Met Office

Diagnosis of exaggerated impacts in adjoint-based sensitivity studies

Forecast Research
Technical Report no.564

March 2012

Sangwon Joo, Andrew Lorenc,
and Richard Marriott

Contents

Abstract..2

1. Introduction 2

2. Equations..3

3. Investigation of Exaggerated Impacts 4

4. Summary and future work...12

5. References...13

Abstract

Abnormally large observation impacts which are occasionally found in Met Office adjoint-based observation sensitivity studies are investigated. These few large impacts can account for more than 20% of the total global impact. It was found that this is due to problems with computational stability of the adjoint model's time-integration scheme and that the large impacts can be removed by increasing the temporal resolution of the adjoint model. To avoid misleading interpretation of observation impact results, a simple way of monitoring for exaggerated impacts is proposed. We believe that a modified version of PF and adjoint model system will alleviate the instability in future.

1. Introduction

The adjoint-based observation impact method (Baker and Daley 2000) measures reductions in a quadratic measure of forecast error from one forecast run to the next. It uses a linear approximation to attribute the reduction in error (known as the impact) to each observed datum in the assimilation cycle. A good estimate of the impact due to each observing system can be calculated by summing over many observations and runs.

The method is used in several NWP centres – this note is based on the Met Office system documented by Lorenc and Marriott (2012), which gives references to other systems. The system estimates a forecast impact for each piece of observational information assimilated. All impacts are produced simultaneously and can be easily aggregated. This capability is extremely useful in evaluating the impact of satellite observations, which consist of many sub-types, to the forecasts of a state-of-art global NWP system assimilating more than 10^6 observations each forecast cycle.

Occasionally the method can give impact values for single observations which are very large (Joo 2011); so large that the relative impacts of observation systems can be affected even when results are accumulated over several runs. These large impacts will hereafter be known as 'exaggerated' observation impacts. The purpose of this note is to discuss possible causes of exaggerated observation impacts and to propose methods for diagnosing their presence in results.

This note briefly describes the adjoint-based observation impact method implemented at the Met Office (section 2) and shows the results of an investigation into the causes of exaggerated observation impacts in a particular case (section 3). There is further discussion in section 4 and a summary in section 5.

2. Equations

We follow the notation of Lorenc and Marriott (2012) and summarise the linearised equations used to calculate the adjoint sensitivities. The NWP process starts from a 4D-Var data assimilation which calculates an analysis increment, $\delta\mathbf{w}_0$, from a set of observation innovations ($\mathbf{y}^o - H\left(\mathbf{x}^b\right)$) and the background forecast state, $\mathbf{x}^b$. The matrix of weights given to each observation at each gridpoint is traditionally called the Kalman Gain matrix, **K**. (4D-Var is efficient because it avoids calculating **K** explicitly.)

$$\delta\mathbf{w}_0 = \mathbf{K}(\mathbf{y}^o - H(\mathbf{x}^b)) \qquad (1)$$

This analysis increment is then added to the background and a new forecast run out to time *t*. By comparison with a forecast, also valid at time *t*, but initiated from the previous analysis (i.e. an extended version of the forecast which provided the background for the current analysis cycle) we can calculate the change in the forecast state due to the assimilation of observations. This can be approximated by

$$\delta\mathbf{w}_t = \mathbf{M}_{0,t}\delta\mathbf{w}_0 \, . \qquad (2)$$

The change in error due to the current analysis step is given by

$$J = \left(\delta\mathbf{w}_t^{fa} + \delta\mathbf{w}_t^{fb}\right)^T \mathbf{C}\,\delta\mathbf{w}_t \qquad (3)$$

where $\delta\mathbf{w}_t^{fa}$ and $\delta\mathbf{w}_t^{fb}$ are differences from the verifying analysis of the forecasts from analysis and background, and the inner-product matrix, **C**, gives a total energy norm. Partial differentiation with respect to $\delta\mathbf{w}_t$ whilst holding $\delta\mathbf{w}_t^{fa} + \delta\mathbf{w}_t^{fb}$ constant gives the adjoint sensitivity with respect to the forecast difference:

