is further to the east in the 1800 UTC CONTROL forecast than in the analysis, reducing the westward extent of the cold-air outbreak. The TMEM_18 forecast also has the synoptic-scale system advanced too far east, although the magnitude of the surface northerly winds is reduced. The TNOMEM_18 has the synoptic-scale system positioned further west than the CONTROL forecast, and the magnitude of the low-level northerly winds is of the same magnitude as the TMEM_18 forecast. The TMEM_18 and CONTROL forecasts lack the region of southerly winds to the east of the polar low centre, that is captured by the TNOMEM_18 forecast. The TNOMEM_12 forecast looks similar to the TMEM_18 forecast, and does not contain a PV anomaly, but does lack the region of southerly winds east of the polar low centre.

7.8 Conclusions

In this targeting experiment a developing polar low in the Norwegian sea was targeted with a dense network of sondes. The polar low was located within a larger ETKF-predicted sensitive region which covered the north-Norwegian sea. The targeting methodology employed here differed from conventional targeting in that only the dynamical feature within the sensitive region, i.e. the polar low, was targeted, which left a large fraction of the sensitive

region not covered by targeted observations. The polar low was targeted twice: as it was forming, and once a closed-contour was visible on the mslp chart.

The targeted observations from the first flight at 1200 UTC on 03 March were not assimilated operationally as the data cut-off time was missed (therefore the operational analyses used to verify the forecasts also do not contain these sondes). Hindcasts that assimilated this targeted sonde data showed that the inclusion of these sondes caused a small forecast degradation within the verification region, in terms of total energy, at the time of polar low landfall. The forecast for the verification region was initially improved in terms of total energy, as the targeted forecast reduced the over-development of a secondary cyclonic feature seen in the CONTROL forecast that was positioned to the north of the polar low, within the verification region. The forecast of the polar low itself was not improved in terms of track or intensity, as the system was too shallow throughout its lifetime and made landfall too early. The system-relative winds showed that compared to the analysis, the system cyclonic circulation was elongated rather than being almost circular.

The targeted observations from the second flight at 1800 UTC on the same day were assimilated operationally into the 1800 UTC forecasts. When the targeted observations were assimilated from both flights, and the forecast run from 1800 UTC (the TMEM_18 forecast), the forecast for the verification region was improved in terms of total energy, by 10% as the polar low made landfall. This improvement occurred because the strength of the winds in the south-west quadrant of the polar low were strengthened relative to the CONTROL forecast, thus some improvement was seen in terms of total energy. However the forecast of the polar low landfall was not improved relative to the CONTROL forecast, as the system made landfall too early and therefore lost its intensity as it moved over land. From the cross-sections through the centre of the polar low it was clear that the synoptic-scale system was advanced too far east in the forecast 6-hours before landfall, and that this was responsible for confining the westward extent of the region of strong surface northerly winds.

Hindcasts for targeted observations assimilated from the second flight only (TNOMEM_18) showed that these sondes caused an improvement to the verification region in terms of total energy of 17% after 24-hours (as the polar low made landfall). This improvement was due to improving the intensity of the system, particularly with respect to the position and strength of the winds in the south-west quadrant of the polar low. The forecast of

the polar low landfall was improved in terms of position and intensity over the control fore-cast, and the structure of the polar low was also improved, although the 10m system-relative winds showed that the circulation was too strong to the east of the polar low with respect to the analysed circulation.

Whether or not a forecast is considered to have been improved by the targeted observations depends on the metric and scale used to assess improvement. Here, looking at the large-scale, the forecast can be considered improved in both the TNOMEM_18 and TMEM_18 runs, but not in the TNOMEM_12 run at the time the polar low made landfall. On the scale of the polar low, only the TNOMEM_18 forecast can be considered a real improvement over its CONTROL forecast, although in terms of the strong winds in the south-west quadrant the TMEM_18 forecast can also be considered an improvement over the CONTROL forecast. It is clear that in this case the memory of previous targeted observations results in a poorer fore-cast, however the forecast appears to be dynamically very sensitive to the initial conditions and therefore it is difficult to draw strong conclusions about the effect of memory on this case.

The mechanism for the improvement or degradation of the forecast by the targeted sondes for this case is the modification of the initial state by the targeted sonde data. As the targeted sondes directly sampled the main dynamical feature in the sensitive region, it is not surprising that the impact moved and amplified with this dynamical feature. The addition of the targeted sondes caused analysis increments in the region where the polar low was developing, and therefore modified both the position and intensity of the developing polar low and the region of strong winds associated with the cold-air outbreak. The forecast of this event is obviously highly sensitive to the initial conditions; a small change caused by the assimilation of targeted observations has led to a deviation from the CONTROL forecasts, that has been shown to be randomly positive or negative when observations were assimilated at different times. This has ramifications for data assimilation for an ensemble of forecasts where the aim is to look at the change in a plume of tracks, resulting from the assimilation of targeted sonde data. The fact that there was no clear modification to the baroclinicity of the system, or the PV distribution which modified the way that the polar low developed does not mean that the targeted sonde data did not contain new information on the baroclinicity of the system, for example. Rather it is likely that the 3D-Var system used to assimilate the data was not able to make full use of the information contained within the observations to constrain the analysis. A 4D-Var system

would have been more suited to this case, for the reasons outlined in Section 7.3.1, but it was unfortunately not possible to run this case with a 4D-Var assimilation system, like that used to run the GFDex targeted cases. It is speculated that if this case were to be re-run using a 4D-Var system, it should enable the data assimilation scheme to make better use of the sonde data, as evolving forecast error covariances during the data assimilation window should lead to shorter correlation structures in the vicinity of the polar low, which should improve how the sonde data is used to modify the forecasts.

