



Forecast Impact of Targeted Observations: Sensitivity to Observation Error and Proximity to Steep Orography

E. A. IRVINE, S. L. GRAY, AND J. METHVEN

University of Reading, Reading, United Kingdom

I. A. RENFREW

University of East Anglia, Norwich, United Kingdom

(Manuscript received 14 April 2010, in final form 27 July 2010)

ABSTRACT

For a targeted observations case, the dependence of the size of the forecast impact on the targeted dropsonde observation error in the data assimilation is assessed. The targeted observations were made in the lee of Greenland; the dependence of the impact on the proximity of the observations to the Greenland coast is also investigated. Experiments were conducted using the Met Office Unified Model (MetUM), over a limited-area domain at 24-km grid spacing, with a four-dimensional variational data assimilation (4D-Var) scheme. Reducing the operational dropsonde observation errors by one-half increases the maximum forecast improvement from 5% to 7%–10%, measured in terms of total energy. However, the largest impact is seen by replacing two dropsondes on the Greenland coast with two farther from the steep orography; this increases the maximum forecast improvement from 5% to 18% for an 18-h forecast (using operational observation errors). Forecast degradation caused by two dropsonde observations on the Greenland coast is shown to arise from spreading of data by the background errors up the steep slope of Greenland. Removing boundary layer data from these dropsondes reduces the forecast degradation, but it is only a partial solution to this problem. Although only from one case study, these results suggest that observations positioned within a correlation length scale of steep orography may degrade the forecast through the anomalous upslope spreading of analysis increments along terrain-following model levels.

1. Introduction

The aim of making targeted observations is to improve the forecast for a specified region through the addition of information in regions where the forecast is sensitive to initial-condition errors. Over the past 10 years or so field campaigns and idealized modeling studies have tested the idea that adding a small number of profile observations, over a limited area, can have a significant (positive) downstream impact on the forecast. The results of these studies have been mixed, showing that on average targeted observations improve the skill of short-range forecasts but that the impact is a mixture of both forecast improvement and degradation (Langland 2005). This response is characteristic of the forecast impact that

would be expected from assimilating a small number of observations of any type.

The dependence of the impact size on various details of the experimental design has also been assessed. For example Leutbecher et al. (2002) looked at the targeted observation coverage and number using a two-dimensional sampling pattern. Targeted observations were spaced according to the horizontal correlation length scales assumed by the European Centre for Medium-Range Weather Forecasts (ECMWF) four-dimensional variational data assimilation (4D-Var) scheme, and the number of observations and size of target region were varied. Taking targeted observations over a larger area was found to be more effective. The proximity of the targeted observations to regions of dense observation coverage is also important. Bergot (1999) found that targeted observations taken closer to the data-rich U.S. coast had a smaller impact than those taken on trans-Atlantic research flights between Ireland and Canada

Corresponding author address: E. A. Irvine, Dept. of Meteorology, University of Reading, P.O. Box 243, Reading, Berkshire RG6 6BB, United Kingdom.
E-mail: e.a.irvine@reading.ac.uk

where the observations were farther from a data-rich region. Targeted observations taken in regions where the routine observing network (where routine observations are defined as the regular radiosonde, aircraft, station, and satellite observations) is sparse (i.e., has few components) are of greater value than those taken in regions where there are already many routine observations (Buizza et al. 2007). The data assimilation scheme used to assimilate the targeted data can influence the impact from targeted observations; Bergot (2001) and Liu and Zou (2001) found that on average a greater forecast improvement was seen when targeted observations were assimilated with 4D-Var rather than with three-dimensional variational data assimilation (3D-Var). This contrasts with the results of modeling studies by Kelly et al. (2007) who showed that 4D-Var was better able to cope with gaps in the routine observing network, propagating observational information from areas with many observations to areas with fewer observations (i.e., implying that a larger impact from targeted observations might be expected from 3D-Var than from 4D-Var).

3D-Var and 4D-Var data assimilation schemes use least squares approximations that take into account the errors in the background field and observations. The relative size of these errors is important, because this determines the relative weights given to the background field and observations in the analysis. The use of climatological background errors, while practical, can lead to an overweighting of the background relative to observations because the background errors may not be appropriate for the meteorological situation. Reducing the observation error attributed to targeted observations will give the targeted observations greater weighting in the analysis and will increase the size of the analysis increment. Szunyogh et al. (1999) demonstrated with a 3D-Var scheme that this can lead to larger values of forecast improvement by assimilating targeted dropsondes with the observation error covariances reduced to 25% of their original values. Here, we use a 4D-Var scheme in which the initial background errors are climatological and are implicitly evolved during the 6-h window.

The Greenland Flow Distortion Experiment (GFDex) included a field campaign that took place in February and March 2007. The aim was to advance our understanding of the flow deformation by Greenland and its effect on downstream predictability (Renfrew et al. 2008). A specific objective of the campaign was to make targeted observations in the region around southern Greenland and Iceland with the aim of improving the 24–48-h weather forecasts over northern Europe. Targeted observations were made for four different cases, the results of which are presented in Irvine et al. (2009). Using the Met Office operational 4D-Var system, the overall

forecast impact was small and forecasts were improved and degraded by similar magnitudes (up to 5%, measured in terms of total energy).