$$\left(\frac{\partial J}{\partial \delta\mathbf{w}_t}\right)^T_{\delta\mathbf{w}_t^{fa}+\delta\mathbf{w}_t^{fb}} = \delta\hat{\mathbf{w}}_t = \mathbf{C}(\delta\mathbf{w}_t^{fa} + \delta\mathbf{w}_t^{fb}) \qquad (4)$$

where the ^ symbol conventionally denotes an adjoint vector of sensitivities to the corresponding vector in the forward calculation, in our case always using an identity matrix inner product[1]. (This is equivalent to taking the finite gradient of J over the interval $\delta\mathbf{w}_t$.) The inner-product of (4) with the vector of forecast differences gives an exact expression for the total forecast impact:

$$J = \left(\delta\hat{\mathbf{w}}_t\right)^T \delta\mathbf{w}_t \,. \qquad (5)$$

We can express the forecast differences in terms of the observation innovations by substituting (1) into (2):

$$\delta\mathbf{w}_t = \mathbf{M}_{0,t}\mathbf{K}\left(\mathbf{y}^o - H\left(\mathbf{x}^b\right)\right) \,. \qquad (6)$$

If we then substitute (6) into (5) and rearrange, we obtain a linear approximation to the total observation impact in terms of observation innovations.

$$J = \left(\mathbf{K}^T\mathbf{M}_{0,t}^T\delta\hat{\mathbf{w}}_t\right)^T\left(\mathbf{y}^o - H(\mathbf{x}^b)\right) = \left(\delta\hat{\mathbf{y}}^o\right)^T\left(\mathbf{y}^o - H(\mathbf{x}^b)\right) \qquad (7)$$

where $\delta\hat{\mathbf{y}}^o$ is the vector of adjoint sensitivities with respect to observation innovations. In order to analyse individual contributions to the total impact from within the observation-set we evaluate (7) for the k^{th} observation, (8).

$$J_k = \left(\delta\hat{\mathbf{y}}^o\right)_k^T\left(\mathbf{y}^o - H(\mathbf{x}^b)\right)_k \,. \qquad (8)$$

Substituting (6) back into (8) and rearranging gives

$$J_k = \delta\hat{\mathbf{w}}_0^T\mathbf{K}_k\left(\mathbf{y}^o - H(\mathbf{x}^b)\right)_k \,. \qquad (9)$$

This expression helps to understand the character of the observation impacts. I.e. they are equivalent to the inner-product of the analysis sensitivity vector, $\delta\hat{\mathbf{w}}_0$, with the vector of analysis weights for the k^{th} observation, $\mathbf{K}_k$, multiplied by the innovation for the k^{th} observation.

3. Investigation of Exaggerated Impacts

We investigate exaggerated observation impacts produced by the Met Office adjoint sensitivity system. The nonlinear model is the Unified Model (UM) run at a horizontal resolution of N320 with 70 vertical levels. The adjoint model is the adjoint of the UM's Perturbation Forecast (PF) model and is run at a horizontal resolution of N108. Observation impacts on a global 24-hour forecast dry energy norm initiated from the

[1] Since derivatives such as $\frac{\partial J}{\partial \delta\mathbf{w}_t}$ are conventionally treated as row vectors, while adjoint model states are conventionally column vectors, we transpose one before defining them to be equivalent.

18UTC 5 June 2010 analysis are shown in Figure 1. Total observation impacts for 2.5 by 2.5 degree grid boxes are shown. The observation impacts in each grid box are expressed as a percentage of total global forecast impact. The observation impact at the coast of Antarctica accounts for more than 10% of the total global impact. This is obviously too large to be interpreted as the impact of such a small set of observations.

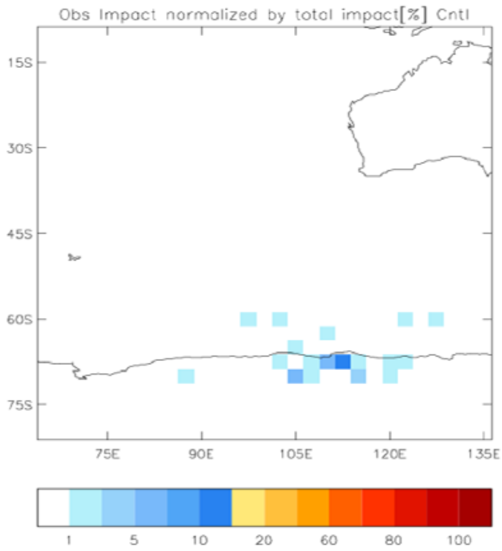

Figure 1. Percentage values of the observation impact normalised by the total global impact at 18UTC 5 June 2010 on the coast of Antarctica.