CHAPTER 8

Conclusions

This thesis has analysed the impact of targeted observations on the forecast for northern Europe, with a focus on addressing the questions set out in Section 1.2:

- What is the impact of targeted observations on the forecast for Northern Europe?

- How do targeted observations impact the forecast?

- Do targeted observations have the potential to have a large impact on the forecast?

The main conclusions to this thesis, which address the above questions, are summarised in Section 8.1. The known caveats to these conclusions are presented in Section 8.2. There is scope to expand the work done in this thesis; potential lines of enquiry arising from results presented in this thesis are discussed in Section 8.3.

8.1   Main Conclusions

The first section of this thesis analysed the impact of observations targeted in the seas around southern Greenland and Iceland on the forecasts for Northern Europe, specifically Scandinavia and the UK. The observations were made during the GFDex (Renfrew et al., 2008) and assimilated into the operational Met Office forecasts.

In Chapter 4 the impact of the GFDex targeted observations on the 1-3 day forecasts for Scandinavia and the UK was assessed. Three flights were targeted in TESV-predicted sensitive regions and one in an ETKF-predicted sensitive region, and there was a further null case. The impact of the targeted observations on the forecast was assessed using a limited-area model on a 24km grid; previous studies used global models with grid spacings of 60km or greater (Langland, 2005). The observations were assimilated using 4D-Var with a six-hour observation window. For two cases it was shown that the targeted observations contained little 'new' information and that this resulted in little to no forecast impact. For cases where the targeted observations did provide new information not contained in the background field, the

data assimilation scheme was able to use this information to reduce the initial condition error. This resulted in small forecast impact (measured in terms of total energy), with a maximum improvement of 5% after 24 hours forecast time and maximum degradation of 3% after 48 hours forecast time, verified against MetUM analyses. The size of the impact is consistent with recent targeting studies such as ATReC (Petersen and Thorpe, 2007) and implies that targeting observations adds little value to the forecast.

In Chapter 4 how targeted observations impact the forecast was also investigated, for the two cases where the targeted observations had a positive impact on the forecast. For one case, the mechanism for impact was the modification of the position of a tropopause fold, which was forcing the development of a polar low. By modifying the position of the tropopause fold, the associated wind field was modified, leading to an improved forecast as the polar low moved through the Scandinavian verification region. In the second case the forecast impact was related to the downstream development of an upper-level trough, therefore in both cases modifying the upper-level fields was crucial in obtaining forecast improvement. A small ensemble of forecasts was used to test the robustness of the forecast impact seen in these two cases. This showed that the forecast improvement was robust in the first case but not the second. Where the observations were cycled in the background field this led to larger impact through the introduction of random perturbations in the initial conditions, some of which grew, and improved or degraded the forecast. This result implies that targeting studies that cycle observations in the background field when running hindcast studies do not measure the true impact of targeted observations.

The potential to increase the forecast impact was assessed in Chapter 5 for one GFDex case (that for which the small improvement was shown to be robust), by increasing the weighting given to the dropsonde observations during the assimilation process. Observation errors were first calculated using the GFDex dropsonde data. Compared with the operational observation errors the calculated GFDex errors were smaller by a factor of two for u, v and T (although of the same order of magnitude for relative humidity) which motivates the reduction of the errors in this experiment. Reducing the dropsonde observation errors gave the dropsonde observations greater weighting during the assimilation and this increased the impact of the dropsondes on the forecast; there was an increase in maximum forecast improvement from 7% to 10% when targeted dropsondes were assimilated, and from 15% to

17% when all dropsondes were assimilated. This corroborates Szunyogh et al. (1999) who demonstrated that a closer initial fit to the observational data increased the impact of the observations on the forecast. It was also seen that the impact on the analysis increment was to locally magnify the impact of the targeted sonde data without greatly changing the pattern of impact, either locally or remotely. It was shown that two dropsondes released close to the steep orography of Greenland had degraded the forecast. When the original targeted set of dropsondes was modified, replacing the two sondes on the Greenland coast with two in the Denmark Strait, the maximum improvement rose from 5% to 17%, and there was (averaged over the verification region) no forecast degradation. The degradation of the forecast was caused by the anomalous spreading of observational data up the steep slopes. A potential solution was explored by removing the lower part of the observation profile for the two sondes on the Greenland coast before the data were assimilated. This did increase the impact of the set of targeted observations but is not an ideal solution; the low-level data should be useful if it were spread in a more realistic manner, into the Denmark Strait where the sonde data is more highly correlated with the atmospheric conditions. This issue could therefore be solved by reformulating the background error covariances so that data are not spread up steep slopes, alternatively sonde data from below the orography height could be rejected and not assimilated. The magnitude of forecast improvement resulting from reducing the observation error covariances and removing the Greenland sondes brings the impact from this case in line with previous targeting experiments over the Atlantic, such as FASTEX (Montani et al., 1999; Szunyogh et al., 1999), and is larger than individual forecast improvements seen in ATReC (Petersen and Thorpe, 2007).