The study presented here takes one targeting case from Irvine et al. (2009), in which dropsondes were targeted in a total-energy singular-vector sensitive region in the lee of Greenland to improve the 24-h forecast over Scandinavia. Irvine et al. (2009) showed that the targeted observations had a small positive impact (approximately 5% in total energy) on the forecast up to 30 h, after which the forecast was degraded. The impact was caused by the modification of the position of a tropopause fold, which was associated with the development of a polar low at the surface. In the current study the dependence of the impact size on the dropsonde observation errors used in the data assimilation scheme is assessed by rerunning the impact experiment with reduced dropsonde observation errors. Two of the dropsondes were released on the coast of Greenland, where the land rises sharply from the sea toward the Greenland plateau, which at its highest point is 3500 m above sea level. The dependence of the impact size on observation location, and specifically to proximity to this steep orography, is assessed by removing these observations from the dataset and replacing them with observations sited farther from the orography. This is motivated by the fact that observations taken close to orography may measure local flow effects that the model cannot represent and therefore may degrade the forecast.

The structure of the paper is as follows. In section 2 the experimental setup is described, including details of the representation of dropsonde observation errors in the Met Office 4D-Var scheme. Results from the experiments are presented in section 3. In section 3a the forecast impact from assimilating different sets of observations with reduced observation errors is presented. In section 3b the sensitivity of the forecast impact to the proximity of observations to steep orography is assessed. Conclusions are presented in section 4.

2. Method

a. Model setup and error specification

The hindcasts were run using the Met Office Unified Model (MetUM), version 6.1, over a limited-area domain covering the North Atlantic Ocean and Europe (the operational NAE domain). The model is run on a rotated grid with horizontal grid spacing of 24 km (2 times that of the operational model) and 38 vertical levels. This version of the model is nonhydrostatic and uses the new dynamics formulation (Davies et al. 2005) for the dynamical core and a semi-implicit, semi-Lagrangian numerical scheme. The global control forecast

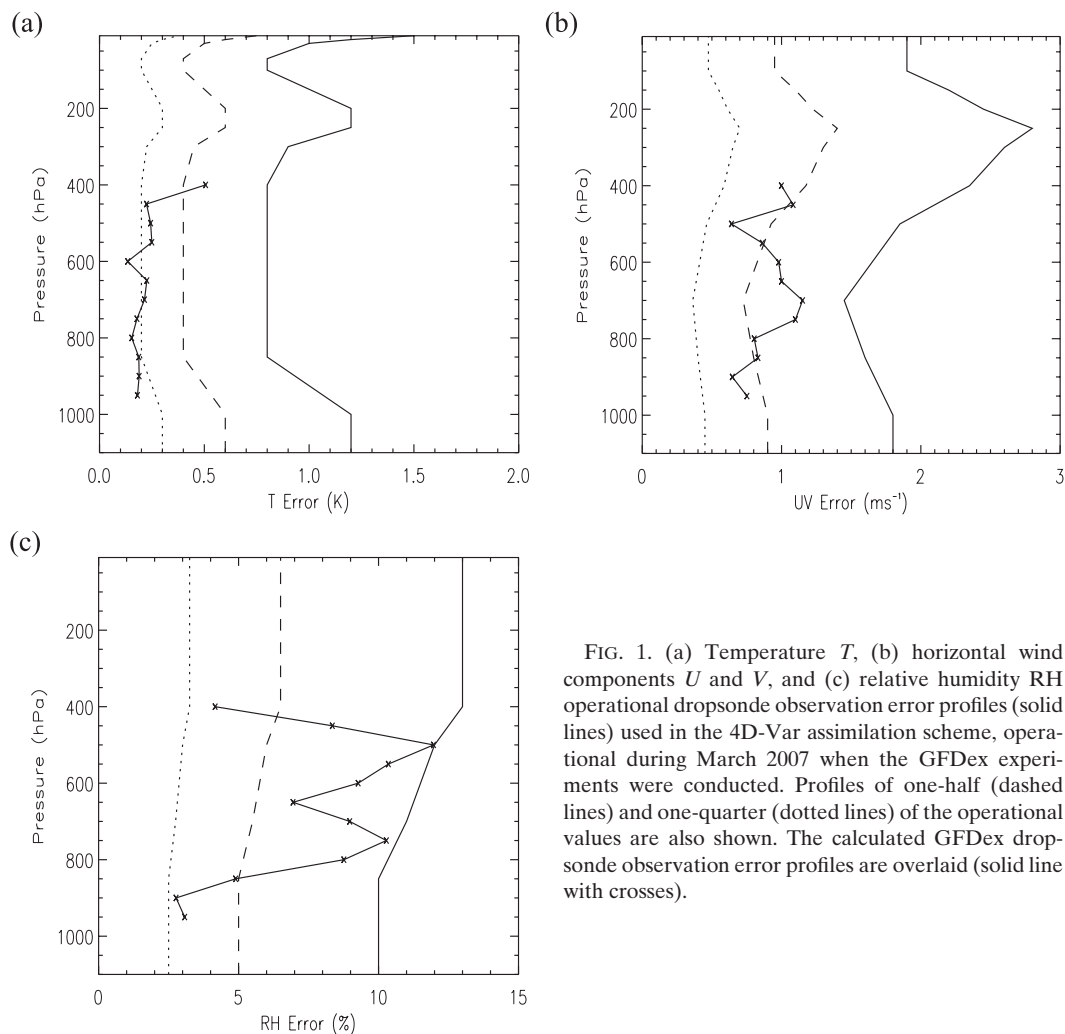


FIG. 1. (a) Temperature T , (b) horizontal wind components U and V , and (c) relative humidity RH operational dropsonde observation error profiles (solid lines) used in the 4D-Var assimilation scheme, operational during March 2007 when the GFDex experiments were conducted. Profiles of one-half (dashed lines) and one-quarter (dotted lines) of the operational values are also shown. The calculated GFDex dropsonde observation error profiles are overlaid (solid line with crosses).

from the Met Office Global and Regional Ensemble Prediction System (MOGREPS) provided lateral boundary conditions for the hindcast.