It was thought that the large observation impacts seen might be caused by so-called 'super-sensitivity' which is defined by Baker and Daley(2000) as following.

Under these conditions, when the observation density is low or there is an abrupt change in observation density, the magnitude of the observation and/or background sensitivities may greatly exceed the analysis sensitivity. We have defined this phenomenon as 'super-sensitivity'

Figure 2 shows the forecast sensitivity to observations and to analysis corresponding to the impacts in Figure 1. For simplicity's sake we only show the forecast sensitivity to analysis (i.e. analysis sensitivity) for the variable theta and the observation sensitivity for ATOVS data which covers most of the horizontal and vertical region and which is related to the theta field. We can see that the observation sensitivity does not exceed the analysis sensitivity in the region and that the intense coastal feature is similar in each. This shows that the large observation impact in this case is not caused by super-sensitivity but is a result of large analysis sensitivities in the region.

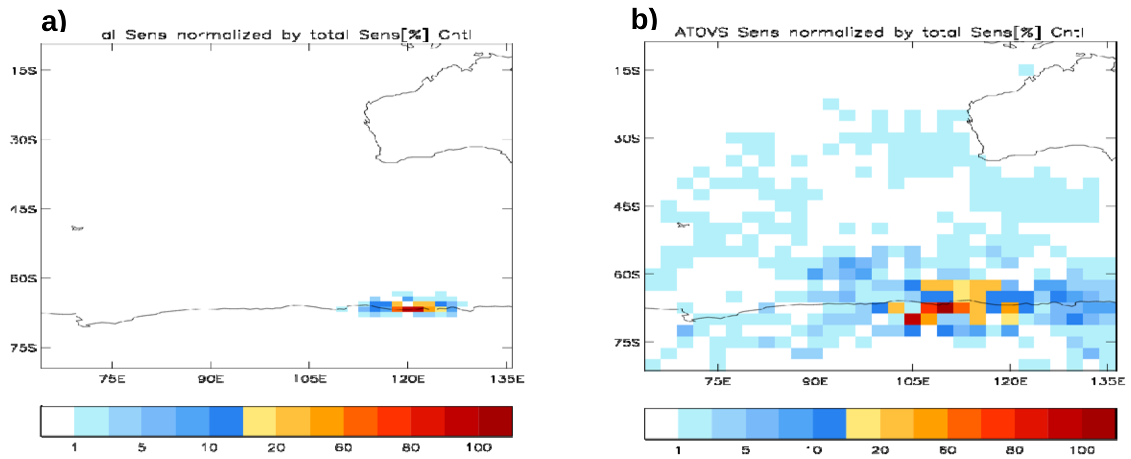

Figure 2. Percentage values of the analysis sensitivity for theta field(a) and observation impact of ATOVS(b) normalised by the total global sum at 18UTC 5 June 2010 on the coast of Antarctica.

The contribution of the g^{th} grid point to the total analysis impact is

$$J_g = \delta \hat{\mathbf{w}}_0^T \left(\delta \mathbf{w}_0 \right)_g . \qquad (10)$$

The analysis sensitivity, $\delta \hat{\mathbf{w}}_0$, and analysis increment, $\delta \mathbf{w}_0$, near to the area of exaggerated observation impacts are examined to identify the main factor making the analysis impacts large. The analysis increment combines contributions from the Kalman gain and observation innovations as shown in (9). Figure 3 shows profiles of globally averaged analysis impacts, sensitivities and increments and the horizontal distribution of those fields at the level of maximum analysis impact for each element in (10). For the sake of simplicity, only the theta field is shown. Other fields display similar properties. Values are normalised by the level averages so that outliers can be seen easily. It is clear that the large fluctuation in analysis impact matches closely the analysis sensitivity for the lower levels above the coast of Antarctica. We can say that the analysis weights (**K**) and observation innovations are not the reason for the large analysis impact in this case, meaning that approximations in background error covariances and observation error variances don't contribute to the large observation impact and that the exaggerated impacts are not related to the misspecification of errors in analysis weights as discussed by Andersson et al. (2000), the problem actually lying in some component of the analysis sensitivity.