The potential of targeted observations to improve the forecast was also explored in Chapter 6. A hypothesis was posed that little impact was seen from the GFDex targeting experiments because of the inability of the data assimilation scheme to extract important dynamical structure from the observations and use it to remove initial condition error. This was tested using experiments with varying numbers of pseudo-targeted observations containing a singular vector structure that was not present in the background field or routine observations, thus assuming an initial condition error of baroclinic structure. The results clearly disproved the hypothesis, as the initial condition error was reduced. A tilted increment was produced which had a similar structure to the SV contained in the pseudo-observations, although with

a reduced tilt. The amount by which the initial condition error was reduced was dependent on how well the SV structure was represented by the targeted sondes. To fully capture the SV structure, 55 sondes over an area of $1{\times}10^6 km^2$ were needed; when only 8 sondes were assimilated along a GFDex targeting flight track covering a smaller area there was less reduction in initial condition error. The implications of this result for the GFDex targeting experiments are that if the initial condition error had a baroclinic structure then the number of sondes released would have been insufficient to fully represent the structure and this would have resulted in the data assimilation scheme being unable to produce a sufficiently tilted increment. That the 4D-Var scheme could reduce this initial condition error at all is maybe surprising given that 4D-Var has been shown to struggle to produce a tilted increment with only a 6 hour data assimilation window (Thépaut et al., 1996). The results from this experiment support the results of Johnson et al. (2006) who demonstrated that the same 4D-Var scheme used in these experiments was capable of correcting growing structures. The initial energy contained in the forecast perturbations was smaller than the SV perturbations for all experiments, and the energy decreased as the SV was less well-described by the pseudo-observations (i.e. as fewer pseudo-observations were assimilated). This gap in energy is due to the data assimilation scheme not adjusting the background fully towards the observations, as it has to take into account the errors in the observations. This was shown by reducing the pseudo-observation errors, which resulted in greater adjustment towards the observations and reduced this initial energy gap, but resulted in the addition of spurious decaying structures to the analysis increment.

Chapter 7 also looked at the impact of targeted observations on the forecast for Northern Europe, this time from observations targeted in the Norwegian sea during the Norwegian IPY-THORPEX field campaign. The observations were targeted within an ETKF sensitive region, but covered only the polar low within this region with a dense network of sondes. This is a different approach to conventional targeting, where a large area is covered with sondes spaced on the order of the horizontal correlation length-scales assumed by the data assimilation scheme. A major difference between the analysis of the Norwegian IPY-THORPEX targeting experiments and GFDex targeting experiments was the use of a 3D-Var data assimilation system. The polar low was targeted twice with dropsondes, at 1200 and 1800 UTC on the same day, as the polar low formed. Only the targeted observations from 1800 UTC improved the fore-

cast of the polar low landfall, in terms of the position and intensity of the polar low (relative to a forecast which did not contain the targeted observations). The forecast improvement as the polar low made landfall was 17% in terms of total energy, which is comparable to the 01 March GFDex case (with the Greenland sondes removed). When targeted observations from 1200 and then 1800 UTC were assimilated and the forecast run from 1800 UTC, the larger-scale forecast was also improved at the time the polar low made landfall, leading to a 10% forecast improvement, which is smaller than when observations were assimilated only at 1800 UTC, but larger than the GFDex forecast impact. The targeted observations from the first flight at 1200 UTC did not improve the forecast of the polar low itself, although there was some improvement to the larger-scale synoptic circulation. The larger forecast impacts from the Norwegian IPY-THORPEX experiments (averaged over the verification region) are comparable in size to the forecast impacts from the WSR and FASTEX campaigns who recorded forecast impacts of 10-20% and also used 3D-Var to assimilate the observations (Montani et al., 1999; Szunyogh et al., 1999, 2000, 2002).

Unlike the GFDex targeted experiments analysed in Chapter 4 where there was a clear dynamical mechanism for the forecast impact, the mechanism for the forecast impact for the polar low case was solely the modification of the initial state (in terms of the position and intensity of the polar low) by the targeted dropsonde data. The forecast of the polar low was therefore highly sensitive to the initial conditions, and therefore the different targeting runs showed forecast tracks going in different directions, some closer to the analysed track and some not, without an obvious signature at initial time. This has implications for the analysis of an ensemble plume of tracks, looking at the change in the tracks resulting from assimilating extra observations.

The GFDex and Norwegian IPY-THORPEX targeting experiments were performed with memory of previous targeted observations and without memory of previous targeted observations. Some discussion of the different results is merited here. The runs that included memory are more relevant to operational targeting, as analysis systems are a continuous succession of data assimilation cycles that therefore run with memory. Of course, if targeted observations are made infrequently (i.e. less than once a week) then the TMEM and TNOMEM approaches are equivalent, as there is no memory left in the system of the previous observations. In GFDex the TMEM run included targeted observations spaced at 2-9 day intervals over a

period of three weeks, and showed large differences between the TNOMEM and TMEM forecasts. The Norway targeting experiments, where targeted observations were assimilated into successive analysis cycles also showed large differences between the TMEM and TNOMEM forecasts. The results from the GFDex TMEM targeting experiments are most relevant to sub-daily targeting, however the Norway results are relevant to using additional satellite or AMDAR data as targeted observations, where the targeted observations could be placed in successive analysis cycles. Therefore unless the aim of a targeting study is to isolate the impact of a particular set of observations on the forecast quality, then a run without memory would be more appropriate. If the aim is to mimic an operational setup, then targeting studies should use memory, particularly where observations are put into successive analysis cycles.

The main conclusions from this thesis, that answer the questions set out in Section 1.2 can be summed-up as follows.

- For the GFDex cases the impact of targeted observations on the forecast for northern Europe is on average small. Targeted observations which did not add 'new' information to the background field had no impact. Targeted observations which did contain 'new' information showed a postive impact of between 5-17% after 18-24 hours, with a maximum degradation 5% after 48 hours.

