Met Office

Meteorology Research and Development

An assessment of deployment strategies for targeting observations

Technical Report No. 515

Keir Bovis, Gareth Dow, Richard Dumelow and Mike Keil

email:nwp_publications@metoffice.gov.uk

An assessment of deployment strategies for targeting observations

Keir Bovis, Gareth Dow, Richard Dumelow and Mike Keil

25th March 2008

Abstract

The primary focus of this study is in the deployment of additional radiosonde observations within targeting regions with the aim of improving a forecast for a pre-defined verification region and forecast range. Two approaches two identifying targeting regions are evaluated. The first is undertaken by a forecaster using forecast products and in the second, regions are identified using the Ensemble Transform Filter Kalman Filter (ETKF). The quality of the targeting guidance is determined in two ways. By running a series of Observation System Experiments (OSEs), the Root Mean Square (RMS) error from each may be compared for a month-long trial period. Secondly, RMS error from each OSE is calculated in the presence of high impact case studies. As a secondary research aim, we undertake an evaluation of different satellite data thinning strategies.

1 Introduction

Global Met centres have recognised the benefit for some years of complementing the routine observing network with a deployment of targeted observations in an attempt to improve the quality of Numerical Weather Prediction (NWP) forecasts [19, 12, 16]. The location of these observations may be fixed, as in the case of requesting additional radiosonde ascents, or mobile, for example the deployment of dropsondes from aircraft or remotely sensed satellite observation data. Regardless of the location, each of these targeted observations are obtained from, or deployed in, a pre-defined area or *targeting region*.

Results previously presented by Dumelow *et al.* [9], subsequently referred to as D2006, have demonstrated the benefit of deploying targeted radiosonde observations in Met Office NWP models. By running a series of Observation System Experiments (OSEs), the benefit was quantified by determining the impact on the quality of forecasts verifying within a pre-defined *verification area*. The authors concluded that for maximum impact, the location of the targeted observations were dependent on the geographical definition of the verification area and the *target forecast range* they sought to improve. Using the verification region and target forecast range as parameters, the study included the identification of targeting regions within which to deploy radiosonde observations comprising an optimal targeting network. The process of identifying the targeting regions is termed *sensitive area prediction* and in D2006, the authors undertook a T+0 500 hPa flow analysis to achieve this. They found that the optimal targeting regions were typically located upstream of the verification area and incorporation of these targeted observations into NWP models in the presence of the routine network, resulted in the largest reduction in forecast error when compared with other targeting strategies. The operational implementation of this deployment, based on a T+0 flow analysis, would be impractical in reality. This is because the data time of sensitive area prediction and targeting were identical, in other words the *lead time* was zero. In addition, the overhead of manually performing such an analysis would be costly in terms of human forecasting resource and possibly prone to error.

This study presents results from experiments undertaken to fulfil three specific research aims:

1. To evaluate the impact on NWP forecasts by deploying further configurations of targeted radiosondes based on sensitive area predictions defined by forecasters. This will follow on from the work undertaken in D2006.

2. To determine the impact on NWP forecasts of the deployment of targeted radiosondes within sensitive area predictions computed by statistical methods using an operationally realistic lead time. The process of identifying the optimal adaptive network is know as *observation targeting,* the main elements of which are summarised in Table 1 and their connectivity shown in Figure 1.

3. To investigate the impact on NWP forecasts of targeting satellite data compared with an operational baseline.

This report is set out as follows. The experimental design for all experiments is presented in Section 2. Results from the experimental runs are given in Section 3 comprising, verification of standard meteorological fields and

tropical cyclone impact studies. A discussion of the key results and conclusions can be found in Section 4 together with recommendations for future work.

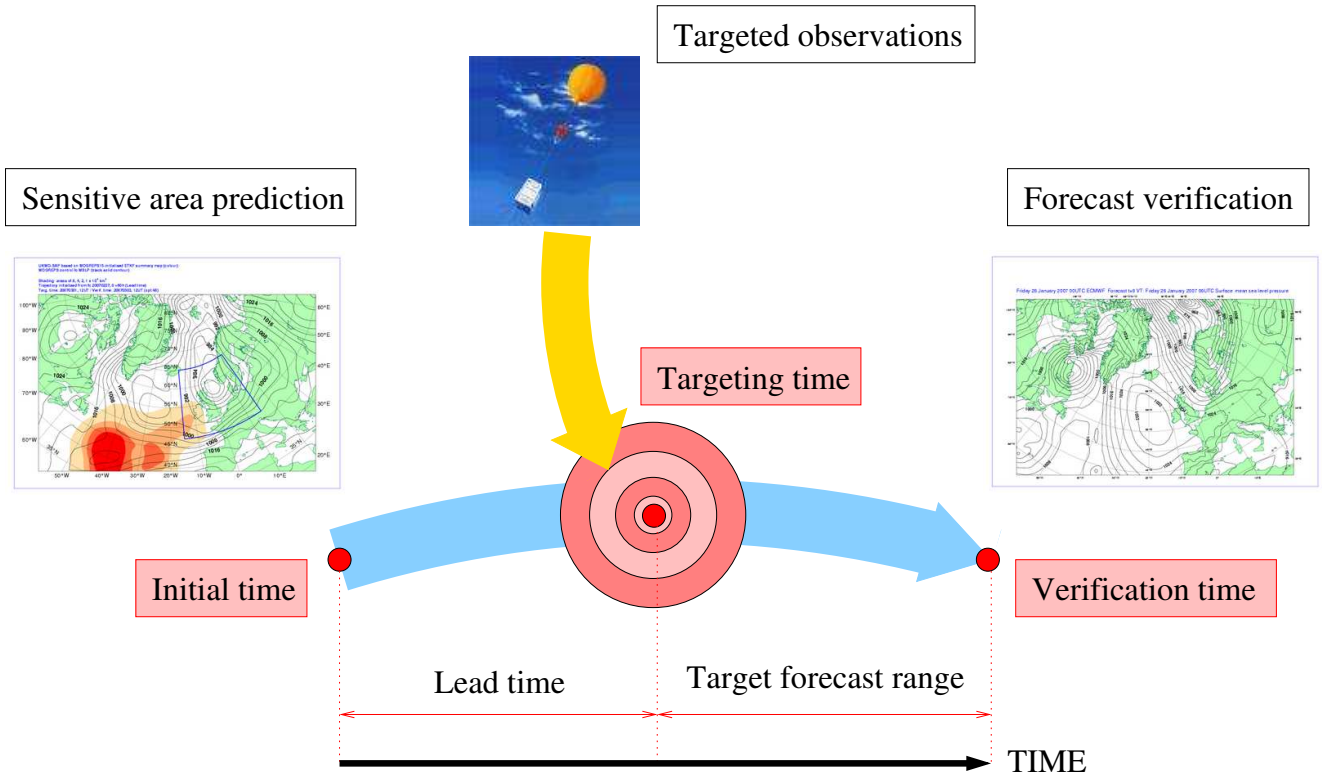


Figure 1: Main elements of the observation targeting paradigm (adapted from Doerenbecher *et al.* [7]).

Initial time	The time at which the sensitive area prediction is generated.
Targeting time	The time of deployment of the targeted observations.
Verification time	The time the target forecast is to be verified.
Sensitive area prediction	The method used to identify the targeting region.
Targeted observations	Targeted observations are made at targeting time deployed in a targeting region.
Forecast verification	The forecast is verified within the verification region where an improvement in forecast impact is sought.
Targeting region	The location of where the targeted observations are to be deployed.
Target forecast range	The forecast range to be improved, typically the length of time between the targeting and verification times.
Lead time	The length of time between production of the sensitive area prediction and the targeting time.

Table 1: Main elements of an observation targeting paradigm.

2 Experimental set-up

This section describes the methods used in the experiments to investigate the three broad research aims introduced in Section 1. To investigate each candidate deployment strategy, we undertake a series of OSEs. Each OSE is based on the Met Office's North American Crisis Area Mesoscale (CAM) model [11].

Tropical cyclone	Landfall	Date
Hurricane Katrina	$30^oN/89^oW$	29 August 2005 12 UTC
Hurricane Ophelia	$34^oN/77^oW$	14 September 2005 00 UTC
Hurricane Rita	$30^oN/93^oW$	24 September 2005 12 UTC
Hurricane Otis	$20^oN/110^oW$	30 September 2005 12 UTC

Table 2: Tropical cyclones occurring during the period within the CAM domain.

2.1 OSE configuration

The CAM is used to evaluate different deployments of targeted observations. Its domain definition is shown in Figure 2(a) together with the verification area whose horizontal spatial coordinates bottom left and top right-hand points are $\{35^oN/98^oW/45^oN/85^oW\}$. The verification region is chosen such that radiosondes are available for deployment upstream in any direction depending on the location of the sensitive areas. The CAM resolution is based on a 17 km horizontal rotated grid with 38 vertical model levels and uses a 3D-variational assimilation scheme. It generates forecasts in the ranges T+12, T+24, T+36 and T+48 at 00 and 12 UTC cycles of each trial day. The CAM configuration is different to that used in D2006. At each cycle, all OSEs are initialised from an identical observation and assimilation background valid at that time[1]. By doing this, the impact of each deployment of targeted radiosonde observations is apparent and not masked by the evolving model background. Boundary conditions for the regional model are obtained from the Met Office's Global model. The OSE trial period spans 29 August 2005 12 UTC through to 1 October 2005 12 UTC and is made up of 66 forecast cycles. The period includes the four tropical cyclones listed in Table 2. These tropical cyclones will be the subject of case studies for which results are presented in Section 3.

A routine radiosonde network is constructed using the method defined in D2006. Radiosonde station locations are subjectively selected from the existing North American network and thinned to an approximate resolution of one per 10^o latitude/longitude box. Targeted radiosonde observations may then be added by selecting from those remaining stations in the existing North American network. No routine observations are deployed within the verification area. Throughout the trial period, all aircraft observations have been removed to improve the radiosonde observation targeting signal. The impact of aircraft observations has been shown to lead to a positive effect on the reduction of forecast error in regional models [8]. Re-running these experiments with the inclusion of these observations may lead to a reduction in any subsequently reported positive impact. Figure 2(a) shows the locations of radiosonde stations that are used routinely (red dots) and those that may be considered for observation targeting (black crosses). At each cycle, the routine radiosonde network may be deployed (depending on the strategy used) together with targeted radiosonde observations whose selection is dependant on the deployment strategy being evaluated.

2.2 Evaluation of targeted radiosonde observations deployed within forecaster defined regions

We consider the following strategies to deploy targeted radiosonde observations and label them appropriately for subsequent reference in this paper:

VER_ONLY In this configuration a fixed deployment of targeted radiosonde observations are deployed within the *verification area only* at each cycle as shown in Figure 2(b). No routine radiosonde network is deployed. The aim of the experiment is to demonstrate the limitations of a strategy that simply deploys within the verification area.

UPS_ONLY This strategy contains no deployment of the routine observing network. Targeted observations are only deployed upstream of the verification area in targeting regions based on the 500 hPa flow analysis valid at targeting time as described and undertaken in D2006. This experiment will demonstrate the utility of the routine baseline observing network. Stations are selected with the aim of improving forecasts in the range of 36 hours.

BASE+T36 In this strategy a routine observation network is deployed together with 10 targeted observations at each cycle. The deployment location of the targeted radiosondes is within regions based on a 500 hPa flow

[1]The observation and assimilation background are taken from an operationally configured equivalent OSE labelled BASE+ALL, run as a meta-control.

analysis using a *36 hour forecast* valid at the targeting time. Radiosonde stations are selected with the aim of improving forecasts in the range of 36 hours.