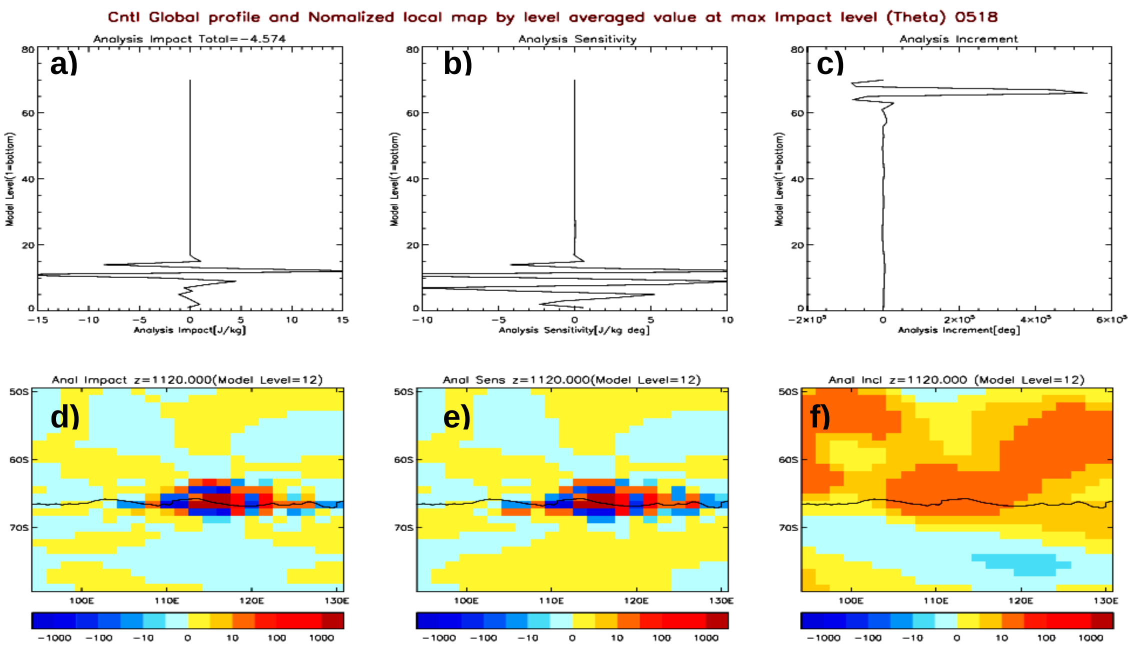

Figure 3. Analysis impacts (a, d), analysis sensitivities (b, e), and analysis increments (c, f) for the theta field of the analysis for 18UTC 5 June. Vertical profiles are normalised by the level-averaged value (a, b, c) and horizontal distributions are normalized by the level averaged value at 1120m (d, e, f) where the analysis impact is largest.

The analysis sensitivity is calculated by applying the adjoint of the linearised model to the adjoint sensitivity with respect to the forecast difference. We can diagnose the time at which large analysis sensitivities appear by plotting domain-averaged values of adjoint sensitivity from time *t* to *0*.

Figure 4 shows the average sensitivity of 24-hour forecast error to model perturbations at 3-hour intervals throughout the adjoint PF-model integration. The sensitivity field is averaged over the same horizontal domain as in Figure 3 at the level of 1120m, the level at which the largest analysis impacts are found. Met Office 4D-Var produces an analysis increment valid at the start of its 6-hour analysis window and so the adjoint integration is performed over 27 hours to find the analysis impact. The value at time 0 is the average adjoint sensitivity to the forecast difference - $\delta \hat{\mathbf{w}}_t$ in (4). It seems that the adjoint sensitivity fluctuates abnormally after 21 hours and then abruptly increases after 27 hours of integration. The reason for the abrupt change in the mean theta field is not clear but it is worth checking the linearity assumption made in equation (8). This can be done by comparing the linearly evolved perturbation, $\delta \mathbf{w}_t^L$, with the actual perturbation which is the difference between the two nonlinear forecasts from analysis and background, $\delta \mathbf{w}_t^{NL}$, as is shown in equation (11).