The Met Office incremental 4D-Var assimilation scheme (Rawlins et al. 2007) was used to assimilate the data. This scheme uses observations taken over a 6-h period centered on the analysis time (1200 UTC for these experiments). The operational background errors were used during the hindcasts; these are initially climatological and have been calculated using the National Meteorological Center method (Parrish and Derber 1992) with some modifications (Ingleby 2001). The operational radiosonde observation error profiles were also used for the dropsonde data. The radiosonde observation error profiles operational at the time of the field campaign (March 2007) are shown in Fig. 1. These were calculated using observation minus background (using a global forecast model) difference statistics over 10 years ago (note that they have recently been revised to slightly

smaller values, after the completion of these experiments). These show a uniform temperature error of 0.8 K through the depth of the troposphere (above the boundary layer), and a wind error that increases with height from 1.5 m s^{-1} at 800 hPa to 2.8 m s^{-1} at the tropopause.

To compare with the operational error profiles, dropsonde observation error profiles were created using some of the GFDex dropsonde (“sonde”) data (a total of 71 profiles, excluding three profiles from malfunctioning sondes) and model data on a 24-km grid, using a similar method to that of the Met Office. Note that errors were only calculated between 400 and 950 hPa where there were greater than 50 data points available. The resulting error profiles (Fig. 1) show more vertical structure than the operational profiles, as a comparatively small number of observations have been averaged over and no smoothing has been applied. The calculated error profiles have values of one-half to one-quarter of those of the operational error profiles, which indicates that the

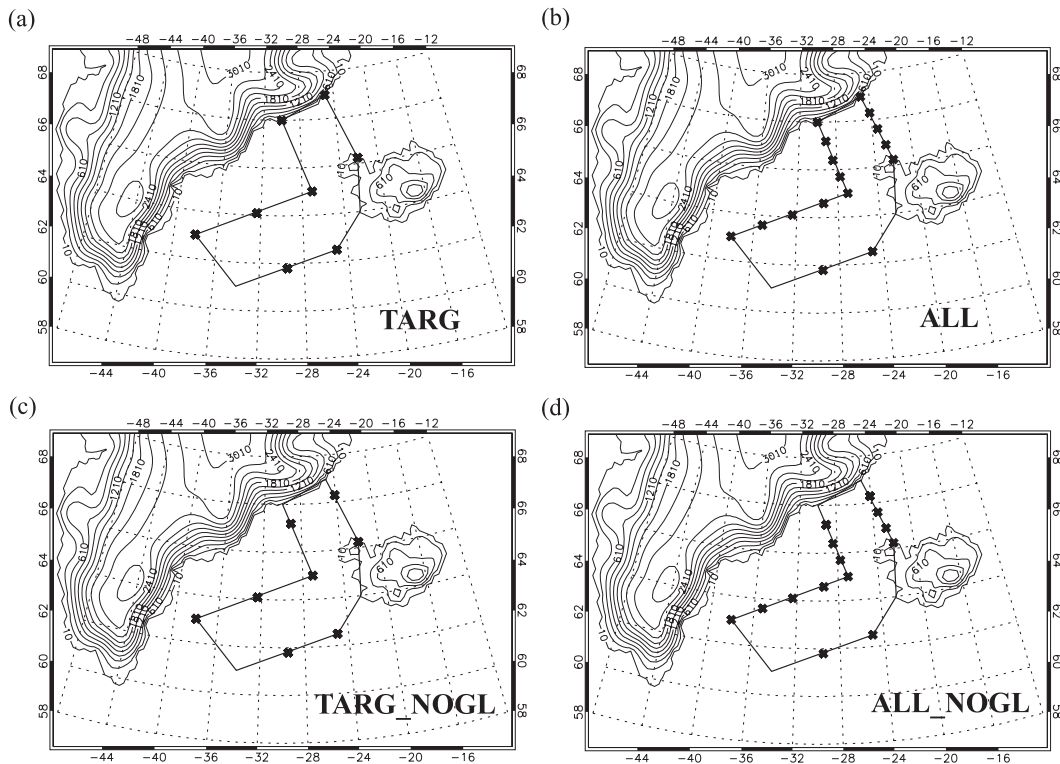


FIG. 2. The configurations of observations for the hindcasts: (a) the original TARG experiment, (b) all sondes (ALL), (c) the TARG set with the two sondes on the Greenland coast replaced by sondes released farther from the coast (TARG_NOGL), and (d) all sondes except the two sondes on the Greenland coast (ALL_NOGL). The flight track for the 1 Mar 2007 flight (solid line) and the model orography (contours) are overlaid. The orography height is contoured every 300 m, with the first contour at 10 m above mean sea level.

operational values were too large. There may be several reasons for this. First, dropsonde data may be more representative of average conditions in a grid box than radiosonde data, because dropsondes do not drift as far as radiosondes (maximum drift for a dropsonde is approximately 10 km as compared with around 200 km for a radiosonde). Second, model resolution has greatly increased (and model formulation improved) in the past 10 years; therefore, a point observation should now be more representative of average conditions in the grid box and observation–model differences should now be smaller. The GFDex errors are also likely negatively biased because of the small sample size and the short period (three weeks) over which the errors were calculated. There is therefore some justification to using reduced dropsonde observation errors in the following experiments.

b. Hindcast experiments

The full forecast cycle (observation processing, data assimilation, and forecast model) was rerun for the period of the field campaign, assimilating only routine observations, to give a set of forecasts that were not

influenced by the dropsonde observations (the dropsonde observations were assimilated into the operational forecasts). These are referred to as the CONTROL forecasts. Hindcasts were then run for the forecast starting at 1200 UTC 1 March 2007, using the background from the CONTROL forecast and assimilating both routine observations and the targeted dropsonde observations.