We compare the impact of these deployment strategies with the OSEs defined originally in D2006. For completeness they are described below:

BASE+UPS At each cycle routine radiosonde observations are deployed together with 10 targeted observation in regions based on the 500 hPa flow analysis valid at targeting time (T+0) described and undertaken in D2006. Radiosonde stations are selected with the aim of improving forecasts in the range of 36 hours.

BASE This strategy comprises the routine network only . The location of the baseline network is shown in Figure 2(c).

BASE+ALL This deployment strategy contains all radiosonde observations used operationally. We run this configuration to provide an analysis for verifying case study forecasts presented later. It corresponds to all of the radiosonde stations shown in Figure 2(a).

2.3 Generation of sensitive area predictions using the Ensemble Transform Kalman Filter (ETKF)

We seek to quantify the impact on NWP forecasts of deploying additional targeted radiosonde observations within regions computed using statistical methods. In this study the method considered is the Ensemble Transform Kalman Filter (ETKF) first proposed by Bishop *et al.* [1, 15]. A complete description of the ETKF can be found in Bovis (2008) [4] but a brief account of its capability follows. Preliminary results of its use have already been previously presented [2].

The ETKF seeks a deployment of targeted observations in addition to a routine network that will minimise the forecast error in a pre-defined verification area and forecast range. It is based on the Kalman Filter [14] and utilises an ensemble of forecasts to estimate model uncertainty within a probabilistic framework. In all our experiments we use ECMWF's Ensemble Prediction System (EPS) [6] to initialise the ETKF and estimate the analysis error covariance at targeting and verification times. The ETKF utilises a novel data transformation technique facilitating the identification of optimal locations of the targeted observation in the presence of a large combination of deployment possibilities.

The OSE configuration used to evaluate each deployment is identical to that described in Section 2.2. At each forecast cycle, the ETKF generates a model of sensitivity with respect to a deployment of targeted observations in the presence of the routine network. In the ETKF this is called the *signal variance*. The signal variance is interpolated to the locations of all the candidate radiosonde locations available for targeting. Radiosonde observations from the top-10 ranked radiosonde locations maximising the ETKF signal variance are automatically selected for deployment. Figure 3 shows an example ETKF sensitive area prediction for deploying targeted observations on 29 August 2005 12 UTC.

We evaluate the following strategies for the deployment of targeted radiosonde observations in regions identified using the ETKF and label them appropriately for subsequent reference in this paper:

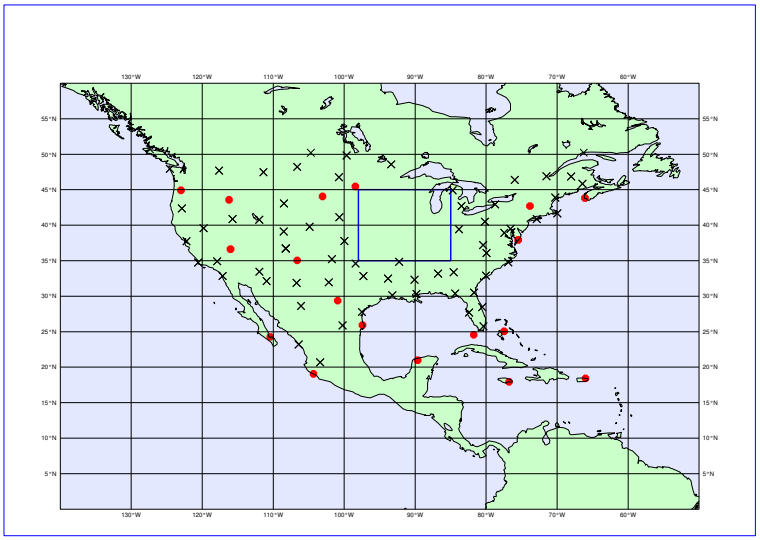
ETKF(24) In this approach the ETKF is used to create targeting guidance for each cycle during the trial period. The lead time used is 24 hours and we seek to deploy targeted observations with the aim of improving 36 hour forecasts.

ETKF(48) This configuration of the ETKF is identical to ETKF(24) except a longer lead time of 48 hours is used.

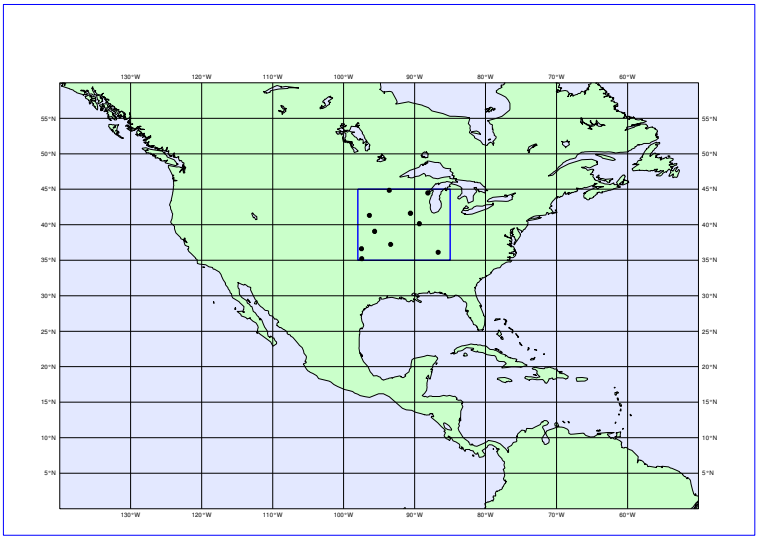
We compare the impact from these targeting strategies with OSEs BASE+T36 and BASE+UPS defined in Section 2.2.

2.4 Experimental methods for the adaptive thinning of satellite data

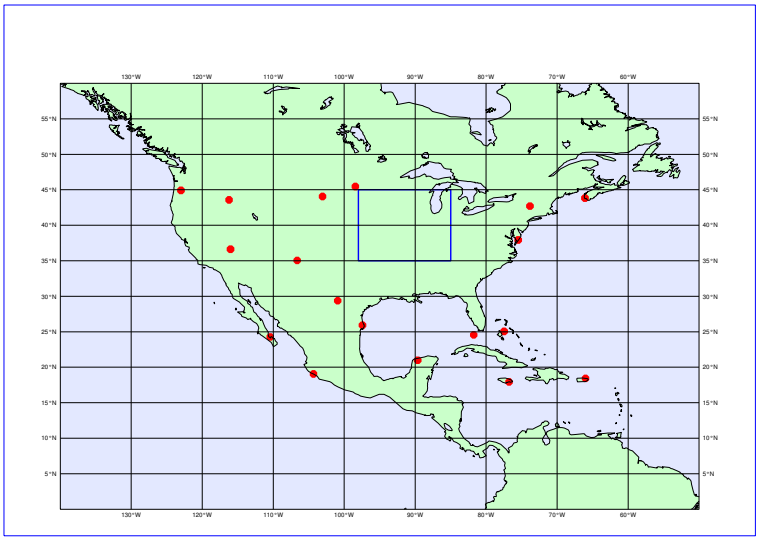
In this section we describe experiments undertaken to evaluate the utility of targeting satellite data. Previous studies have shown that increasing the observation density for satellite observations with uncorrelated error improves the analysis accuracy [17]. We assess the utility of incorporating adaptively thinned Advanced Microwave Sounding Unit (AMSU) radiance data obtained from satellites with that of an operational NWP baseline. To evaluate each



(a)



(b)



(c)

Figure 2: (a) CAM domain with verification area shown as a blue box (candidate surface stations for observation targeting and routine surface network shown as black crosses and red circles respectively); (b) location of the 10 static radiosonde stations located within the verification area; (c) the routine radiosonde network used in BASE deployment strategy.

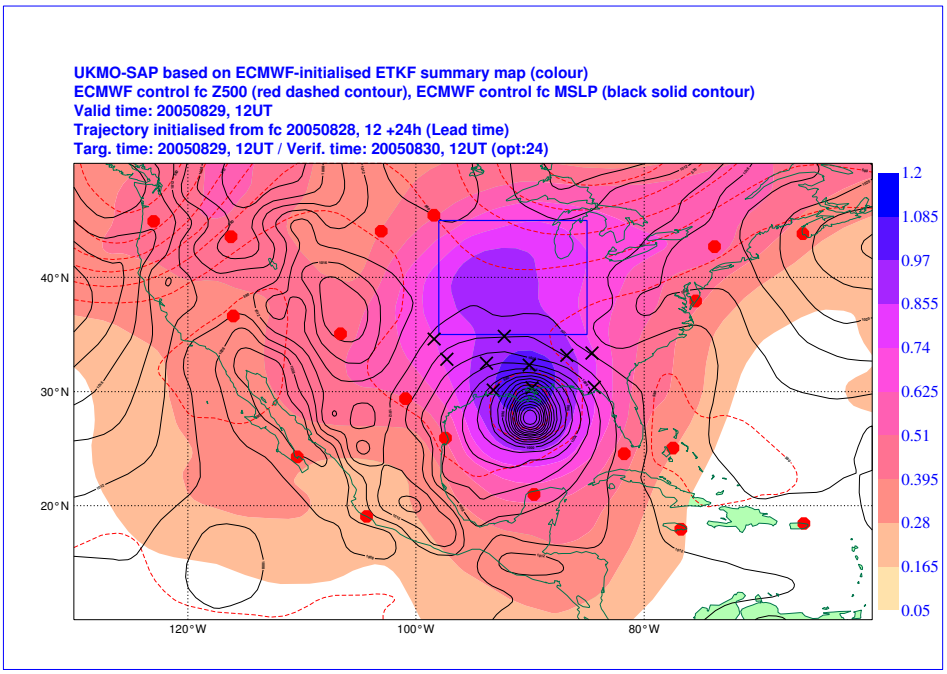


Figure 3: Example sensitive area prediction generated using the ETKF for targeting on 29 August 2005 12 UTC showing the ETKF signal variance (filled coloured contour) and forecast Mean Sea Level Pressure (MSLP, black contour) and 500 hPa geopotential height (Z500, red contour). Radiosonde stations selected for targeting are shown as black crosses and the routine observing network shown in red.

AMSU data thinning strategy, we run an OSE based on the Met Office's CAM described in Section 2.2 utilising a routine radiosonde observation network with no targetable radiosondes. Throughout the OSE trial period, surface and satellite observations are assimilated as normal except that the thinning of AMSU data is undertaken according to different strategies. Preliminary results have already been presented [5, 3]. We define the following OSEs corresponding to different AMSU data thinning strategies evaluated in this study:

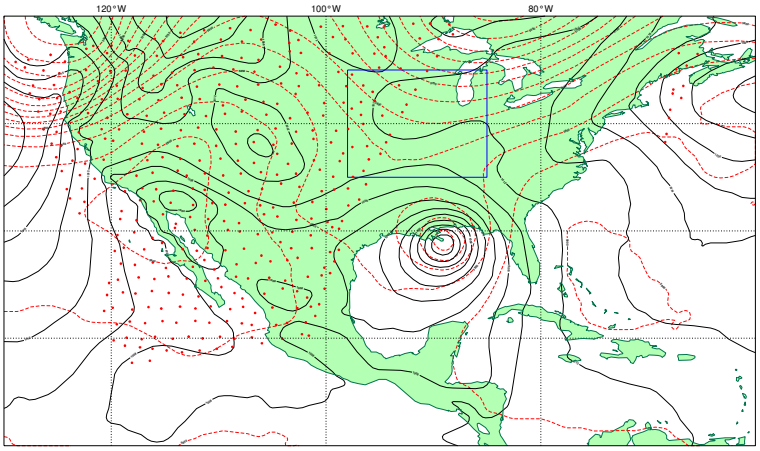
D154 AMSU data thinning is undertaken globally at the resolution of one observation every 154 km. An example swathe is shown in Figure 4(a). This acts as a baseline to quantify the impact of other AMSU data thinning strategies.