$$\delta w_t^{NL} = \delta w_t^{fa} - \delta w_t^{fb}, \ \delta w_t^{L} = \mathbf{M}_{0,t} \delta w_0 . \qquad (11)$$

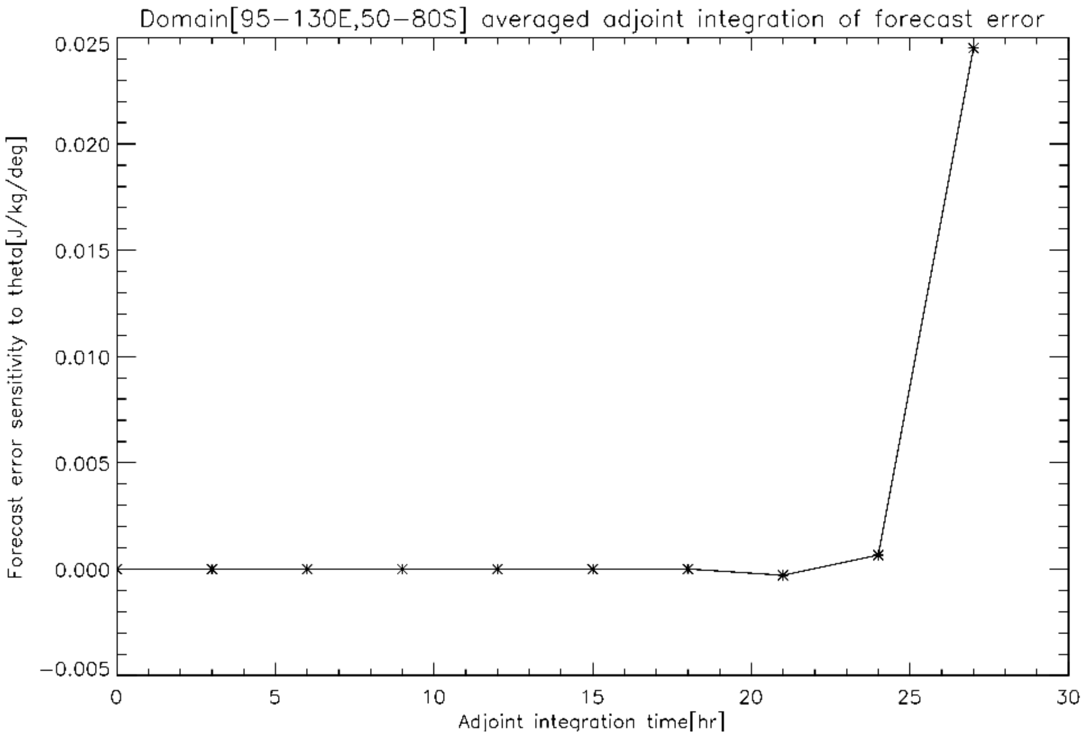

Figure 4. Domain-averaged values of adjoint sensitivity to perturbations in the theta field at 3-hour intervals. The domain is from 95E to 130E and 50S to 80S.

The difference between linear and nonlinear calculations of 24-hour forecast change due to analysis is shown in Figure 5. The mean temperature difference at 1120m exceeds 50 degrees and its standard deviation is about 10,000 degrees. The difference is confined to the region of exaggerated observation impacts along the coast of Antarctica, shown in Figure 1.