Four different observation sets were assimilated in the hindcast runs: TARG, ALL, TARG_NOGL, and ALL_NOGL. The TARG observation set is the eight dropsondes (Fig. 2a) that were designated as targeted sondes and are the same dropsondes as were assimilated in the TNOMEM hindcast in Irvine et al. (2009) (note that additional sondes were released for other purposes). The ALL observation set includes all dropsondes released during the flight (Fig. 2b). This increases the spatial resolution of the dropsondes (the separation decreases from a minimum spacing of 220 km down to 85 km), which should better capture gradients in the Denmark Strait but does not change the area sampled by the dropsonde data. The TARG_NOGL set is the same as the TARG set except that the two sondes on the Greenland coast are replaced by two sondes released farther

TABLE 1. Description of hindcast runs. The RI is the relative impact, defined in (2), where the positive values indicate a forecast improvement.

Obs set	Dropsonde obs errors (\times operational)	Max RI (%)	Min RI (%)	Avg RI (%)
TARG	1.0	6.9	-4.6	0.4
	0.5	8.0	-7.4	-0.3
	0.25	9.9	-5.8	0.4
ALL	1.0	15.3	-5.4	3.0
	0.5	17.6	-1.0	4.7
	0.25	18.0	-4.4	4.1
TARG_NOGL	1.0	18.0	0.0	7.2
	0.5	22.6	2.7	10.4
ALL_NOGL	1.0	19.8	1.4	8.5

away from the coast (Fig. 2c). The ALL_NOGL includes all dropsondes except the two on the Greenland coast (Fig. 2d). The ALL_NOGL and TARG_NOGL experiments test whether the two Greenland sondes were beneficial to the forecast quality. They were released close to steep orography and so the profiles may contain features that are not resolved by the model. The observation sets were assimilated with operational [i.e., the values used in the TNOMEM experiment from Irvine et al. (2009)], one-half operational or one-quarter operational dropsonde observation errors; the errors were applied to the model-grid-oriented horizontal wind components U and V , temperature T , and relative humidity RH (and the error profiles are shown in Fig. 1). Nine different hindcasts were run. The combinations of observation sets and dropsonde observation errors are given in Table 1.

c. Verification of targeted forecasts

To assess the forecast impact of modifying the dropsonde observation errors and observation set, the forecasts are compared against a forecast containing no targeted observations (the CONTROL forecast), and both forecasts were verified. We follow previous targeting studies and use analyses rather than observations as the best estimate of the true state. ECMWF analyses on a 25-km grid (T799) are used for this purpose. It is preferable to verify against analyses rather than observations in this case because there are few radiosonde observations in the verification region. ECMWF analyses are used in place of MetUM analyses so that the forecasts are verified using independent analyses; verification against MetUM analyses produces qualitatively similar results (not shown). The forecast error [relative to an analysis (A)] was calculated in terms of the total energy of the difference between the forecast and analysis, TE_{F-A} :

$$TE_{F-A} = \frac{1}{2}(U_{F-A}^2 + V_{F-A}^2) + \frac{1}{2} \frac{c_p}{T_{\text{ref}}}(T_{F-A}^2), \quad (1)$$

where F is either the CONTROL (C) or targeted (Ta) forecast, T_{ref} is a reference temperature of 300 K, and c_p is the specific heat capacity. The total energy was calculated at 850-, 500-, and 250-hPa levels and summed. This was then used to calculate the relative impact RI of the targeted observations, defined as

$$RI = \frac{TE_{C-A} - TE_{Ta-A}}{TE_{C-A}} \times 100, \quad (2)$$

so that $RI > 0$ implies that the targeted sondes have improved the forecast and $RI < 0$ implies that they have degraded the forecast. The multiplication factor of 100 converts RI to a percentage; a positive RI can be interpreted as the percentage reduction in forecast error due to the targeted sondes. The forecasts are verified for a region over Scandinavia (an approximately square region 1000 km on a side, which extends from 54° to 72°N and from 0° to 40°E). This verification region was used in the calculation of the sensitive areas (that were used to determine where to target observations) and was predetermined and fixed for the duration of GFDex.

3. Results

a. Impact of modifying the configuration of observations and observation error specification

The impact of reducing the observation errors in the assimilation can be seen by examining the analysis increments. By expressing the analysis increment $\delta\mathbf{x}$ as the best linear unbiased estimate (Kalnay 2003) it is seen that the size of the analysis increment is proportional to the departure of the observation from the background field $\mathbf{y} - \mathbf{h}(\mathbf{x}^b)$ weighted by the observation and background errors (\mathbf{R} and \mathbf{B}):

$$\delta\mathbf{x} = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}[\mathbf{y} - \mathbf{h}(\mathbf{x}^b)], \quad (3)$$

where \mathbf{y} are the observations, \mathbf{x}^b is the background state, and \mathbf{h} interpolates to observation space (\mathbf{H} is the linearized version). Because \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. The analysis increment due to the targeted dropsondes is shown in Fig. 3 for observation sets TARG and TARG_NOGL and operational and half-operational dropsonde observation errors. The impact of the targeted sondes (TARG) assimilated with operational errors (Fig. 3a) is to strengthen a cyclone in the lee of Greenland (centered