D40 AMSU data is thinned globally but at the finer resolution of one observation every 40 km. An example of typical data is shown in Figure 4(b). This was the operational configuration used in all Met Office local area models at the outset of this study.

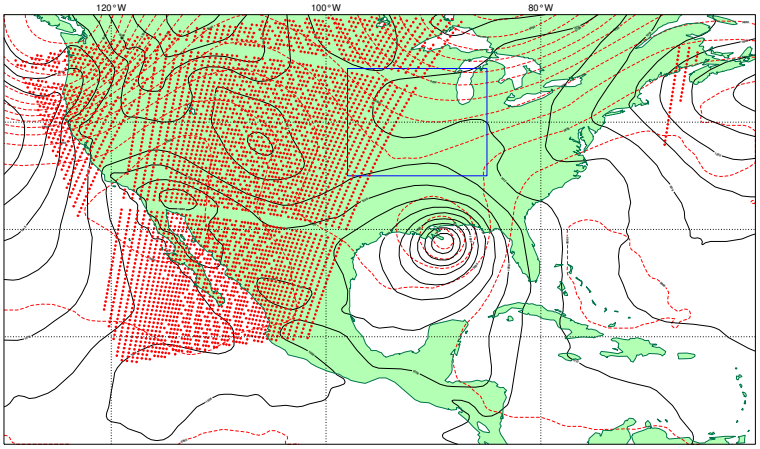
ETKF(24) In this strategy, AMSU data thinning is undertaken at two different densities. Within regions maximising the signal variance generated by the ETKF, AMSU data thinning is carried out at a resolution of one observation every 40 km, the effective maximum data usage. This results in a higher density of AMSU observations occurring in these regions. Outside of these areas, a thinning resolution of one observation every 154 km is used. Thinning at the finer resolution is undertaken within regions of an approximate size 2×10^6 km. The ETKF is configured to run with a lead time of 24 hours and with the aim of improving 24 hour forecasts in the previously defined verification region. The adaptive thinning of AMSU data is performed for every cycle during the trial period. An example of typical data is shown in Figure 4(c). From this figure, it can be seen that adaptively thinned AMSU data may occur within the verification region.

3 Results

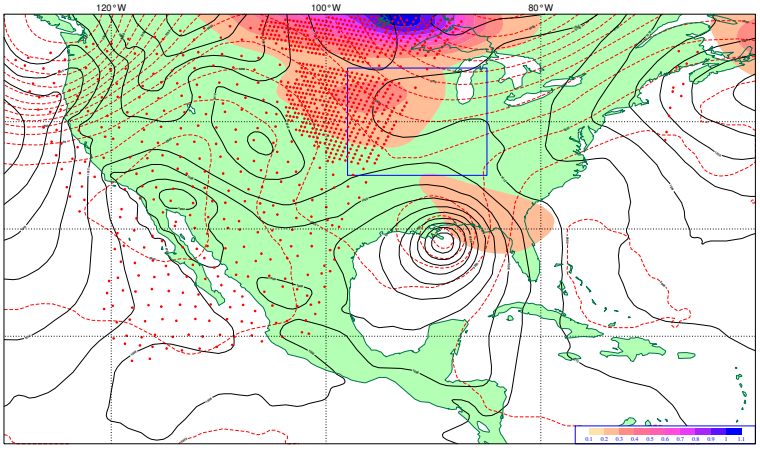
In this study we evaluate experimental impact in three different ways. In Section 3.1 we assess the performance of each deployment strategy over the whole trial period by computing the forecast Root Mean Square (RMS) error



(a)



(b)



(c)

Figure 4: AMSU data thinned observation swathe examples; (a) globally at one observation every 154 km used in D154; (b) globally at one observation every 40 km used in D40; (c) adaptive thinning at one observation every 154 km outside sensitive area prediction (shown as filled coloured contour) and one observation every 40 km within this area.

for a selection of forecast fields. This is followed in Section 3.2 by undertaking a series of case studies examining the forecasting of tropical cyclones with tracks that pass through the verification region.

3.1 Experimental results for forecast RMS error

To determine the impact of the different deployment strategies, we verify the forecasts generated at each 00 and 12 UTC forecast cycle throughout the trial period against radiosonde observations deployed within the verification region. Key fields considered in this study are geopotential height, temperature, wind speed and relative humidity at 500 hPa. Each of these fields may be used to diagnose the modelling of synoptic weather conditions within the CAM. Forecast RMS error is computed for a set of candidate forecast ranges $\{T+0, T+12, T+24, T+36, T+48\}$. To assess the performance of each strategy at different heights in the atmosphere, we present forecast RMS error results for different pressure levels at the target forecast range of T+36 for targeted radiosonde experiments and at T+24 for AMSU data thinning experiments.

Targeted radiosonde deployment strategies

Figures 5(a)-(d) show the mean forecast RMS error for each of the deployment strategies discussed in Section 2.2. From the verification of 500 hPa geopotential height in Figure 5(a), we see that OSE VER_ONLY gives the smallest mean forecast error up to T+24 but then BASE+T36 shows the largest impact beyond this forecast range. This result highlights that at short forecast ranges, the upstream flow is less important and that a good analysis of the verification region itself is more important. At T+48 there is little to separate the competing deployment strategies. A broadly similar pattern is observed in the verification of the other forecast fields. This is shown in Figures 5(b)-(d) but we observe that the performance of OSE VER_ONLY is superior only to T+12. From these results we can see that OSEs that undertake the deployment of the routine and targeted observations generally out-perform those that don't deploy both networks. For example, OSEs BASE+T36 and BASE+UPS out-perform OSEs VER_ONLY, BASE and UPS_ONLY beyond the T+12 forecast range. Differences in forecast error between each approach at different levels for the target 36-hour forecast range are harder to identify. This is illustrated by the results for forecast geopotential height and wind speed shown in Figures 6(a) and (c) respectively. Discernible differences do exist for forecast temperature as shown in Figure 6(b). Here we see the smallest forecast error in OSE VER_ONLY up to 700hPa and above this pressure level, OSE BASE+T36 gives the smallest forecast RMS error. A similar trend can be observed for relative humidity scores shown in Figure 6(d). Better representation of the boundary layer temperature and humidity at this time in OSE VER_ONLY may be as a result of frequent inversions resulting in poor modelling of the verification region at lower levels compared with other OSEs. Above 400 hPa, OSE BASE+UPS gives the smallest RMS error for temperature and relative humidity shown in Figures 6(b) and (d). From Figures 6(a)-(d), we see that the utility of OSE VER_ONLY decreases with height.

The performance of targeted radiosonde deployment strategies in ETKF targeting regions

We now evaluate the utility of the ETKF in identifying targeting regions. We compute the mean forecast RMS error for the forecast fields presented in the previous section at the same forecast ranges. We compare the OSEs utilising ETKF targeting guidance, ETKF(24) and ETKF(48), with OSEs based on a 500hPa flow analysis, BASE+UPS and BASE+T36. Both ETKF approaches aim to improve 36 hour forecasts. At T+36 the results in Figures 7(a)-(d) highlight that at least one ETKF-based OSE out-performs those based on the 500hPa flow analysis for each forecast field with the exception of 500 hPa temperature when equal performance is observed. From the mean forecast RMS error results for these fields presented in Figures 7(a)-(d), it is apparent that the relative merit of OSE ETKF(24) over OSE ETKF(48) is dependent on forecast range and forecast field under consideration. Figures 8(a)-(d) show results for each forecast field verifying at different pressure levels for 36-hour forecasts. There is little to distinguish the different approaches for the verification of forecast geopotential height and wind speed shown in Figures 8(a) and (c). For temperature and relative humidity, Figures 8(b) and (d) respectively, we see a reduced forecast RMS error from one of the ETKF OSEs below 500 hPa. Above this height forecast RMS error for both OSEs utilising ETKF targeting guidance are inferior to OSE targeting from 500hPa flow analysis for forecast relative humidity. Results presented in this figure for temperatures above 500 hPa are inconclusive.

AMSU data thinning strategies

We finally present an assessment of the utility of the adaptive thinning of AMSU data used in OSE ETKF(24) compared with the static global thinning controls, OSEs D154 and D40. For each OSE, Figures 9(a)-(d) show the mean forecast RMS error for the verification of the four key forecast fields at different forecast ranges. From

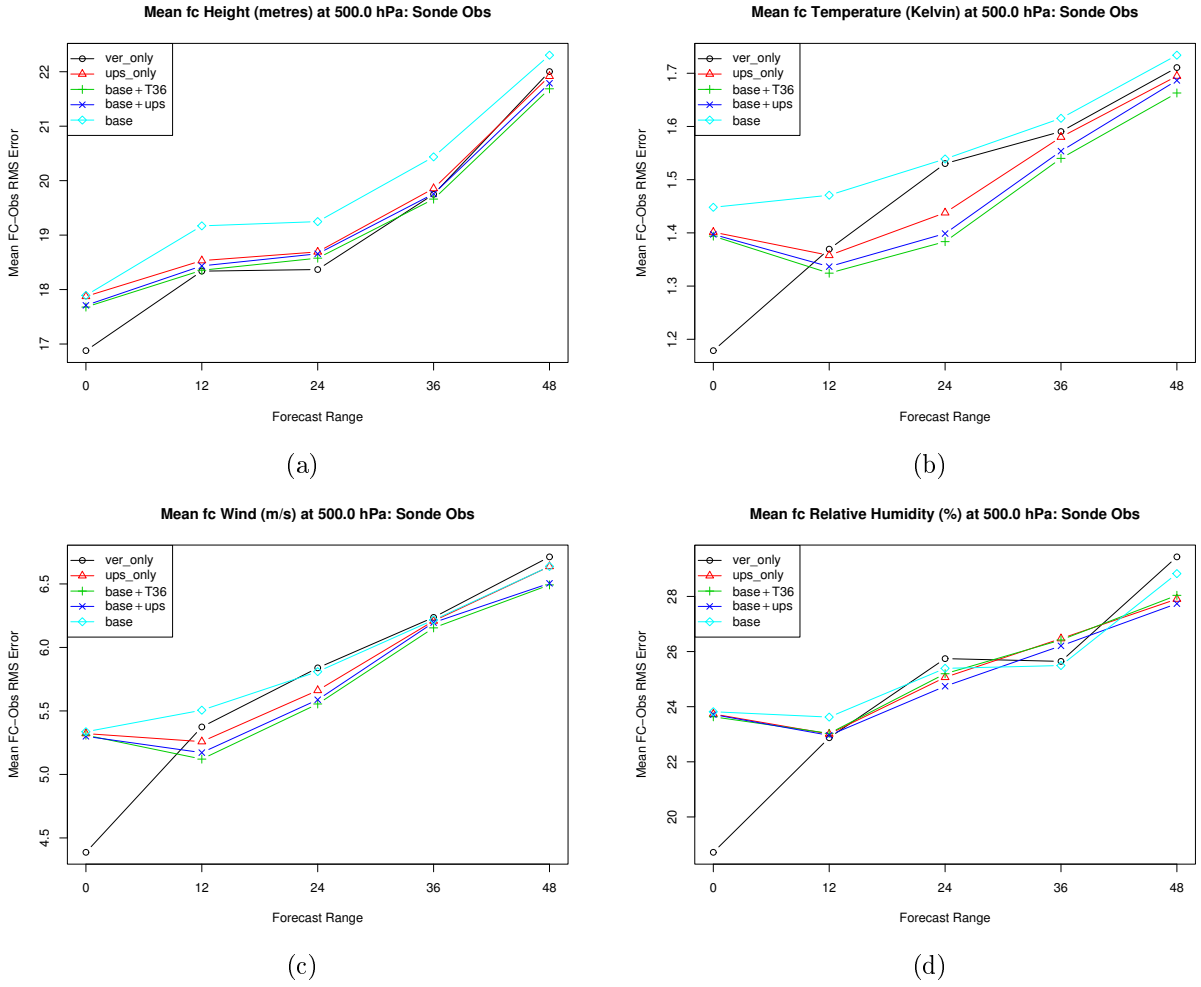


Figure 5: Targeted radiosonde deployment strategies: mean forecast - analysis RMS error verified using radiosonde observations within the verification region for different forecast ranges: (a) geopotential height at 500 hPa; (b) temperature at 500 hPa; (c) wind speed at 500 hPa and (d) relative humidity at 500 hPa.