- Targeted observations from GFDex impacted the forecast through modifications to the upper-level fields. For one case the impact was associated with the modification of the position of an upper-level PV anomaly, for another the impact was associated with an upper-level trough. For the Norwegian IPY-THORPEX experiments the impact was caused soley by the modification of the initial state, changing the location and strength of the incipient polar low.

- This thesis has demonstrated that, despite improvements to the resolution and formulation of NWP models and assimilation schemes, targeted observations can have a large positive impact on the forecasts. The size of the impact has been shown to be dependent on a number of factors. Firstly that the targeted observations are given sufficient weighting in the data assimilation scheme to allow the background field to be adjusted to correct initial condition error. Secondly, approximately 50 observations covering an area of $1 \times 10^6 \text{km}^2$ are required to sufficiently describe any baroclinic structures that might

not be well-represented by the background or routine observations. Lastly, that the targeted observations are not positioned within a correlation length-scale (e.g. 200km) of steep orography, which could lead to an anomalous spreading of analysis increments along terrain-following model levels.

8.2 Caveats

A major drawback of assessing the impact of targeted observations using field campaigns is the small number of targeting cases to analyse. As discussed by Kelly et al. (2007), this makes it hard to draw statistically significant conclusions about the success of targeting, as to do so would require an extremely large number of cases. The field experiment closest to meeting this requirement is the WSR, which has targeted landfalling cyclones in the west Pacific every year since 1999. The conclusions presented in this chapter have been drawn from the analysis of a maximum of 6 targeting cases (this number includes one null case) which means that the results are not necessarily representative of all regions and seasons. At attempt to address statistical significance was made in Chapter 4 by using ensemble methods to test the robustness of the forecast impact; this method may be useful to obtain significant conclusions from a small number of cases.

Both the GFDex and Norwegian IPY-THORPEX experiments were run using limited-area models with a 24km grid-spacing. Whilst this is a higher resolution than previously used to run targeting experiments (e.g. ATReC used 60km (Petersen and Thorpe, 2007)) it is at lower resolution than the operational forecast models over the same domains. The Met Office run the MetUM at 12km resolution over the NAE domain, and have plans to increase the resolution to 4km over the same domain. Over a similar area as the limited-area domain used in the IPY-THORPEX experiments, the Norwegian Meteorological Office run their versions of the High Resolution Limited-Area Model (HIRLAM) with 12km grid-spacing, and a 4km grid-spacing is used for a smaller domain. This means that the results presented here could quickly become out-of-date as the operational forecast centres move to higher and higher resolution. Data assimilation at these higher resolutions requires different assumptions, for instance balance assumptions become less appropriate at higher resolutions, therefore the addition of targeted observations would not necessarily have the same impact for a higher resolution model as for a 24km grid model.

Throughout this thesis, dry TESVs have been used. This was the default provided by ECMWF for the field campaigns, and it was not possible to obtain moist total energy singular vectors in their place. The experiments conducted in chapter 6 used the leading SV from the TESV SAP for that targeting case, and therefore also used a dry TESV. Coutinho et al. (2004) found that the inclusion of latent-heat release in the computation of singular vectors resulted in singular vectors on a smaller scale, that exhibited enhanced growth rates compared to dry TESVs. The resolution that the SVs are computed at is therefore important, as SVs computed at higher resolution will exhibit smaller-scale structure. There were some similarities in the locations of the dry and moist TESVs: both identified baroclinic regions, but moist TESVs also occured in frontal regions, causing an enhancement or reduction in precipitation. Like the dry TESVs, the main growth mechanism for the moist TESVs was found to be PV-unshielding, but it was aided by latent-heat release instead of being solely dependent on the wind-shear. Overall, using dry TESVs instead of moist TESVs in this study may have affected the exact areas identified as sensitive by the singular vectors. However as the structure and growth mechanisms are similar for both types of TESV it should not compromise the results of the study. The choice of norm can also affect the structure of the SV, however HSVs computed with background error statistics based on an ensemble of 4D-Vars have similar structure to TESVs (Lawrence et al., 2009). The SVs used in this thesis start from a 24 or 36 hour forecast (see Section 4.2.2 for a discussion of this), therefore the location they identify may differ from an SV started from an analysis. However it was found during GFDex that there was good correspondence between SVs that started at longer forecast lead times and SVs that started from shorter forecast lead times therefore the location of the SVs does not seem to be very sensitive to the forecast lead time. Nonlinear processes have also been neglected in the computation of the TESVs; this is likely to have some effect on the structure and location identified by the TESVs.

The assessment of the impact of targeted observations of a polar low is limited by the use of 3D-Var. Some comparison of the magnitude of forecast impact from the GFDex and Norwegian IPY-THORPEX targeting experiments has been made, but it must be remembered that different types of assimilation schemes were used to assimilate the data. Some studies have shown that a greater impact is seen when 4D-Var is used to assimilate targeted data (Liu and Zou, 2001; Bergot, 2001), as 4D-Var schemes take into account both the time and

position of the observation, and allow for time-varying covariances whereas 3D-Var schemes account for only the position and use static error covariances. Kelly et al. (2007) showed with a 12-hour window 4D-Var scheme that these differences may limit the impact of extra data in a 4D-Var scheme, due to the ability of 4D-Var to propagate information from data-rich to data-poor regions. There are very few observations in the Arctic region, i.e. there is a lack of information to propagate, so in this case because of the time-varying covariances a larger impact may have been seen by using a 4D-Var scheme.

8.3 Future Work

This thesis has provided a contribution to the understanding of how targeted observations impact a forecast, and their potential to do so. The scope to extend this study further is discussed here.