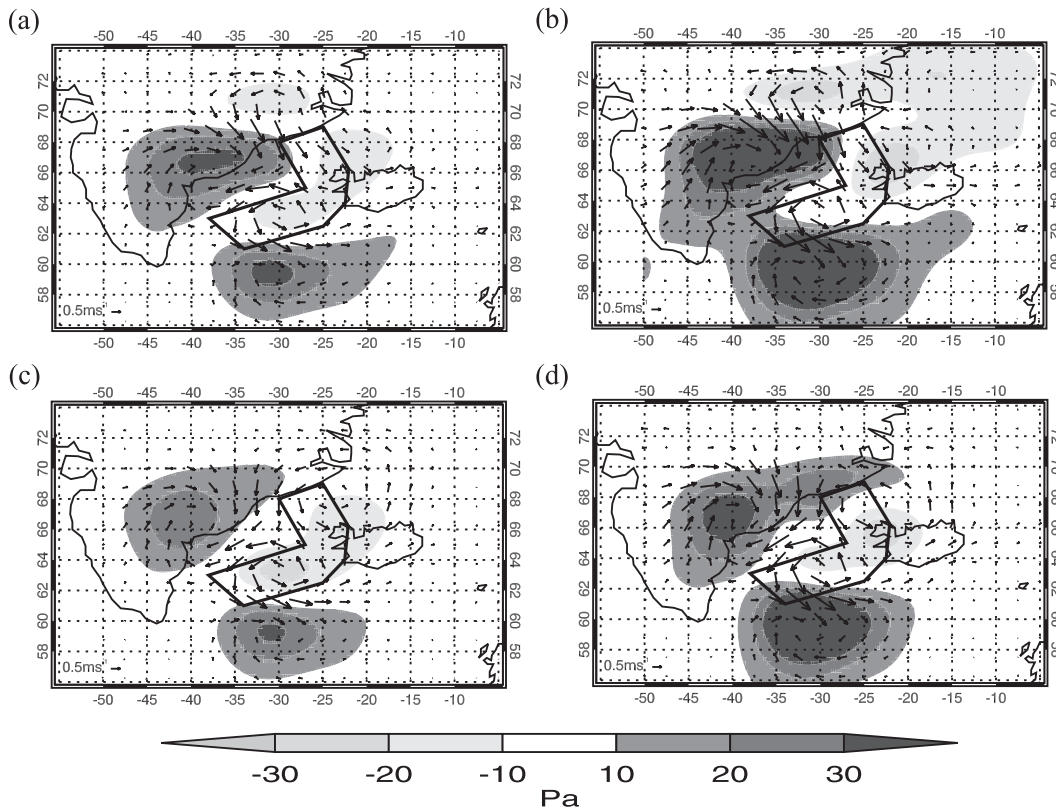


FIG. 3. Targeted minus CONTROL analysis increment in pressure (shaded) and wind strength and direction (arrows) for (a) TARG with operational dropsonde errors, (b) TARG with the operational dropsonde errors halved, (c) TARG_NOGL with operational dropsonde errors, and (d) TARG_NOGL with operational dropsonde errors halved. 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 500 hPa). The flight track is overlaid.

about 63°N, 28°W) through a more negative pressure increment and a stronger cyclonic wind increment in the targeted forecast relative to the CONTROL forecast (see also Irvine et al. 2009).

Increasing the fit of the observations to the analysis by reducing the observation errors (Fig. 3b) increases the magnitude of the analysis increment, as expected. The analysis increments will not be magnified by a constant factor; the magnification factor will vary spatially as it is dependent on the relative size of the observation and background errors [(3)]. Increasing the number of sondes assimilated (from TARG to ALL) has the same effect on the increments as reducing the observation errors (not shown). However, when observation set TARG_NOGL, which does not include sondes on the Greenland coast, is assimilated the pattern of impact is modified (Figs. 3c,d); there is no cyclonic increment introduced to the north of the Greenland sondes (centered about 71°N, 30°W).

Figure 4 shows RI calculated using (2) for all observation sets where the observations were assimilated with operational dropsonde observation errors. For TARG

the RI increases to a maximum of 7% after 18-h of forecast and then decreases so that the forecast is degraded after 30 h. At 0-h forecast time (immediately after the observations have been assimilated) the size of the difference between the forecasts is extremely small; therefore, the differences between the initial RI values for the different forecasts are not significant. Assimilating all observations increases the maximum RI to 15%, but changing the set of observations so that the two sondes on the Greenland coast are not included gives a maximum RI of 18%, implying that these two sondes have a negative impact on the forecast. The maximum, minimum and average RI for each hindcast has also been computed (Table 1), using only the 6–48-h forecast period to exclude the impact at $t + 0$ and after 48 h when the impact from the targeted sondes has moved out of the verification region (not shown). Assimilating the same set of observations with reduced errors increases the maximum RI by 2%–5% but does not necessarily increase the average RI over the forecast (Table 1). The largest forecast impacts are clearly obtained by removing