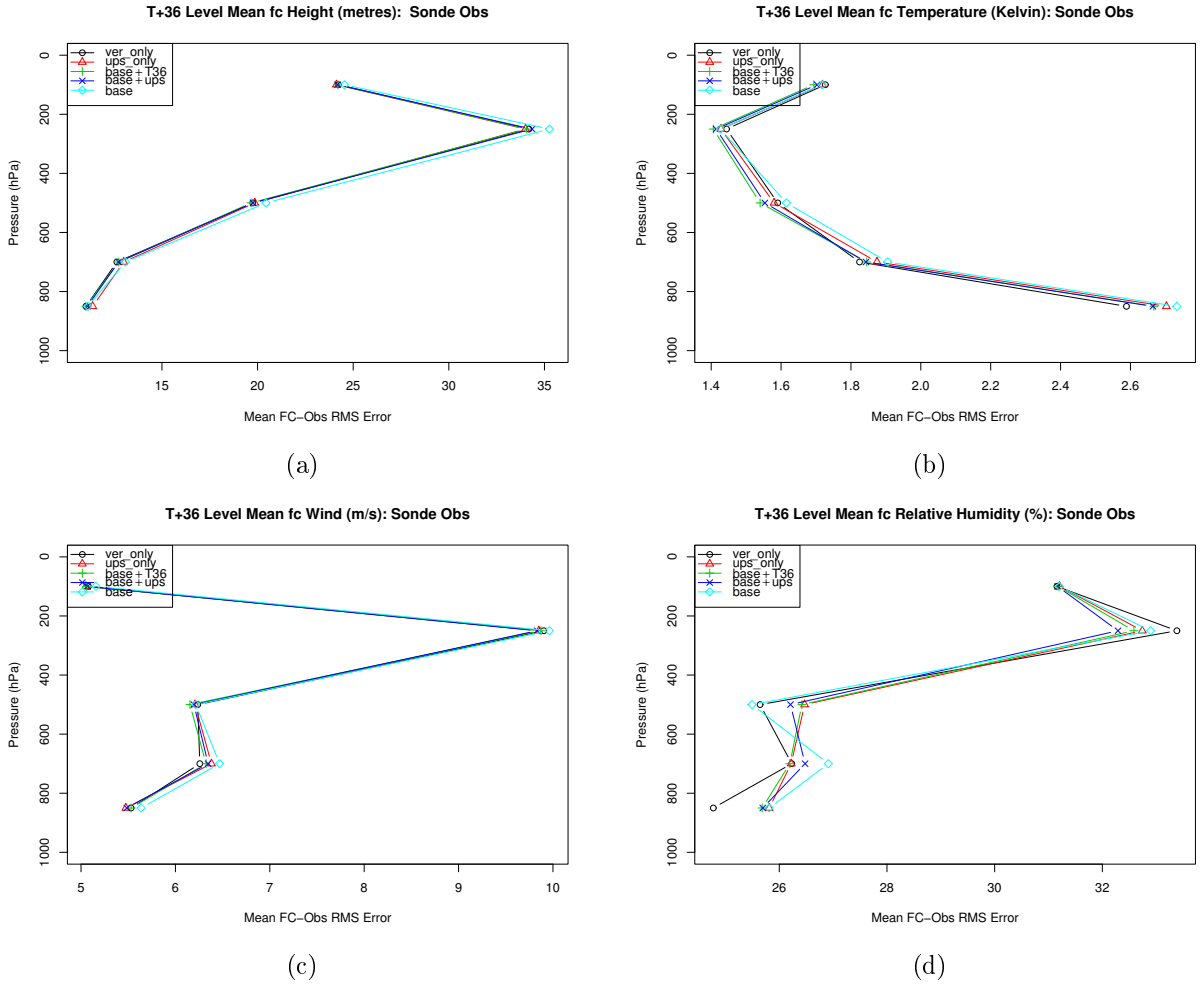


Figure 6: Targeted radiosonde deployment strategies: mean T+36 forecast - observation RMS error verified using radiosonde observations within the verification region at different pressure levels: (a) geopotential height; (b) temperature; (c) wind speed and (d) relative humidity.

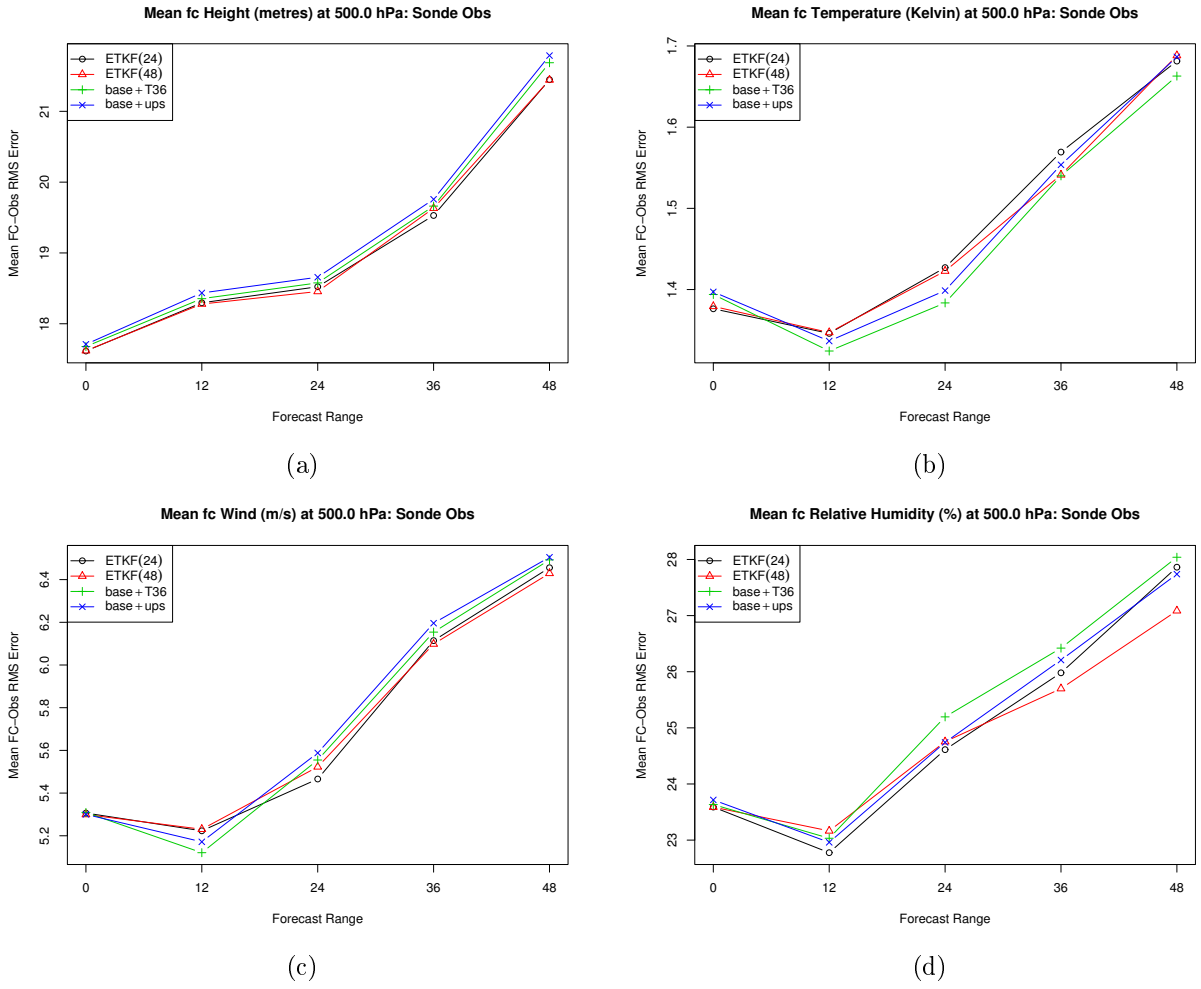
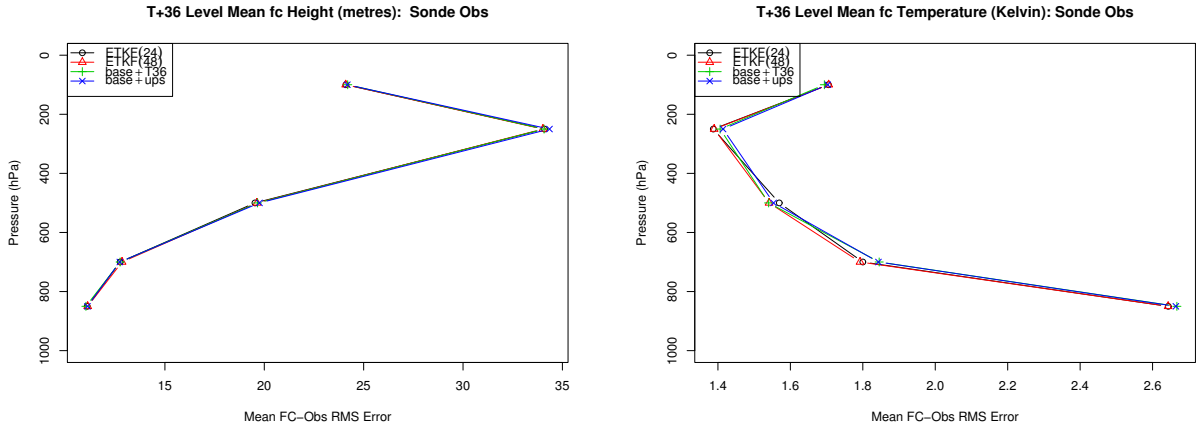
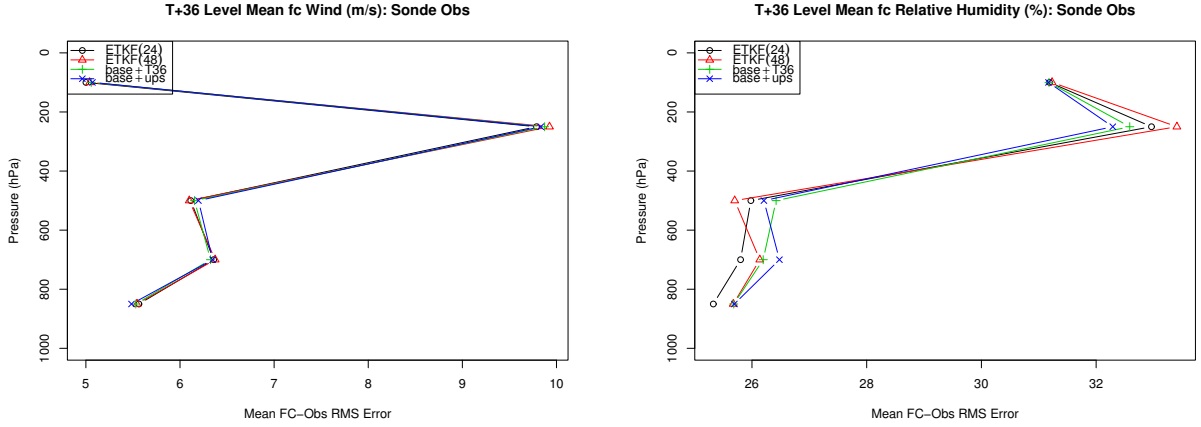


Figure 7: ETKF based sensitive area prediction radiosonde deployment strategies: mean forecast - analysis RMS error verified using radiosonde observations within the verification region for different forecast ranges: (a) geopotential height at 500 hPa; (b) temperature at 500 hPa; (c) wind speed at 500 hPa and (d) relative humidity at 500 hPa.