In Chapter 4 a poor man's ensemble was used to assess the robustness of the forecast impact from the GFDex targeted sondes. It was shown that this poor man's ensemble had similar spread to the operational MOGREPS ensemble. The poor man's ensemble was used in lieu of the MOGREPS ensemble as it is not initialised at 1200 UTC (the start time for the GFDex forecasts) therefore ensemble perturbations were not available for 1200 UTC. Analysis of future targeting experiments could use the ECMWF or MOGREPS global ensembles, or MOGREPS limited-area ensemble to test the robustness of the forecast impact.

The covariance experiments demonstrate the potential increase in improvement that could be achieved by forcing a closer fit to the targeted observations in sensitive regions. An extension to this study would be to reduce the observation error covariances for the wind components only, or temperature only, instead of reducing the errors for u, v, T and RH at the same time. It was only possible to conduct this experiment for one case study; it would be interesting to repeat this experiment for different case studies and using different assimilation schemes to see if the same results are achieved.

The singular vector experiments detailed in Chapter 6 showed that the 4D-Var scheme was capable of producing tilted analysis increments that resembled a singular-vector structure. The singular vector used in the experiment was evolved using a linear model; in a nonlinear model the wave can be distorted as the PV contours wrap-up, therefore nonlinear processes

are an additional complication to understanding the growth of the singular-vector structure. Despite this, it might be possible to use a channel Eady model to investigate whether the evolved PV structures can be explained from the initial PV, temperature and wind fields, neglecting nonlinear processes.

The forecast degradation by two sondes close to the steep orography of Greenland is an important result as it has consequences for routine observations used to make operational weather forecasts. To determine the extent to which this problem affects operational forecasts, an extended study is required to examine whether routine radiosonde observations situated close to steep or high orography can degrade the forecast, and under what synoptic conditions the degradation occurs. Further study is required to assess whether this problem is unique to the Met Office data assimilation scheme; it is likely that operational centres which also use terrain-following height co-ordinates and a data assimilation scheme which does not have flow dependent error covariances will experience similar problems.

References

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.*, **131**, 1613–1628.

———, 2008: Large forecast degradations due to synoptic surveillance during the 2004 and 2005 hurricane seasons. *Mon. Wea. Rev.*, **136**, 3138–3150.

Aebischer, U. and C. Schär, 1998: Low-level potential vorticity and cyclogenesis in the lee of the Alps. *J. Atmos. Sci.*, **55**, 186–207.

Arakawa, A. and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.*, **17**, 173–265.

Arzel, O., T. Fichefet, and H. Goosse, 2006: Sea ice evolution over the 20th and 21st centuries as simulated by current AOGCMs. *Ocean Modelling*, **12**, 401–415.

Badger, J. and B. J. Hoskins, 2001: Simple initial value problems and mechanisms for baroclinic growth. *J. Atmos. Sci.*, **58**, 38–49.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*, **134**, 1951–1970.

Barkmeijer, J., R. Buizza, and T. N. Palmer, 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction Scheme. *Q. J. R. Meteorol. Soc.*, **125**, 2333–2351.

Bergot, T., 1999: Adaptive observations during FASTEX: A systematic survey of upstream flights. *Q. J. R. Meteorol. Soc.*, **125**, 3271–3298.

———, 2001: Influence of the assimilation scheme on the efficiency of adaptive observations. *Q. J. R. Meteorol. Soc.*, **127**, 635–660.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Bishop, C. H. and Z. Toth, 1999: Ensemble transformations and adaptive observations. *J. Atmos. Sci.*, **56**, 1748–1765.

Blechschmidt, A.-M., 2008: A 2-year climatology of polar low events over the Nordic Seas from satellite remote sensing. *Geophys. Res. Lett.*, **35**, L09 815.

Bouttier, F. and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Q. J. R. Meteorol. Soc.*, **127**, 1469–1488.

Bowler, N. E., A. Arribas, K. B. Mylne, K. R. Robertson, and S. E. Beare, 2008: The MO-GREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **134**, 703–722.

Bracegirdle, T. J. and S. L. Gray, 2008: An objective climatology of the dynamical forcing of polar lows in the Nordic seas. *Int. J. Climatol.*, **28**, 1903–1919.

Buizza, R., C. Cardinali, G. Kelly, and J. Thepaut, 2007: The value of targeted observations - Part II: the value of observations taken in singular vectors-based target areas. *Q. J. R. Meteorol. Soc.*, **133**, 1817–1832.

Buizza, R. and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.*, **56**, 2965–2985.

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

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Q. J. R. Meteorol. Soc.*, **135**, 239–250.

Cardinali, C. and R. Buizza, 2003: Forecast skill of targeted observations: A singular-vector-based diagnostic. *J. Atmos. Sci.*, **60**, 1927–1940.

Cardinali, C., R. Buizza, G. Kelly, M. Shapiro, and J. Thepaut, 2007: The value of targeted observations - Part III: weather regimes influence. *Q. J. R. Meteorol. Soc.*, **133**, 1833–1842.

Cardinali, C., S. Pezzulli, and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130**, 2767–2786.

Charney, J. and A. Eliassen, 1964: On the growth of the hurricane depression. *J. Atmos. Sci.*, **21**, 68–75.

Chelton, D. B., M. H. Freilich, J. M. Sienkiewicz, and J. Von Ahn, 2006: On the use of QuikSCAT scatterometer measurements of surface winds for marine weather prediction. *Mon. Wea. Rev.*, **134**, 2055–2071.