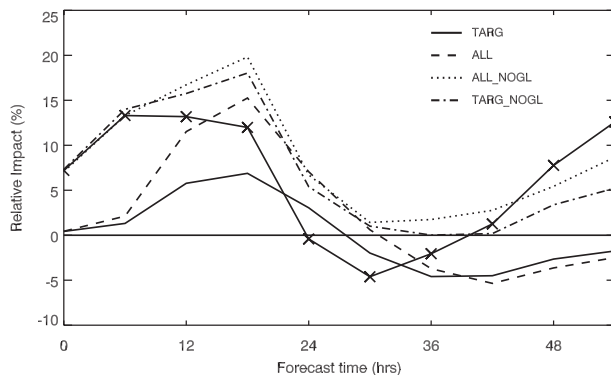


FIG. 4. Relative impact averaged over the Scandinavian verification region, calculated using (2), for the TARG hindcast (solid line), the ALL hindcast (dashed line), the TARG_NOGL hindcast (dashed-dotted line), the ALL_NOGL hindcast (dotted line), and a TARG hindcast that rejected data from the Greenland sondes below 850 hPa (solid line with crosses). All were assimilated with operational dropsonde observation errors.

the Greenland sondes from the dataset, because this not only increases the maximum and average RI but also results in a minimum RI that is above (or equal to) zero (i.e., these sets of sondes do not degrade the forecast for Scandinavia at any time).

These results indicate that the two sondes located on the coast of Greenland degrade the forecast over Scandinavia. To confirm this result, an additional hindcast was run assimilating all sondes except the two sondes on the coast of Greenland (ALL_NOGL; Fig. 2d). This gave a maximum RI in the Scandinavian verification region after 18 h of 20%, as compared with 15% when all observations were assimilated (Fig. 4 and Table 1), and the RI is positive at all forecast times. This proves that for this case dropsondes released on the coast of Greenland have a negative forecast impact over Scandinavia.

b. Cause of forecast deterioration by sondes adjacent to Greenland

In this section the cause of the forecast degradation by the two sondes adjacent to Greenland is determined. Two hypotheses are tested. The first is that two observations on the Greenland coast degrade the forecast because they contain structure that the model is not capable of resolving. The second hypothesis tested is that it is the spreading of observational data from the Greenland sondes that causes the forecast degradation. The first hypothesis is motivated by the effect of Greenland's orography on the local airflow, blocking and distorting it, creating southerly barrier winds, creating tip jets in the lee of Greenland, and altering the temperature profiles by introducing strongly stratified elevated layers

(Petersen et al. 2009). Sondes placed close to any such steep orography may measure these local effects, which would then make the sonde data unrepresentative of the larger area around it. Unrepresentative data can be removed when the observations are processed (before data assimilation), by rejecting data that are significantly different from the background field (assumed to be due to measuring local effects).

Figure 5 shows model profiles from the CONTROL forecast at three locations: at the location of one of the Greenland coastal dropsondes (G), on the Greenland plateau to the northwest of the dropsonde (NW), and in the Denmark Strait to the southeast of the dropsonde (SE). The points are separated by approximately 200 km; this is the separation distance used for the targeted sondes in the GFDex targeting experiments and is approximately the horizontal correlation length scale for temperature assumed by the data assimilation scheme. There are some differences in the near-surface temperature between the profiles, but the wind profiles are similar (Fig. 5b). The largest difference in the profiles is that the NW profile starts at 750 hPa, because it is located 2700 m above sea level on the Greenland plateau. The dropsonde profile at G is also plotted; the dropsonde and model profiles have a similar structure although the modeled temperature is up to 5°C too warm and the modeled wind speed is up to 5 m s⁻¹ too slow throughout the depth of the troposphere. This shows that the dropsonde profile does not contain structure that cannot be resolved by the model (remember we are comparing the dropsonde profile with a model profile that does not contain the observation); therefore, this is not the cause of the forecast degradation.

The second hypothesis is motivated by the large difference in low-level conditions between the coastal sonde and the Greenland plateau, which is within a correlation length scale of the observation and therefore the maximum distance to which the sonde data could be spread. It is clear from Fig. 5 that in particular the lowest part of the dropsonde profile at G is representative of conditions in the Denmark Strait, not of conditions over the Greenland plateau.

To test this hypothesis, the analysis increment resulting from the assimilation of the two sondes on the coast of Greenland was calculated by taking the difference of the analysis increment resulting from assimilating all sondes (ALL) and that from assimilating all except the two Greenland sondes (ALL_NOGL) (both with operational observation errors). Figure 6 shows a cross section through the analysis increment difference for the V-wind component, which cuts through the position of one of these sondes, at 30°W. Neglecting the influence of the other coastal sonde, which is several hundred