(a)

(b)



(c)

(d)

Figure 8: ETKF based sensitive area prediction radiosonde deployment strategies: mean T+36 forecast - observation RMS error verified using radiosonde observations within the verification region at different pressure levels: (a) geopotential height; (b) temperature; (c) wind speed and (d) relative humidity.

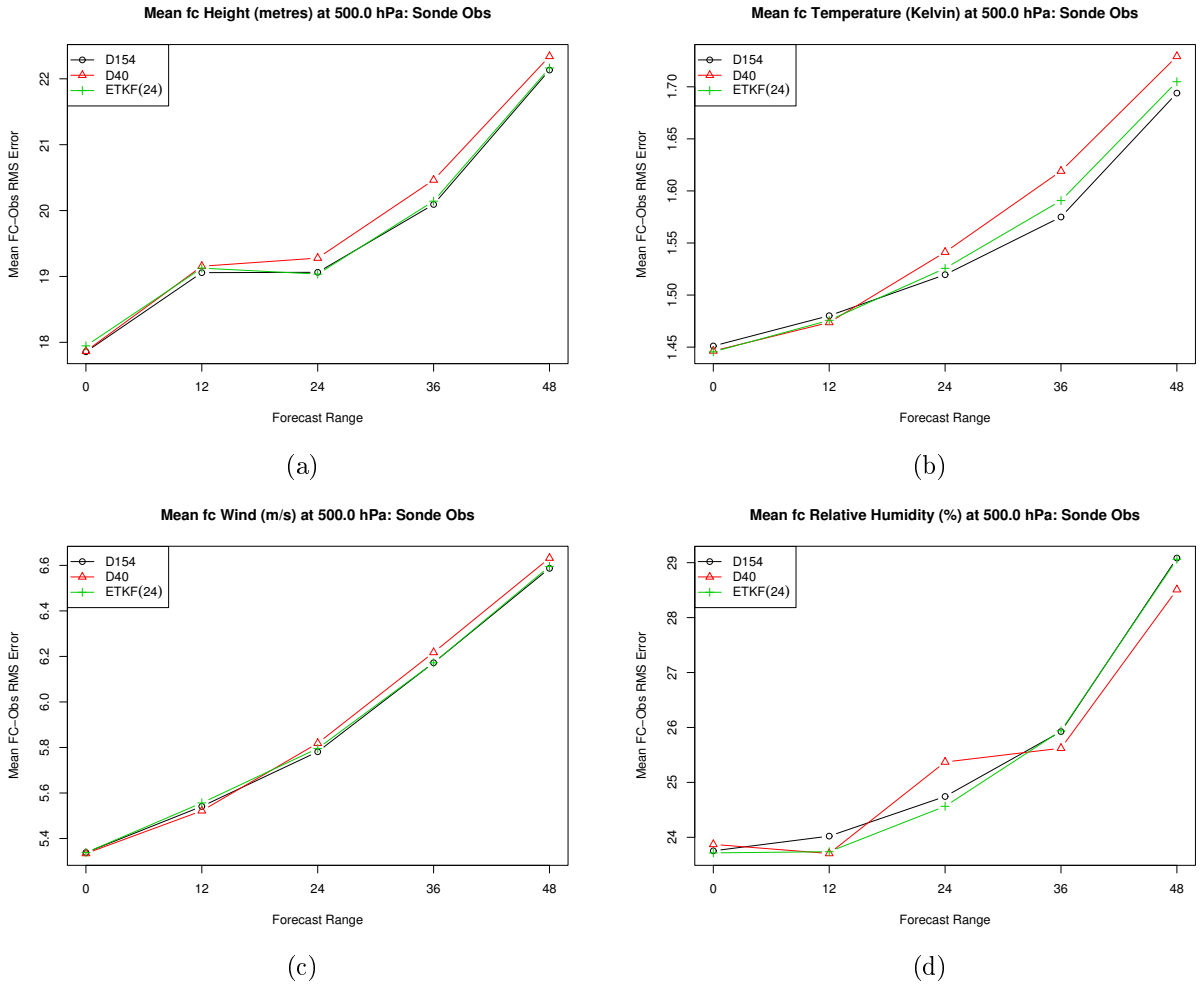
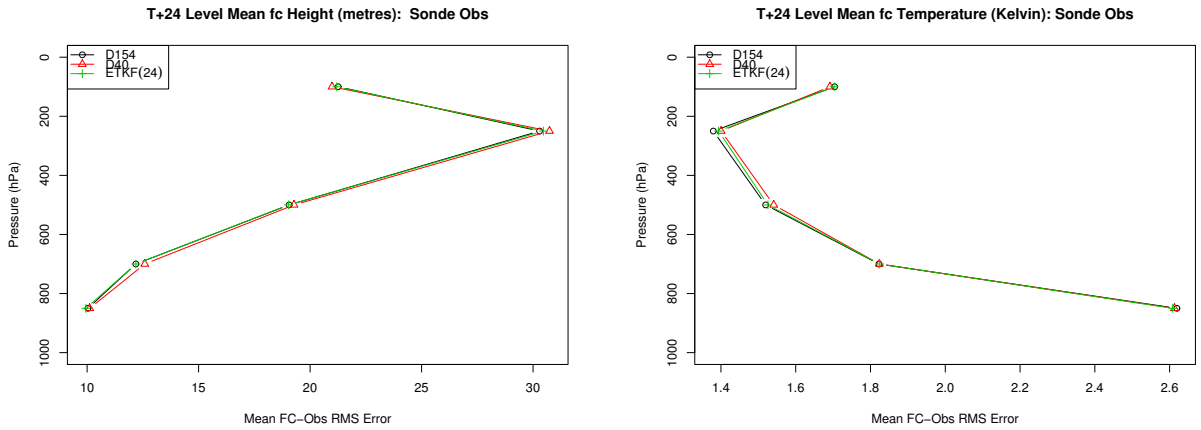


Figure 9: AMSU data thinning strategies: mean forecast - analysis RMS error verified using radiosonde observations within the verification region for different forecast ranges: (a) geopotential height at 500 hPa; (b) temperature at 500 hPa; (c) wind speed at 500 hPa and (d) relative humidity at 500 hPa.

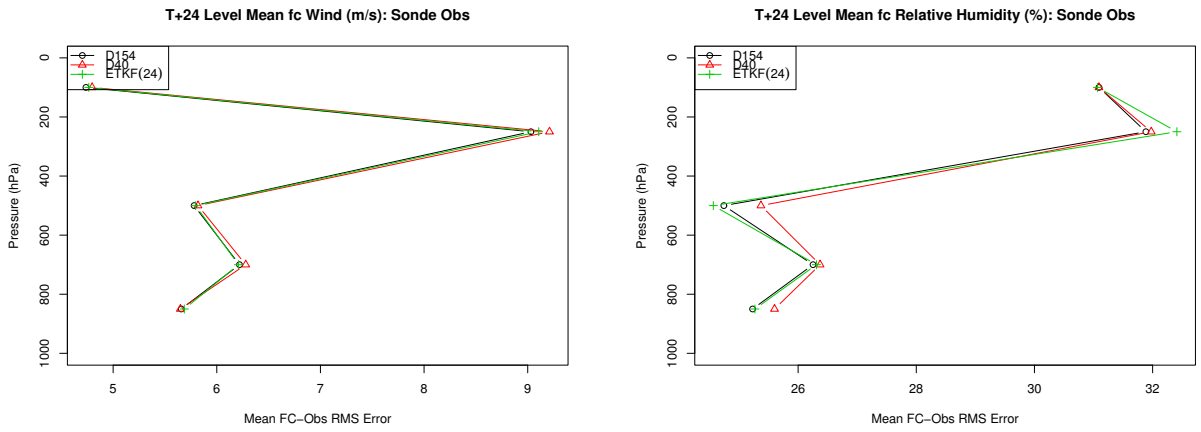
these figures, we can see that for each forecast field at the target 24-hour forecast range, the current operational global setting, OSE D40, is inferior to all other approaches. The horizontal correlation of observation errors in this OSE directly leads to the poorer performance. This result has prompted changes to current Met Office operational regional models. The performance of the AMSU data thinning approach OSE D154 and OSE ETKF(24), appear broadly similar for all fields with the exception of relative humidity shown in Figure 9(d). For this field at 24-hour forecast ranges, OSE ETKF(24) exceeds the performance of D40 and D154. When viewing the results of the mean T+24 forecasts at all pressure levels in Figures 10(a)-(d), there is little discernible difference in impact for geopotential height or mean wind speed at all levels. For T+24 temperature, OSE ETKF(24) RMS forecast error scores are marginally inferior to OSE D154. D154 exhibits a smaller forecast RMS error than D40 at all pressure levels. For relative humidity, a smaller forecast RMS error is seen at 500 hPa for ETKF(24), but a larger forecast error is apparent at 250 hPa. With little noticeable difference overall in the performance of OSEs ETKF(24) and D154, we conclude that adaptive thinning at 40 km within a targeted region reflects the balance of a positive impact of a targeting strategy with the negative impact associated with increased horizontal error correlation.

3.2 Case studies

In this section we present results from the evaluation of each deployment strategy of additional radiosonde observations in the presence of tropical cyclones as they pass through the verification region. We discount hurricanes Ophelia and Otis whose forecast trajectories are shown in Figures 11(a) and (b) as they do not cross the verification region. By contrast, the tracks of hurricanes Katrina and Rita (Figures 11(c) and (d)) do cross the verification