Condron, A., G. R. Bigg, and I. A. Renfrew, 2006: Polar mesoscale cyclones in the Northeast Atlantic: comparing climatologies from ERA-40 and satellite imagery. *Mon. Wea. Rev.*, **134**, 1518–1533.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387.

Coutinho, M. M., B. J. Hoskins, and R. Buizza, 2004: The influence of physical processes on extratropical singular vectors. *J. Atmos. Sci.*, **61**, 195–209.

Daley, R., (Ed.) , 1991: *Atmospheric Data Analysis*. Cambridge University Press.

Dando, M. L., 2007: The impact of targeted satellite observations on numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **133**, 1945–1957.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, and A. Staniforth, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, **131**, 1759–1782.

Dumelow, R., 2008: The impact of conventional observations on global and regional NWP forecasts. *Met Office Technical Report No. 516*, 1–21.

Duncan, C. N., 1978: Baroclinic instability in a reversed shear flow. *Meteorological Magazine*, **107**, 17–23.

Emanuel, K. A. and R. Rotunno, 1989: Polar lows as Arctic hurricanes. *Tellus A*, **41A**, 1–17.

Fisher, M. and E. Andersson, 2001: Developments in 4D-var and Kalman filtering. *ECMWF Tech. Memo.*, **347**, 36pp.

Fourrié, N., D. Marchal, F. Rabier, B. Chapnik, and G. Desroziers, 2006: Impact study of the 2003 North Atlantic THORPEX Regional Campaign. *Q. J. R. Meteorol. Soc.*, **132**, 275–295.

Gelaro, R., R. H. Langland, G. D. Rohaly, and T. E. Rosmond, 1999: An assessment of the singular-vector approach to targeted observing using the FASTEX dataset. *Q. J. R. Meteorol. Soc.*, **125**, 3299–3327.

Gilmour, I., L. A. Smith, and R. Buizza, 2001: Linear regime duration: Is 24hrs a long time in synoptic weather forecasting? *J. Atmos. Sci.*, **58**, 3525–3539.

Graham, R. J., S. R. Anderson, and M. J. Bader, 2000: The relative utility of current observation systems to global-scale NWP forecasts. *Q. J. R. Meteorol. Soc.*, **126**, 2435–2460.

Gregory, D. and P. R. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506.

Harold, J. M., G. R. Bigg, and J. Turner, 1999a: Mesocyclone activity over the north-east Atlantic. Part 1: Vortex distribution and variability. *Int. J. Climatol.*, **19**, 1187–1204.

———, 1999b: Mesocyclone activity over the north-east Atlantic. Part 2: An investigation of causal mechanisms. *Int. J. Climatol.*, **19**, 1283–1299.

Hartmann, D. L., R. Buizza, and T. N. Palmer, 1995: Singular vectors: the effect of spatial scale on linear growth of disturbances. *J. Atmos. Sci.*, **52**, 3885–3894.

Hollingsworth, A. and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. part I: The wind field. *Tellus*, **38A**, 111–136.

Holton, J. R., (Ed.) , 2004: *An introduction to dynamic meteorology*. Elsevier Academic Press.

Hoskins, B. J., R. Buizza, and J. Badger, 2000: The nature of singular vector growth and structure. *Q. J. R. Meteorol. Soc.*, **126**, 1565–1580.

Hoskins, B. J. and M. M. Coutinho, 2005: Moist singular vectors and the predictability of some high impact European cyclones. *Q. J. R. Meteorol. Soc.*, **131**, 581–601.

Ingleby, A. C., N. B. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Q. J. R. Meteorol. Soc.*, **119**, 1195–1225.

Ingleby, N. B., 1998: OPS scientific documentation paper number 5. Radiosonde processing. *Met Office Internal Report*.

———, 2001: The statistical structure of forecast errors and its representation in the Met. Office Global 3-D variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **127**, 209–231.

Ingleby, N. B. and A. C. Lorenc, 1998: OPS scientific documentation paper number 2. Generic quality control. *Met Office Internal Report*.

Irvine, E. A., S. L. Gray, J. Methven, I. A. Renfrew, K. Bovis, and R. Swinbank, 2009: The impact of targeted observations made during the Greenland Flow Distortion Experiment. *Q. J. R. Meteorol. Soc.*, **135**, 2012–2029.

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

Kalnay, E., (Ed.) , 2003: *Atmospheric Modelling, data assimilaton, and predictability*. Cambridge University Press.

Kelly, G., J. Thepaut, R. Buizza, and C. Cardinali, 2007: The value of targeted observations - Part I: data denial experiments for the Atlantic and the Pacific. *Q. J. R. Meteorol. Soc.*, **133**, 1803–1815.

Klein, T. and G. Heinemann, 2002: Interaction of katabatic winds and mesocyclones near the eastern coast of greenland. *Meteorol. Appl.*, **9**, 407–422.

Kolstad, E. and T. J. Bracegirdle, 2008: Marine cold-air outbreaks in the future: an assessment of IPCC AR4 model results for the Northern Hemisphere. *Clim Dyn*, **30**, 871–885.

Kolstad, E. W., 2006: A new climatology of favourable conditions for reverse-shear polar lows. *Tellus*, **58A**, 344–354.

Langland, R. H., 2005: Issues in targeted observing. *Q. J. R. Meteorol. Soc.*, **131**, 3409–3425.

Langland, R. H., Z. Toth, R. Gelaro, I. Szunyogh, M. A. Shapiro, S. J. Majumdar, R. E. Morss, G. D. Rohaly, C. Velden, N. Bond, and C. H. Bishop, 1999: The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bull. Am. Met. Soc.*, **80**, 1363–1384.