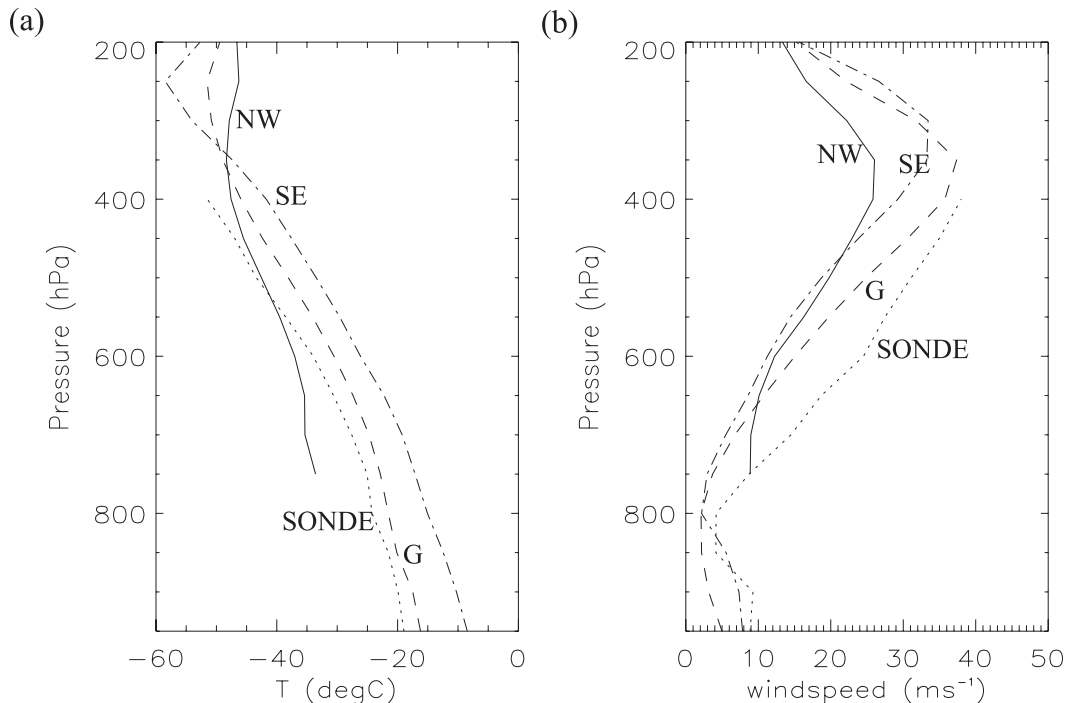


FIG. 5. (a) Temperature and (b) wind speed for model profiles at 1200 UTC from the CONTROL forecast at three locations: centered on the location of one of the coastal dropsondes at 68°N , 30°W (G; dashed line); to the northwest over the Greenland plateau at 70°N , 32°W (NW; solid line); and to the southeast over the Denmark Strait at 66°N , 28°W (SE; dashed-dotted line). The coastal dropsonde observations, interpolated to the same pressure levels as the model data, are also shown (SONDE; dotted line).

kilometers to the northeast of this sonde, we can consider the difference in the analysis increment seen here to be due to the assimilation of this one coastal sonde. Considering the analysis increment to be proportional to the spreading of the observational data by the background errors [(3)], it is clear that the background errors have acted to spread out the information contained in this observation to both the east and west over a distance of several hundred kilometers (the cross section is at 68°N where 10° longitude is approximately 400 km). This behavior is also seen in the U , potential temperature, and specific humidity increments (not shown). The MetUM and 4D-Var systems use a vertical coordinate based on height (rather than pressure), which is terrain following near the ground. This means that the observational information has been spread up along the steeply sloping orography of Greenland, as evidenced by the tilt seen in Fig. 6, 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 utilize observations near steeply sloping orography in a manner that does not result in the degradation of the forecast downstream. Future observations should be sited farther away from steep orography where possible. Feasible solutions for current radiosonde observations sited close to steep orography could be to reject data below the height of the orography, increase the observation error, or decrease the background error of the radiosonde data below the height of the orography (which would reduce the magnitude of the analysis increment). The first of these solutions is the simplest and harshest approach. It has been tested here by rerunning the targeted hindcast (with the TARG observation set and operational observation errors) but excluding data from the two sondes on the Greenland coast below 850 hPa. The choice of the cutoff of 850 hPa was a compromise based on the fact that the model orography reaches 750 hPa only 200 km from the observation but that excluding too much of the profile could itself be detrimental to the forecast because data that is spread in the opposite direction, away from the slope, is likely to be beneficial at all forecast levels. The RI for this case is shown in Fig. 4; the maximum RI is doubled relative to a forecast containing the same

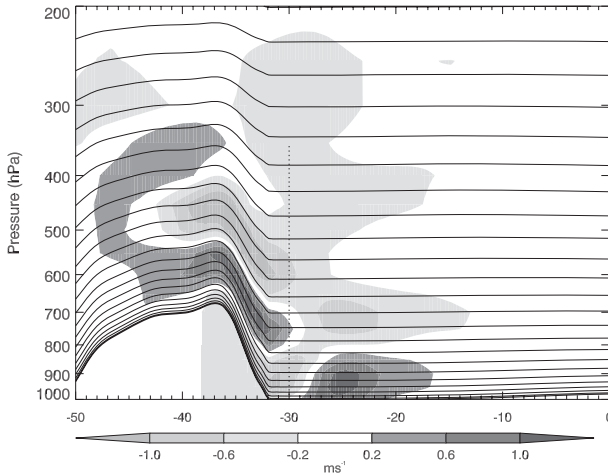


FIG. 6. Cross section through the Denmark Strait at 68°N from 50°W to 0° showing the ALL minus ALL_NOGL analysis increment (shading) for the V -wind component (model-grid oriented) and pressure on model levels (black contours). The model levels are terrain following at the surface, and therefore the lowest model level indicates the orography; the cross section cuts through Greenland on the left. The dropsonde is located at 68°N , 30°W and was released from approximately 350 hPa (dotted line). Note the logarithmic scale on the y axis.

observation set and using the data from the full profile from the Greenland sondes. There is greater degradation after 30 h than for the TARG_NOGL hindcasts; unlike for the TARG_NOGL hindcasts, removal of data below 850 hPa does not remove the cyclonic analysis increment to the north of the Greenland sondes (not shown), which could cause the forecast degradation. This result also gives weight to the conclusion that the spreading of observational data from the lower model levels upslope is the cause of the forecast degradation, because removing this lower part of the profile increases the positive impact this set of sondes has on the forecast.