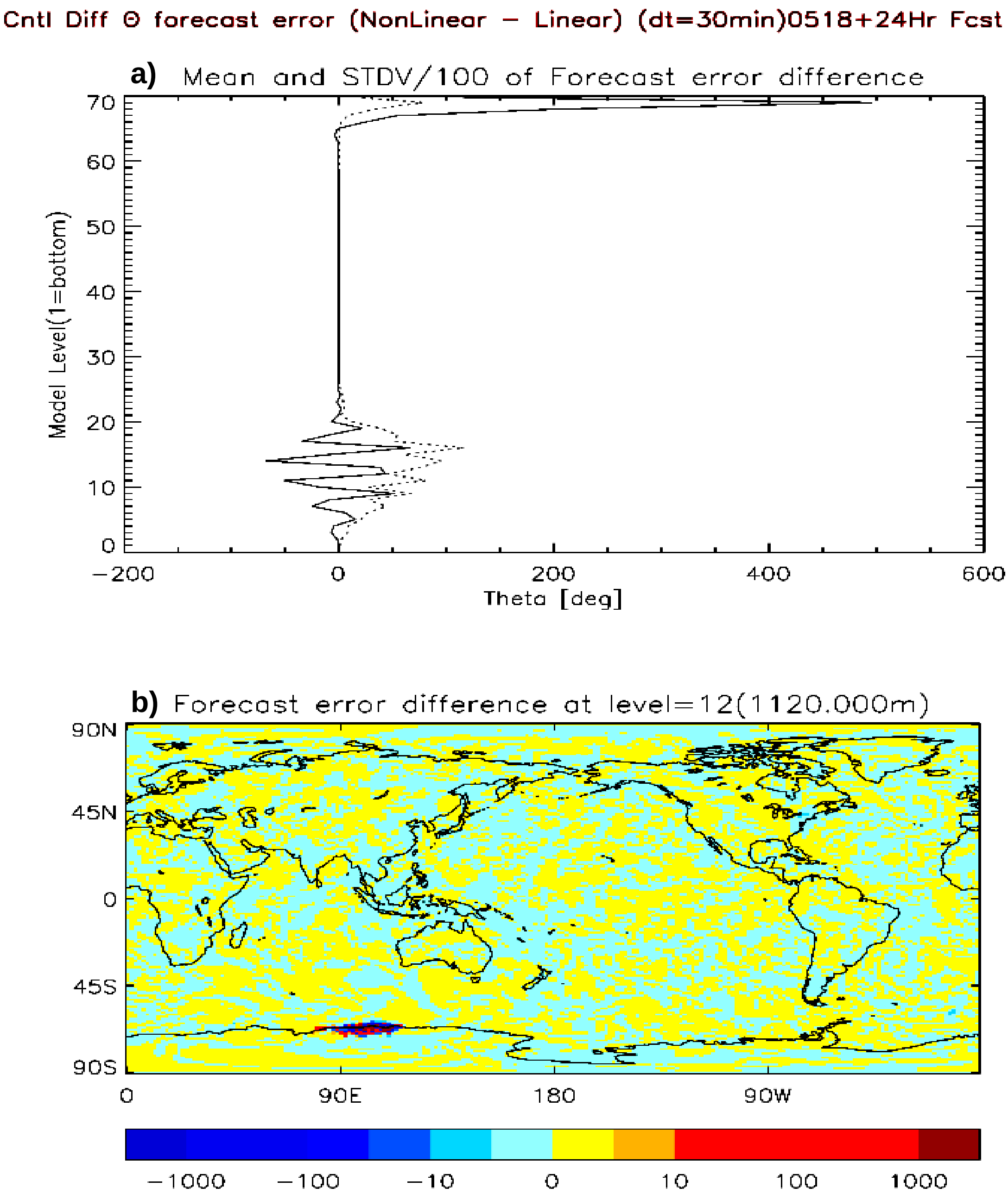

Figure 5. The difference in linear and nonlinear 24-hour forecast (error) change due to analysis. Global averaged mean (solid line) and standard deviation (dotted line) profile (a) and horizontal distribution at the level of maximum analysis impact, 1120m (b). (The standard deviation is divided by 100 for ease of plotting on the same scale as the mean profile.)

These results indicate a computational stability problem in the adjoint model's integration scheme. Similar problems have been observed occasionally, often associated with static instability in the linearization state, small grid lengths E-W near the poles, and steep orography (Tim Payne, personal communication). A modified version of the PF and adjoint model system has been developed to alleviate such problems. We have not yet tested this case with this new version, but we have found that a reduced timestep does eliminate the instability in this case.

Figure 6 shows the effect of performing the adjoint PF-model integrations at different temporal resolutions. The timestep used for the adjoint forecast model step of the observation sensitivity calculation in this study was 30 minutes. This is less than the 45 minutes temporal resolution of the PF-model in the 4D-Var inner loop. At temporal resolutions greater than 15 minutes results are similar to that at 15-minute resolution (not shown here). As the timestep is increased to 20 minutes, the mean difference between nonlinear and linear forecasts becomes apparent after 20 hours; with a 30-minute timestep a large difference in the mean theta field appears after only 10 hours. This shows that in observation sensitivity studies using longer forecasts (e.g. 48 to 72 hours) a shorter timestep would be needed to avoid the excessive growth of unstable modes.