Lawrence, A. R., M. Leutbecher, and T. N. Palmer, 2009: The characteristics of hessian singular vectors using an advanced data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **135**, 1117–1132.

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.

Leutbecher, M. A., A. Doerenbecher, F. Grazzini, and C. Cardinali, 2004: Planning of adaptive observations during the Atlantic THORPEX Regional Campaign 2003. *ECMWF Newsletter, Winter Issue*, **102**, 16–25.

Liu, H. and X. Zou, 2001: The impact of NORPEX targeted dropsondes on the analysis and 2-3 day forecasts of a landfalling Pacific winter storm using NCEP 3DVAR and 4DVAR systems. *Mon. Wea. Rev.*, **129**, 1987–2004.

Lock, A. P., M. R. Brown, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*, **128**, 3187–3199.

Lorenc, A. C., 2007: A study of o-b monitoring statistics from radiosondes, composited for low-level cloud layers. *Met Office R & D Technical Report 504*.

Lorenc, A. C., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders, 2000: The Met Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126**, 2991–3012.

Lorenc, A. C. and O. Hammon, 1988: Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Q. J. R. Meteorol. Soc.*, **114**, 515–543.

Lorenc, A. C., I. Roulstone, and A. White, 2003: On the choice of control fields in Var. *Met Office Internal Report*.

Majumdar, S. J., S. D. Aberson, C. H. Bishop, R. Buizza, M. S. Peng, and C. A. Reynolds, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.*, **134**, 2354–2372.

Majumdar, S. J., C. H. Bishop, R. Buizza, and R. Gelaro, 2002a: A comparison of ensemble transform Kalman filter targeting guidance with ECMWF and NRL total energy singular vector guidance. *Q. J. R. Meteorol. Soc.*, **128**, 2527–2549.

Majumdar, S. J., C. H. Bishop, and B. J. Etherton, 2002b: Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Mon. Wea. Rev.*, **130**, 1356–1369.

Majumdar, S. J., C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast error variance produced by targeted observations? *Q. J. R. Meteorol. Soc.*, **127**, 2803–2820.

McNally, A. P., 2002: A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128**, 2551–2556.

Moninger, W. R., R. D. Mamrosh, and P. M. Pauley, 2003: Automated meteorological reports from commercial aircraft. *Bull. Am. Met. Soc.*, **84**, 203–216.

Montani, A., A. J. Thorpe, R. Buizza, and P. Unden, 1999: Forecast skill of the ECMWF model using targeted observations during FASTEX. *Q. J. R. Meteorol. Soc.*, **125**, 3219–3240.

Moore, G. W. K., R. S. Pickart, and I. A. Renfrew, 2008: Buoy observations from the windiest location in the world ocean, Cape Farewell, Greenland. *Geophys. Res. Lett.*, **35**, L18 802.

Moore, G. W. K. and I. A. Renfrew, 2005: Tip jets and barrier winds: A QuikSCAT climatology of high wind speed events around Greenland. *J. Climate*, **18**, 3713–3725.

Nordeng, T. E., 1990: A model-based diagnostic study on the development and maintenance of two polar lows. *Tellus*, **42A**, 92–108.

Ólafsson, H. and P. Bougeault, 1997: The effect of rotation and surface friction on orographic drag. *J. Atmos. Sci.*, **54**, 193–210.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. *J. Atmos. Sci.*, **55**, 633–653.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis scheme. *Mon. Wea. Rev.*, **120**, 1747–1763.

Persson, A., 1999: Synoptic-dynamic diagnosis of medium range weather forecast systems. *Proc. Seminar on Diagnosis of Models and Data Assimilation Systems*, ECMWF, Reading, United Kingdom, 123–137.

Petersen, G. N., S. J. Majumdar, and A. J. Thorpe, 2007: The properties of sensitive area predictions based on the ensemble transform Kalman filter (ETKF). *Q. J. R. Meteorol. Soc.*, **133**, 697–710.

Petersen, G. N., H. Ólafsson, and J. E. Kristjánsson, 2003: Flow in the lee of idealised mountains and Greenland. *J. Atmos. Sci.*, **60**, 2183–2195.

Petersen, G. N., I. A. Renfrew, and G. W. K. Moore, 2009: An overview of barrier winds off southeastern Greenland during GFDex. *Q. J. R. Meteorol. Soc.*, In press.

Petersen, G. N. and A. J. Thorpe, 2007: The impact on weather forecasts of targeted observations during A-TReC. *Q. J. R. Meteorol. Soc.*, **133**, 417–431.

Pettersen, S. and S. J. Smebye, 1971: On the development of extratropical cyclones. *Q. J. R. Meteorol. Soc.*, **97**, 457–482.

Rabier, F. and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.*, **118**, 649–672.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.*, **122**, 121–150.

Rasmussen, E., 1979: The polar low as an extratropical CISK disturbance. *Q. J. R. Meteorol. Soc.*, **105**, 531–549.

Rasmussen, E. A. and J. Turner, (Eds.) , 2003: *Polar Lows: Mesoscale Weather Systems in the Polar Regions*. Cambridge University Press.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **133**, 347–362.

Renfrew, I. A., G. W. K. Moore, J. E. Kristjánsson, H. Ólafsson, S. L. Gray, G. N. Petersen, K. Bovis, P. R. A. Brown, I. Føre, T. Haine, C. Hay, E. A. Irvine, A. Lawrence, T. Ohigashi, S. Outten, R. S. Pickart, M. Shapiro, D. Sproson, R. Swinbank, A. Woolley, and S. Zhang, 2008: The Greenland Flow Distortion Experiment. *Bull. Am. Met. Soc.*, **89**, 1307–1324.