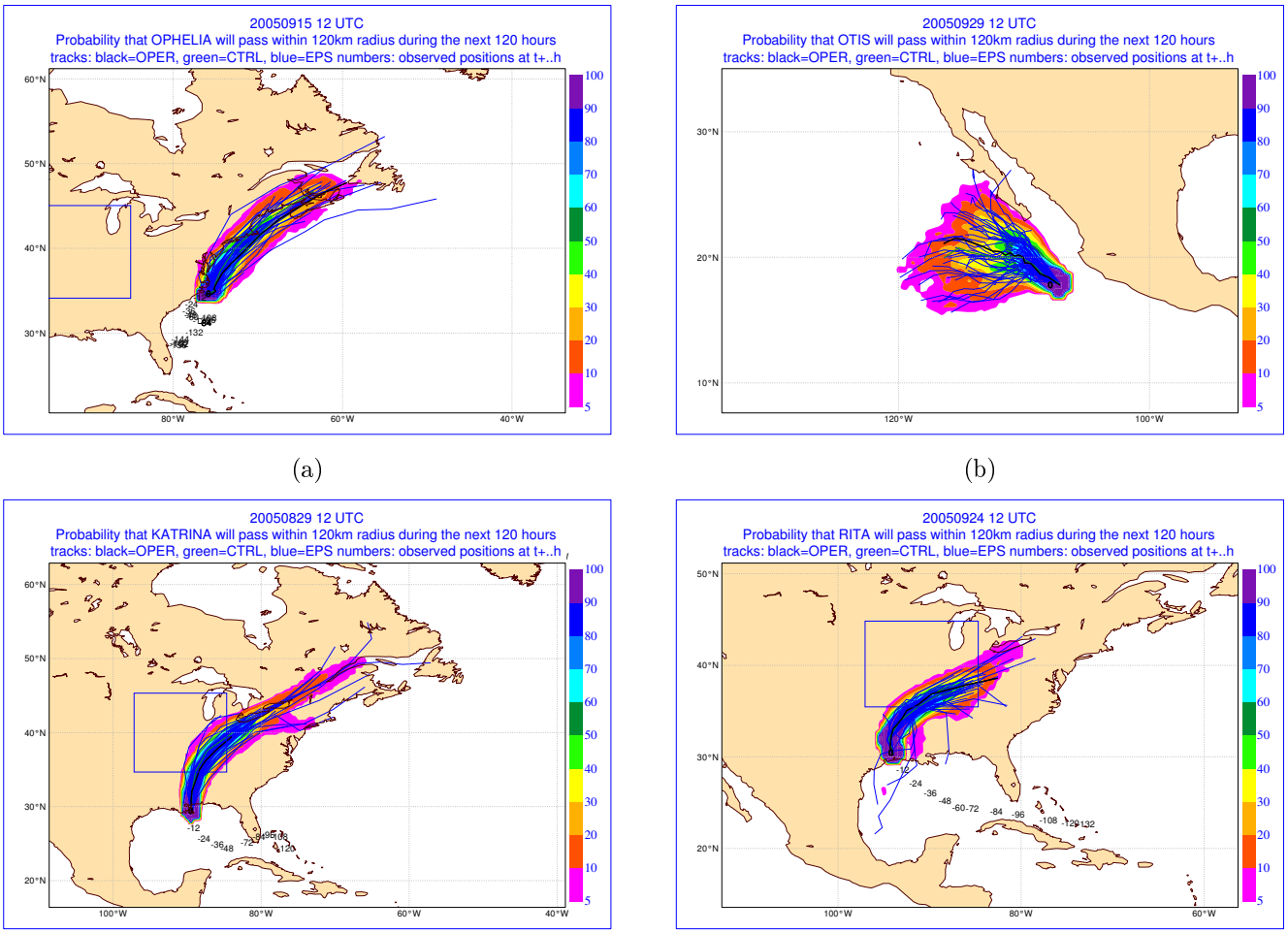
(a) (b)



(c) (d)

Figure 10: AMSU data thinning strategies: mean T+24 forecast - observation RMS error verified using radiosonde observations within the verification region at different pressure levels: (a) geopotential height; (b) temperature; (c) wind speed and (d) relative humidity.

region and are discussed below.



(a) (b)

(c) (d)

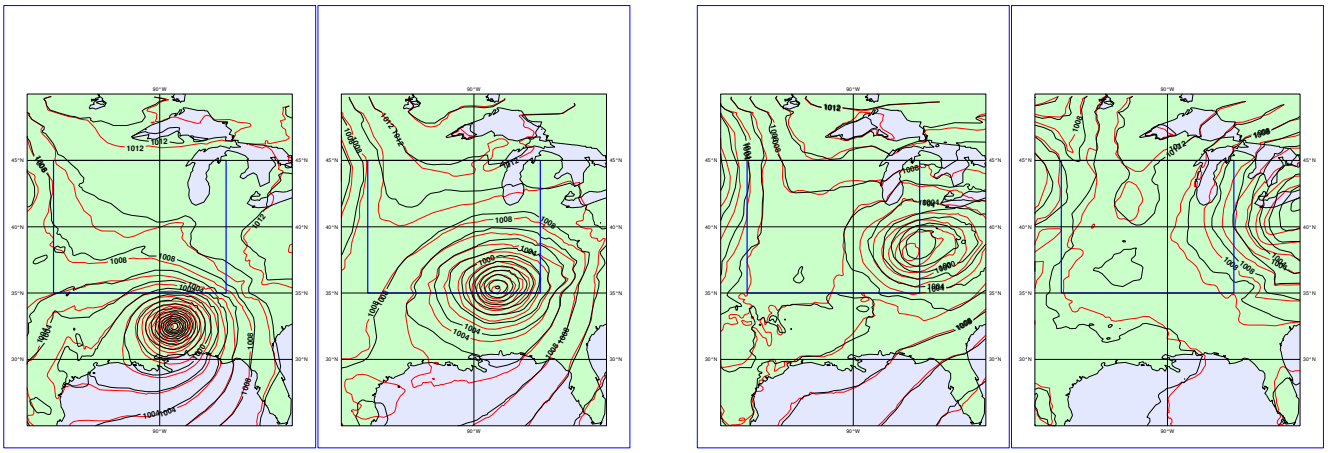
Figure 11: Forecast trajectory probability (coloured contour), verification area (blue box), ECMWF ensemble member trajectory (blue lines), ECMWF operational trajectory (black line) and observed position (digits) for hurricanes: (a) Ophelia on 15 September 2005 12 UTC; (b) Otis on 29 September 2005 12 UTC; (c) Katrina on 29 August 2005 12 UTC and (d) Rita on 24 September 2005 12 UTC.

Hurricane Katrina case

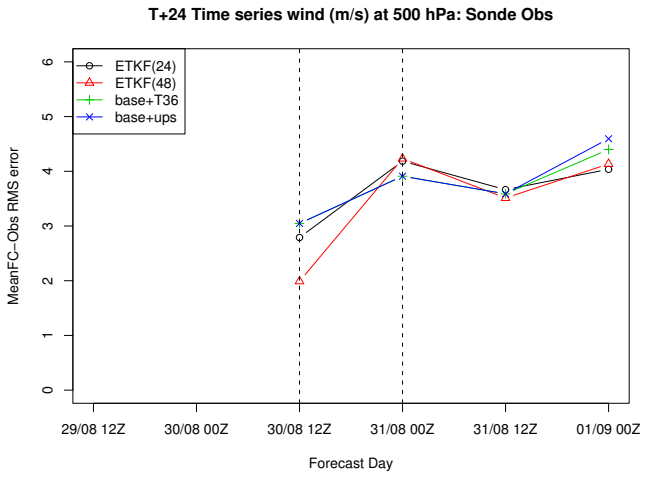
Hurricane Katrina was one of the most expensive and deadly hurricanes on record in the United States [10]. During a 12-hour period between 00 UTC and 12 UTC on 28 August 2005, the maximum sustained winds increased from 100 to 145 kt. Katrina made landfall in the southern states as a category 3 hurricane on August 29 2005. Figures 12(a) and (b) show the analysed synoptic track of Katrina as it enters the verification region on 30 August 2005 12 UTC and leaves, heading north-east 12-hours later on 31 August 2005 00 UTC. In these figures, the synoptic track of Katrina is shown in red using the analysed position from OSE BASE+ALL used in this study and marked as the black track from the equivalent OSE used in D2006. The location of Katrina in both OSEs is broadly similar although it is slightly later in leaving the verification area on 31 August 2005 00 UTC when modelled by OSE BASE+ALL used in this study (Figure 12(b)). These differences are a consequence of the meta-control OSE configuration used in this study differing from that used in D2006. In this study model backgrounds used at the start of each assimilation cycle are identical.

For the period Katrina was active during the trial, Figure 12(c) shows the verification time-series of T+24 forecast RMS error 500 hPa wind speed verified using radiosonde observations deployed within the verification region for each deployment strategy. We are unable to verify at the target 36-hour forecast range because of the proximity to the start of the OSE trial period. In this figure the dates that Katrina enters and leaves the

verification region are shown with vertical dashed lines. The first available 24-hour forecast that can be verified is at the same time Katrina enters the verification region on 30 August 2005 12 UTC. From Figure 12(c) it can be seen that OSE ETKF(48) with a 48-hour lead time, has the smallest forecast RMS error. OSE ETKF(24) has a slightly larger forecast error. The improvement seen in the OSE ETKF(48) utilising the longer lead time have been seen in previous studies [4] and shown to be a result of the spread of ensemble perturbations used by the ETKF thereby better capturing the model uncertainty. Both OSEs targeting in ETKF defined regions out-perform OSEs BASE+UPS and BASE+T36 that deploy additional observations on the basis of a 500 hPa flow analysis when Katrina enters the verification region. At 31 August 2005 00 UTC on leaving the verification region, the converse is true and OSEs BASE+UPS and BASE+T36 appear to out-perform the ETKF OSEs.



(a) (b)



(c)

Figure 12: Analysed position of hurricane Katrina showing synoptic location from BASE+ALL from D2006 (black contour) and BASE+ALL OSE used in this study (red contour). Verification region shown as blue box, analyses times: (a) 30 August 2005 00 UTC & 12 UTC; (b) 31 August 2005 00 UTC & 12 UTC; (c) time-series of T+24 forecast 500 hPa wind speed for the period surrounding Katrina entering and leaving the verification region for each experiment.

Table 3 shows the integrated forecast RMS error for T+24 Mean Sea Level Pressure (MSLP) and 500 hPa temperature (T500) verifying in the verification region on 30 August 12 UTC for each OSE. Verification is undertaken for each OSE with a common verifying analysis (BASE+ALL). We see a larger RMS error for T+24 forecast MSLP in OSE BASE+T36 = 0.758 compared with OSE ETKF(24) = 0.631. This improvement in RMS error is again greater in the ETKF configuration with the longer lead time, OSE ETKF(48) = 0.514 for the same forecast field. For the verification of T+24 T500, the forecast RMS error obtained for OSE BASE+T36 = 0.386 but the shorter lead time ETKF configuration performs more poorly ETKF(24) = 0.414.

OSE	MSLP	T500
BASE+T36	0.758	0.386
ETKF(24)	0.631	0.414
ETKF(48)	0.514	0.378

Table 3: Verification of T+24 forecast Mean Sea Level Pressure (MSLP) and 500 hPa temperature (T500) for hurricane Katrina case verifying 30 August 2005 12 UTC. Figures quoted are forecast RMS error.

Hurricane Rita case

Hurricane Rita was an intense hurricane active between 18 - 26 September 2005. It was notable in that it had the fourth lowest central pressure of 895 hPa ever recorded in the Atlantic basin [10]. During its tropical storm phase (19 - 20 September 2005) rapid intensification occurred such that the cyclone attained a category 5 status with 154 kt winds before weakening to category 3 prior to landfall on 23 September 2005 [13, 10]. In Figures 13(a)-(c), we plot the analysed synoptic track of the hurricane as it makes landfall and crosses the verification region. Analysed locations using BASE+ALL from D2006 are shown in black contours and BASE+ALL used in this study shown as red contours. As with the Katrina case, we see a difference between the synoptic location of Rita. In the BASE+ALL OSE used in D2006, Rita enters the verification region on 25 September 2005 12 UTC. In the BASE+ALL OSE used in this study, it enters the verification region 12 hours later on 26 September 2005 00 UTC as shown in Figure 13(b). The differences in these two runs are highlighted for continuity between this study and D2006 and again can be attributed to the use of a common background in this study.

For the period Rita was active during the trial, Figure 13(d) shows the verification time-series of T+36 forecast RMS error for 500 hPa wind speeds verified using deployed radiosonde observations within the verification region for each strategy. This figure also shows the period of time when Rita enters and leaves the verification region, delineated with a vertical dashed line. Rita enters the verification region on 26 September 2005 00 UTC and during the following 24-hour both ETKF-based strategies exhibit a smaller forecast RMS error for 500 hPa wind speed compared with BASE+UPS and BASE+T36. By contrast, the BASE+UPS case appears to have the largest RMS error for this field at this time.

OSE	MSLP	T500
BASE+T36	0.717	0.850
ETKF(24)	0.760	0.794
ETKF(48)	0.683	0.806

Table 4: Verification of T+36 forecast Mean Sea Level Pressure (MSLP) and 500 hPa temperature (T500) for hurricane Rita case verifying 25 September 2005 12 UTC. Figures quoted are forecast RMS error.