4. Conclusions

The sensitivity of the forecast improvement to the dropsonde observation error and location of dropsonde observations with relation to the coast of Greenland has been assessed, using a targeting case study from the Greenland Flow Distortion Experiment.

The calculated GFDex dropsonde observation errors were shown to be smaller than the operational observation errors, motivating a reduction in these errors for the GFDex dropsondes in the assimilation. The impact of reducing the dropsonde observation errors during assimilation (so as to increase the fit of the analysis to the dropsonde data) was to increase the magnitude of the associated analysis increment without changing the spatial

structure, as expected. This increases the maximum forecast relative impact by a few percent for this case, independent of the number of observations assimilated. This finding is consistent with Szunyogh et al. (1999) who found an increase in the forecast improvement by reducing observation errors within a 3D-Var system. Assimilating targeted observations with reduced observation errors is a viable method of increasing the impact from these observations, for assimilation systems that do not use fully flow-dependent background errors.

Two sondes close to the steep orography of Greenland caused a degradation of the forecast. Replacing these two sondes with two sondes released farther from the orography increased the maximum improvement from 7% to 18%. The forecast degradation by the two sondes was not caused by the sonde data measuring local flow effects that the model cannot represent, but rather by anomalous spreading of the observational data up the steep slopes of Greenland. A partial solution was tested whereby data from the two sondes below 850 hPa were removed; this increased the relative impact of the set of sondes. This is a practical short-term solution; a better solution would be to improve the representation of the background error covariances to stop the upslope spreading of observational data. Given the way that the error covariances are specified, we are restricted to working in model coordinates and therefore spreading data along coordinate surfaces rather than horizontally. An alternative solution would be to assimilate the sondes with increased observation errors. Priority then should be to reduce the background error length scales by using flow-dependent error covariances that are appropriate to the meteorological situation (i.e., by approximating initial background errors using a large ensemble; e.g., Fisher and Andersson 2001). Obtaining a good estimate of the uncertainty in the analysis would require running an extremely large ensemble, which is currently not feasible. It would, however, be possible to use a hybrid method in which background error estimates from a small (e.g., 24 member) ensemble are blended with climatological estimates; this is currently under development at the Met Office (R. Swinbank 2010, personal communication). The use of fully flow-dependent error covariances and errors that are relevant to the current synoptic situation rather than being climatological should generally result in shorter correlation length scales in this region and would therefore limit the horizontal extent to which the sonde data are spread. This result has potential consequences for the routine observing network where observation platforms such as radiosonde stations are sited within a correlation length scale of high or steep orography (e.g., there are several operational radiosonde stations around the coast of Greenland). These

results are based on a single case study, and therefore further work is required to assess the generality of this result and to determine the extent to which this issue affects operational weather forecasts.

Acknowledgments. We thank Richard Swinbank for useful discussions during the study. We acknowledge technical support from the National Centre for Atmospheric Science Computational Modelling Support team. We thank the Met Office for use of the MetUM and Richard Dumelow and Keir Bovis for providing technical help with the Met Office forecast suite. The study was supported by Natural Environment Research Council (NERC) Grant GFDex NE/C003365/1.

REFERENCES

- Bergot, T., 1999: Adaptive observations during FASTEX: A systematic survey of upstream flights. *Quart. J. Roy. Meteor. Soc.*, **125**, 3271–3298.
- , 2001: Influence of the assimilation scheme on the efficiency of adaptive observations. *Quart. J. Roy. Meteor. Soc.*, **127**, 635–660.
- Buizza, R., C. Cardinali, G. Kelly, and J. Thepaut, 2007: The value of targeted observations. II: The value of observations taken in singular-vectors-based target areas. *Quart. J. Roy. Meteor. Soc.*, **133**, 1817–1832.
- 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. *Quart. J. Roy. Meteor. Soc.*, **131**, 1759–1782.
- Fisher, M., and E. Andersson, 2001: Developments in 4D-Var and Kalman filtering. ECMWF Tech. Memo. 347, 36 pp.
- Ingleby, N. B., 2001: The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **127**, 209–231.
- 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. *Quart. J. Roy. Meteor. Soc.*, **135**, 2012–2029.
- Kalnay, E., Ed., 2003: *Atmospheric Modelling, Data Assimilation, and Predictability*. Cambridge University Press, 364 pp.
- Kelly, G., J. Thepaut, R. Buizza, and C. Cardinali, 2007: The value of targeted observations. I: Data denial experiments for the Atlantic and the Pacific. *Quart. J. Roy. Meteor. Soc.*, **133**, 1803–1815.
- Langland, R. H., 2005: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.*, **131**, 3409–3425.
- Leutbecher, M., J. Barkmeijer, T. N. Palmer, and A. J. Thorpe, 2002: Potential improvement to forecasts of two severe storms using targeted observations. *Quart. J. Roy. Meteor. Soc.*, **128**, 1641–1670.
- Liu, H., and X. Zou, 2001: The impact of NORPEX targeted dropsondes on the analysis and 2–3 day forecasts of a land-falling Pacific winter storm using NCEP 3DVAR and 4DVAR systems. *Mon. Wea. Rev.*, **129**, 1987–2004.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis scheme. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Petersen, G. N., I. A. Renfrew, and G. W. K. Moore, 2009: An overview of barrier winds off southeastern Greenland during GFDex. *Quart. J. Roy. Meteor. Soc.*, **135**, 1950–1967.
- 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. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362.
- Renfrew, I. A., and Coauthors, 2008: The Greenland Flow Distortion Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 1307–1324.
- 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. *Quart. J. Roy. Meteor. Soc.*, **125**, 3189–3217.