Renfrew, I. A., G. N. Petersen, D. A. J. Sproson, G. W. K. Moore, H. Adiwidjaja, S. Zhang, and R. North, 2009: A comparison of aircraft-based surface-layer observations over Denmark Strait and the Irminger Sea with meteorological analyses and QuikSCAT winds. *Q. J. R. Meteorol. Soc.*, In press.

Shapiro, M. A. and A. J. Thorpe, 2004: THORPEX international science plan version III. *WMO Publication*.

Shutts, G. J., 1990: Dynamical aspects of the October storm, 1987: A study of a successful fine-mesh simulation. *Q. J. R. Meteorol. Soc.*, **116**, 1315–1347.

Simmons, A. J. and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128**, 647–677.

Szunyogh, I., Z. Toth, K. A. Emanuel, C. H. Bishop, C. Snyder, R. E. Morss, J. Woolen, and T. Marchok, 1999: Ensemble-based targeting experiments during FASTEX: The effect of dropsonde data from the Lear jet. *Q. J. R. Meteorol. Soc.*, **125**, 3189–3217.

Szunyogh, I., Z. Toth, R. E. Morss, S. J. Majumdar, B. J. Etherton, and C. H. Bishop, 2000: The effect of targeted dropsonde observations during the 1999 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.*, **128**, 3520–3537.

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.*, **130**, 1144–1165.

Thépaut, J.-N., P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four-dimensional variational data assimilation: a case study. *Q. J. R. Meteorol. Soc.*, **122**, 535–561.

Vaisala, 2006: Vaisala radiosonde RS92-SGP brochure. http://www.vaisala.com.

Weissmann, M., R. Busen, A. Dörnbrack, S. Rahm, and O. Reitebuch, 2005: Targeted observations with an airborne wind lidar. *J. Atmos. Oceanic Technol.*, **22**, 1706–1719.

Weissmann, M. and C. Cardinali, 2007: Impact of airborne Doppler lidar observations on ECMWF forecasts. *Q. J. R. Meteorol. Soc.*, **133**, 107–116.

Wilhelmsen, K., 1985: Climatological study of gale-producing polar lows. *Tellus*, **37A**, 451–459.

Wilson, D. R. and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.*, **125**, 1607–1636.

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

APPENDIX A

Acronyms and Abbreviations

Acronym or Abbreviation	Definition
3D-Var	Three-Dimensional Variational data assimilation
4D-Var	Four-Dimensional Variational data assimilation
ACARS	Aircraft Communications Addressing and Reporting System
AIREPS	Aircraft Reports
AMDAR	Aircraft Meteorological Data Relay
AMSU-A	Advanced Microwave Sounding Unit -A
ATOVS	Advanced TIROS Operational Vertical Sounder
ATReC	Atlantic THORPEX regional campaign (field campaign)
AVHRR	Advanced Very-High Resolution Radiometer
BLUE	Best Linear Unbiased Estimate
CAMM	Crisis-Area Mesoscale Model (MetUM)
CAPE	Convective Available Potential Energy
CISK	Conditional Instability of the Second Kind
DLR	Deutsches Zentrum fur Luft und Raumfahrt (German Aerospace Centre)
DOTSTAR	Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (field campaign)
DTS	Data Targeting Suite
ECMWF	European Centre for Medium-Range Weather Forecasts
EnKF	Ensemble Kalman Filter
EPS	Ensemble Prediction System
ERA-40	40-year ECMWF Re-analysis
ET	Ensemble Transform
ETKF	Ensemble Transform Kalman Filter
Eurorisk-PREVIEW	Eurorisk Prevention, Information and Early Warning programme
FAAM	Facility for Atmospheric Airborne Measurement (UK)
FASTEX	Fronts and Atlantic Storm-Track Experiment (field campaign)
GFDex	Greenland Flow Distortion Experiment (field campaign)
GOS	Global Observing System
GPS	Global Positioning System
GTS	Global Telecommunications System
HIRLAM	High-Resolution Limited Area Model
HSV	Hessian Singular Vector
IPCC	Intergovernmental Panel on Climate Change
IPY-THORPEX	International Polar Year - THORPEX (field campaign)
LBC	Lateral Boundary Condition

Met.No	Norwegian Meteorological Service
MetUM	The UK Met Office Unified Model
MODIS	Moderate Resolution Imaging Spectroradiometer
MOGREPS	Met Office Global and Regional Ensemble Prediction System
NAE	North-Atlantic European (limited-area domain for MetUM)
NCEP	National Centers for Environmental Prediction (USA)
NMC	National Meteorological Center (USA)
NORPEX	North Pacific Experiment (field campaign)
NWP	Numerical Weather Prediction
OPS	Observation Processing System
OSE	Observing System Experiment
PGE	Probability of Gross Error
PV	Potential Vorticity
QuikSCAT	Quick Scatterometer
SAP	Sensitive Area Prediction
SSM/I	Special Sensor Microwave Imager
SST	Sea-surface Temperature
SV	Singular Vector
TESV	Total Energy Singular Vector
THORPEX	The Observing System Research and Predictability Experiment
TIROS	Television and Infra-red Observational Satellite
VAR	Variational Data Assimilation
WISHE	Wind-Induced Surface Heat Exchange
WSR	Winter Storm Reconnaisance Program (field campaign)