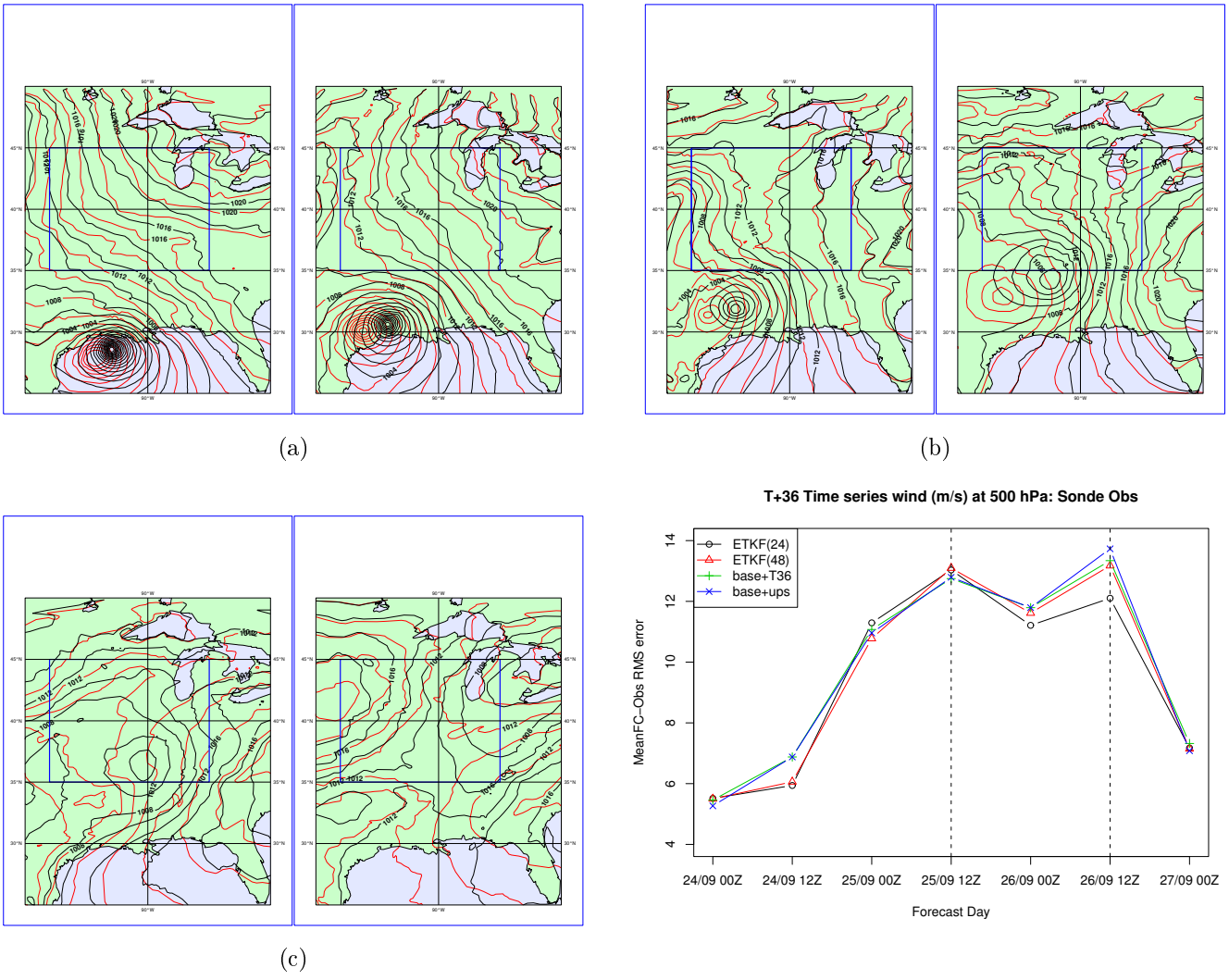


Figure 13: Analysed position of hurricane Rita showing synoptic location from BASE+ALL from D2006 (black contour) and BASE+ALL OSE used in this study (red contour). Verification region shown as blue box, analyses times: (a) 24 September 2005 00 UTC & 12 UTC; (b) 25 September 2005 00 UTC & 12 UTC; (c) 26 September 2005 00 UTC & 12 UTC; (d) time-series of T+36 forecast 500 hPa wind speed for the period surrounding Rita entering and leaving the verification region for each experiment.

Table 4 shows the integrated forecast RMS error for T+36 MSLP and T500 verifying in the verification region on 25 September 12 UTC for each OSE. Verification is undertaken for each OSE with a common verifying analysis (BASE+ALL). From these results we see the forecast RMS error for T+36 MSLP for OSE BASE+T36 = 0.717 and for OSE ETKF(24) = 0.760. An improvement in forecast RMS error is found in the ETKF configuration with the longer lead time with OSE ETKF(48) = 0.683. For T+36 T500, forecast RMS error for OSE BASE+T36 = 0.850 and ETKF(24) = 0.794 for the shorter lead time. For the Rita case, forecast T500 forecast RMS error is larger when targeting using the ETKF with a longer lead time, OSE and ETKF(48) = 0.806.

In the absence of a diagnostic tool to estimate individual observation impact on short range forecast error,

we seek to explain differences in forecast error by qualitatively assessing the location of targeted observation with respect to the underlying 500 hPa flow. For the Katrina case, the underlying mid-level flow is from an area north-west of the verification region shown in Figures 14(a)-(c). Accurate modelling of this air mass (shown as a red dashed ellipse) as it engages with the remnants of Katrina (shown as a black dashed ellipse) entering from the south may be one possible explanation for differences in forecast error observed. For the best performing OSE ETKF(48) shown in Figures 14(c), three observations have been targeted in this area together with four in a region just ahead of Katrina to the south of the verification area. OSEs ETKF(24) and BASE+T36 have fewer targeted observations deployed in the airmass upstream of the verification region shown in Figures 14(b) and (a). Adopting a similar hypothesis that accurate modelling of the air mass upstream of the verification region that engages with the remnants of Rita in Figures 14(d)-(f) (shown as a black dashed ellipse), may also explain differences in forecast error for this case. For OSE ETKF(48) shown in Figure 14(f), no targeted observations were deployed around Rita directly south of the verification region although the upstream air mass was well sampled. By contrast, OSE ETKF(24) shown in Figure 14(e) deploys five observations in the vicinity of Rita and five upstream although none are further west than 110^oW of longitude. The observations deployed in BASE+T36 in Figure 14(d) are broadly similar but include additional observation deployment west than 110^oW of longitude. These differences in observation deployment account for differences in the performance seen for each OSE.

4 Discussion, conclusions and future work

4.1 Discussion and conclusions

In this study we have evaluated different deployment strategies of radiosonde observations and undertaken thinning of AMSU data during one a month long trial period running OSEs based on a regional NWP model. We have evaluated the performance of each OSE by examining a selection of forecast fields at the appropriate forecast ranges and where appropriate evaluating the deployment strategy in the presence of tropical cyclones. From these results we draw the following conclusions:

- The mean forecast RMS errors for fields geopotential height, temperature and wind speed at 500 hPa, are smaller at the target forecast range (T+36) for deployment strategies that include a routine network (OSEs BASE+T36 and BASE+UPS).
- When a routine and targeted radiosonde network is deployed, the T+36 mean forecast RMS errors for forecast fields 500 hPa geopotential height, temperature and wind speed are smaller when targeting additional observations in regions identified with the ETKF (OSEs (ETKF(24) and ETKF(48)) compared with regions based on a 500 hPa flow analysis(OSEs BASE+UPS and BASE+T36).
- We have shown that the global thinning of AMSU satellite data at a distance of one observation every 40 km (OSE D40) leads to an increased forecast error compared with larger distances (OSE D154).
- By considering the hurricane Katrina case study, we have shown that the T+24 forecast RMS vector error for 500 hPa wind speeds[2], is smaller prior to Katrina's entry to the verification region, when targeted observations are deployed in regions identified using the ETKF (OSEs ETKF(24) and ETKF(48)) compared with a those using a 500hPa flow analysis (OSEs BASE+T36 and BASE+UPS). This result is complemented with positive results for 500 hPa temperature and MSLP.
- By considering hurricane Rita at forecast time of entry into the verification region, we also see an improvement in wind speeds by deploying targeted observations using the ETKF (OSEs ETKF(24) and ETKF(48)).

The performance of each OSE can be assessed in terms of an overall forecast skill score. Forecast skill refers to the relative accuracy of a set of forecasts with respect to some set of standard reference forecast [20]. This is compiled from the complete set of forecast fields verified against surface and sonde observations deployed in the verification region at *all* forecast ranges for each cycle of the OSE period. The larger the skill score, the better the performance of the OSE. A visualisation of the computed skill score for each OSE is shown in Figure 15, the best performing OSE lying on the right of the line and poorest performing OSE on the left. From these results we can draw the following conclusions:

- The best performing deployment strategies make use of the routine radiosonde network complimented with a deployment of targeted radiosondes (red and yellow markers).

[2]Termed the 'steering flow', this forecast field is important in determining the track of the tropical cyclone by a forecaster.

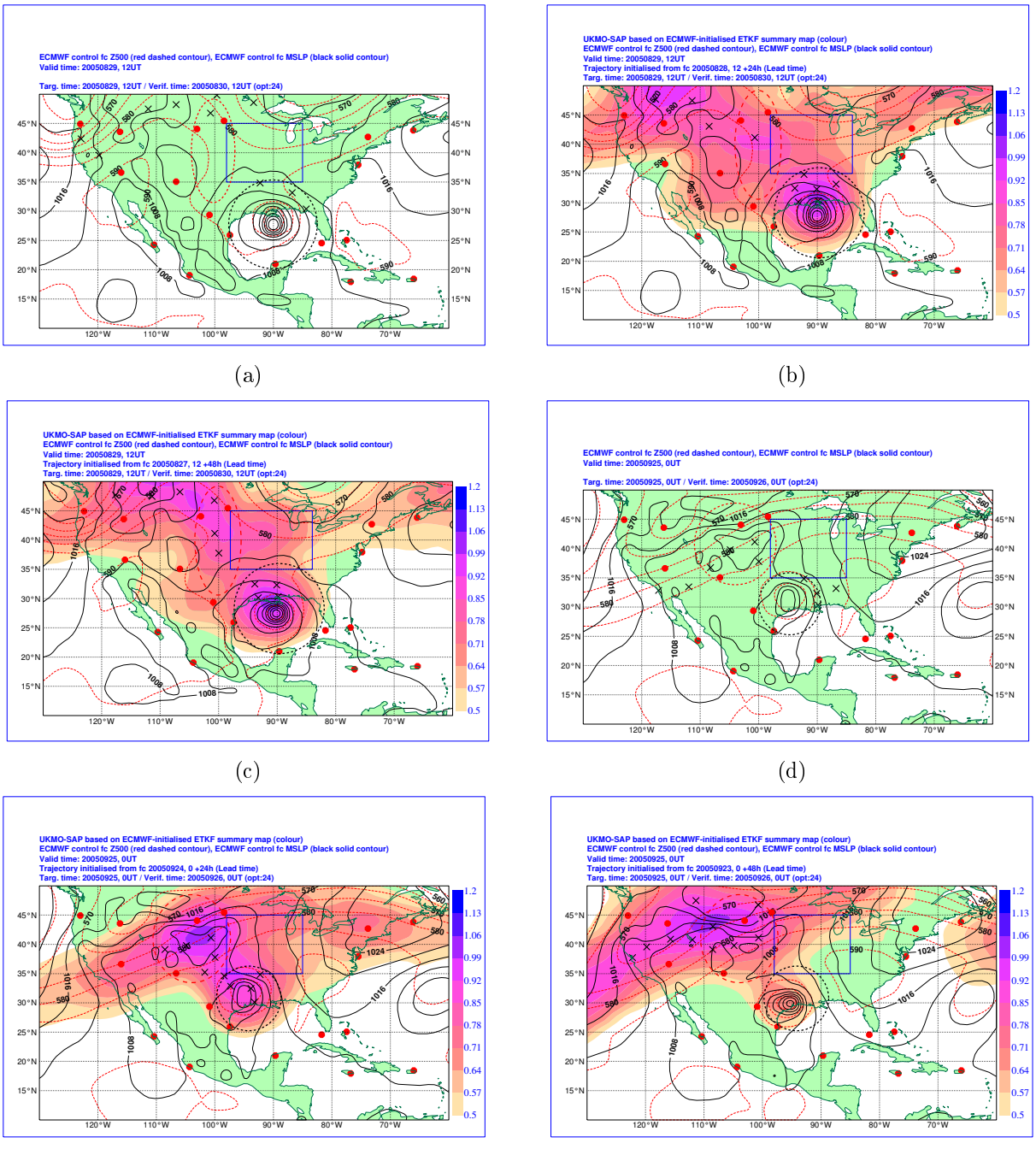


Figure 14: Katrina case ETKF targeting guidance at 29 August 2005 00 UTC targeting guidance for (a) BASE+T36, (b) ETKF(24), (c) ETKF(48). Rita case ETKF targeting guidance at 25 September 2005 00UTC targeting guidance for (d) BASE+T36, (e) ETKF(24) and (f) ETKF(48). The 500 hPa mid-level flow area upstream of the verification region is shown by a red dashed ellipse. Remnants of each tropical cyclone are delineated by a dashed black ellipse.

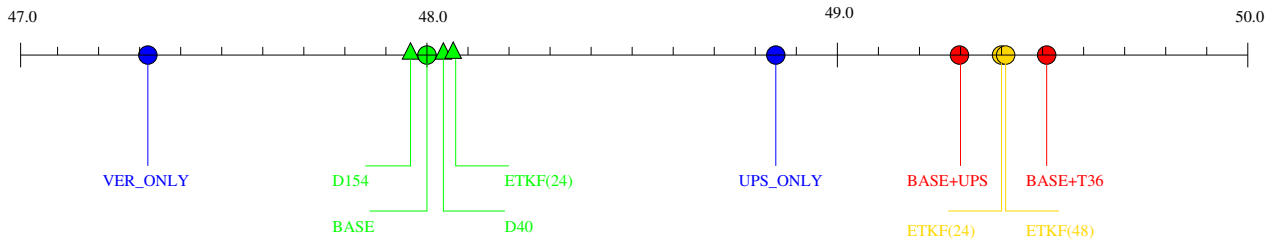


Figure 15: Overall skill scores obtained from each OSE. Evaluation of radiosonde deployments (circular markers) and AMSU data thinning (triangular markers). Colour coding indicates radiosonde deployment strategy grouping: no routine network (blue); routine network only (green); routine and targeted observation network deployed in regions identified using a 500 hPa flow analysis (red); routine and targeted observations deployed in regions identified in ETKF targeting guidance (yellow).

- Strategies utilising a deployment of either the routine network (green markers) or targeted network (blue markers) only, are inferior to the deployment of both routine and targeted radiosonde networks (yellow and red markers). This observation and the previous reinforces the observation made by D2006 that a comprehensive radiosonde network is invaluable in defining the initial conditions for NWP.
- Rather surprisingly, the utility of deploying targeted radiosonde observations in regions identified using a T+36 flow analysis (OSE BASE+T36) is greater than that of a deployment bases on a more up-to-date T+0 500 hPa flow analysis (OSE BASE+UPS). It is important to bear in mind that the overall skill score is computed from RMS scores at *all* forecast ranges and we discuss this result further below.
- Running the ETKF to generate generate targeting guidance using the longer lead time of 48 hours (OSE ETKF(48)), leads to a larger overall skill score than that of the shorter lead time of 24 hours (OSE ETKF(24)).
- The impact, in terms of overall skill score, from the deployment of targeted radiosonde observation within regions identified using a T+36 500 hPa flow analysis is greater than deployment in areas identified by the best performing ETKF configuration.
- The adaptive thinning of AMSU data within areas identified using the ETKF (OSE ETKF(24)) leads to a larger overall skill score than simply thinning globally at one observation every 40 km (OSE D40) or 154 km (OSE D154).
- The impact of attempting to optimise the thinning strategy of AMSU data has a much smaller impact compared with the deployment of an additional targeted radiosonde network.

Although the skill score gives an overall evaluation of each deployment strategy, in an observation targeting paradigm we are primarily interested in improving a forecast at the target range, in this study 36-hours. Table 5 shows the number of improved fields forecast as a percentage at each forecast range for OSEs BASE+T36, ETKF(48) and ETKF(24). These OSEs correspond to those with the best overall skill score shown in Figure 15. We define an improved field as one with a smaller forecast RMS error compared with the corresponding OSE BASE+UPS used as a baseline. We make the following conclusions:

- At the target forecast range of T+36, both ETKF approaches (ETKF(24), ETKF(48)) have a larger percentage of forecast fields improved than OSE BASE+T36.
- Deployment using the T+36 500 hPa flow analysis (OSE BASE+T36) leads to improved shorter-range forecasts 12-24 hours in range compared with other OSEs evaluated.
- Deploying an targeted network using ETKF targeting guidance with a shorter lead time (OSE ETKF(24) leads to a greater percentage of improved fields compared with using a longer lead time (OSE ETKF(48)) at the target forecast range.

Deployment strategy	T+12	T+24	T+36	T+48
BASE+T36	90%	85%	60%	50%
ETKF(48)	75%	75%	70%	55%
ETKF(24)	75%	75%	75%	45%

Table 5: Percentage number of forecast fields with a reduced forecast RMS error compared to corresponding fields in OSE BASE+UPS used as a baseline.

4.2 Future work

We have presented results from the deployment of different configurations of supplemental radiosonde observations and assessed the impact of each by running a series of OSEs. Results presented highlight the utility of a supplemental radiosonde observing network and the need to identify the optimal deployment locations.

Targeted observational data used in this study has been made available by constructing a reduced routine network. In order that algorithms can be fully tested and refined, more targeted radiosonde observation field data is required. By collaborating with forecasting staff, a field campaign would facilitate the construction of a data set to further demonstrate the benefits of observation targeting. Such campaigns are already underway, for example the Greenland Flow Distortion Experiment (GFDex) [18].

The methodology adopted in this study may be used to evaluate other types of targetable observational data. For example future research may include evaluation of the adaptive thinning of aircraft data using the framework described in this study.

With the launch of new satellite instrumentation, an exponential growth in the amount of satellite data available for NWP is expected. Intelligent thinning of satellite data may play an important role in improving target forecasts. Although the impact of this approach described in this report is small compared to that offered by targeted radiosonde observations, a large amount of data is easily manipulated and made available for processing.

Improvements in modelling the physical dynamics of the atmosphere coupled with increases in horizontal and vertical resolution of NWP will undoubtedly lead to improvements in forecasts. The deployment of adaptive observational networks and ingestion into regional NWP models will allow forecasters to better utilise these improvements.

References

[1] C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform Kalman Filter. part I: Theoretical aspects. *MONTHLY WEATHER REVIEW*, 129(3):420–436, 2001.

[2] K. Bovis. Observation targeting using the Met Office Global and regional ensemble prediction system. In *The Second THORPEX International Science Symposium*, Landshut, Germany, December 2006.

[3] K. Bovis. Adaptive thinning of AMSU data in the Met Office Global model. In *The Royal Meteorological Society Conference*, Edinburgh, United Kingdom, September 2007.

[4] K. Bovis. An evaluation of the utility of ETKF sensitive area prediction for observation targeting. Technical Report 513, Met Office Forecasting Research Technical Report, March 2008.

[5] K. Bovis and B. Candy. Adaptive thinning of satellite data using the Ensemble Transform Kalman Filter (ETKF). In *The Second THORPEX International Science Symposium*, Landshut, Germany, December 2006.

[6] R. Buizza, J. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart. The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Quarterly Journal of the Royal Meteorological Society*, 133(624):681–695, 2007.

[7] A. Doerenbecher, M. Leutbecher, and D. S. Richardson. Impact comparison of observation targeting predictions during the (North-) Atlantic TReC 2003. In *The First THORPEX International Science Symposium*, Montreal, Canada, December 2004.

[8] R. Dumelow. Space-terrestrial study at the Met Office - initial report for EUCOS on the summer period experiment. 2007.

[9] R. Dumelow, R. Crocker, G. Dow, K. Bovis, and M. Keil. Impact of radiosonde observations on weather forecasts and artillery trajectories. 2006.

[10] M. Gierach and B. Subrahmanyam. Satellite data analysis of the upper ocean response to hurricanes Katrina and Rita (2005) in the Gulf of Mexico. *IEEE Geoscience and remote sensing letters*, 4(1):132–136, January 2007.

[11] G. Greed. Upgrades to the Crisis Area Mesoscale Model service. Technical Report 464, Met Office Forecasting Research Technical Report, 2005.

[12] A. Joly, K. A. Browning, P. Bessemoulin, J. Cammas, G. Caniaux, J. Chalon, S. A. Clough, R. Dirks, K. A. Emanuel, L. Eymard, R. Gall, T. D. Hewson, P. H. Hildebrand, D. Jorgensen, F. Lalaurette, R. H. Langland, Y. Lemaitre, P. Mascart, J. A. Moore, P. Persson, F. Roux, M. A. Shapiro, C. Snyder, Z. Toth, and R. M. Wakimoto. Overview of the field phase of the fronts and Atlantic storm-track EXperiment (FASTEX) project. *QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY*, 125(561):3131–3163, 1999.

[13] R. Houze Jr., S. Chen, W. Lee, R. Rogers, J. Moore, G. Stossmeister, M. Bell, J. Cetrone, W. Zhao, and S. Brodzik. The hurricane rainband and intensity change experiment. *Bull. American Meterological Society*, 87(11):1503–1521, November 2006.

[14] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[15] S. J. Majumdar, C. H. Bishop, B. J. Etherton, and Z. Toth. Adaptive sampling with the ensemble transform Kalman Filter. part II: Field program implementation. *MONTHLY WEATHER REVIEW*, 130(5):1356–1369, 2002.

[16] D. Mansfield, D. Richardson, and B. Truscott. An overview of the Atlantic THORPEX Regional Campaign. In *The First THORPEX International Science Symposium*, Montreal, Canada, December 2004.

[17] Liu Z. Q. and F. Rabier. The potential of high-density observations for numerical prediction: A study with simulated observations. *Quarterly Journal of the Royal Meteorological Society*, 129(594):3013–3035, 2003.

[18] I. A. Renfrew, G. W. K. Moore, J. E. Kristjánsson, H. Ólafsson, S. L. Gray, G. N. Petersen, K. Bovis, P. Brown, I. Føre, T. Haine, C. Hay, E. A. Irvine, T. Oghuishi, S. Outten, R. S. Pickart, M. Shapiro, D. Sproson, R. Swinbank, A. Wooley, and S. Zhang. The greenland flow distortion experiement. *To be submitted to Bulletin of the American Meterological Society.*

[19] I. Szunyogh, Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson. Propagation of the effect of targeted observations: The 2000 winter storm reconnaissance program. *MONTHLY WEATHER REVIEW*, 130(5):1144–1165, 2002.

[20] Daniel S. Wilks. *Statistical Methods in the Atmospheric Sciences*, volume 59 of *International geophysics series*. 1995.