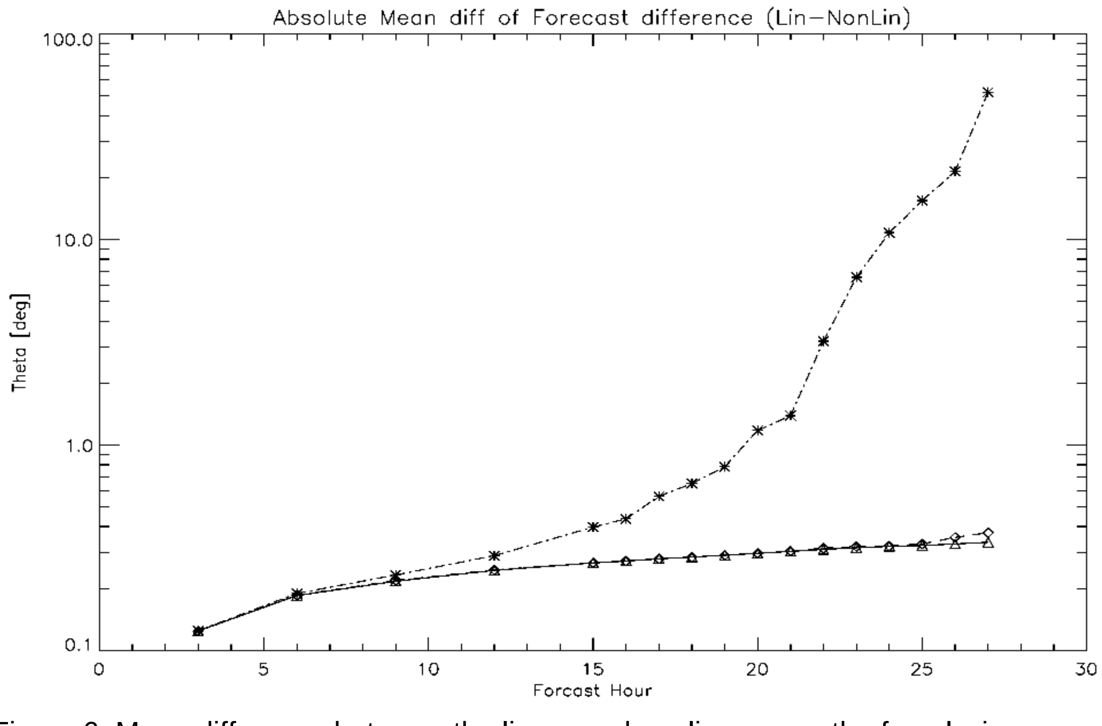

Figure 6. Mean difference between the linear and nonlinear growth of analysis perturbations after 24 hours with timesteps in the linear model of 15 minutes (solid line with triangular marks), 20 minutes (dotted line with diamond marks), and 30 minutes (dotted and dashed line with asterisk marks).

The large difference between nonlinear and linear forecasts is removed for timesteps of 15 minutes or less and the large difference in the low levels over Antarctica (shown in Figure 4) also disappears (not shown here). The analysis and observation impacts with the reduced timestep of 15 minutes are shown in Figure 7. The largest observation impact is less than 0.2% of the total impact and the excessively large impacts in the low levels over Antarctica (see Figure 1) are not present in results from a modified system.

The observation impacts are recalculated with an adjoint PF-model timestep of 15 minutes for the period from 18UTC 1 to 12UTC 7 June 2010 to see the effect of the exaggerated sensitivities on the routine application of the adjoint sensitivity tool. The comparison of total observation impacts between 15 minutes and 30 minutes PF-model timestep is shown in Figure 8. The adjoint sensitivity tool over-represents the impact of ATOVS and SYNOP observations which cover the coast of Antarctica where the large sensitivities were found, while the impact of IASI data is under-represented. This result shows that checks for exaggerated observation impacts should be performed to avoid the effects of computational stability problems if observation impacts are to be interpreted correctly.

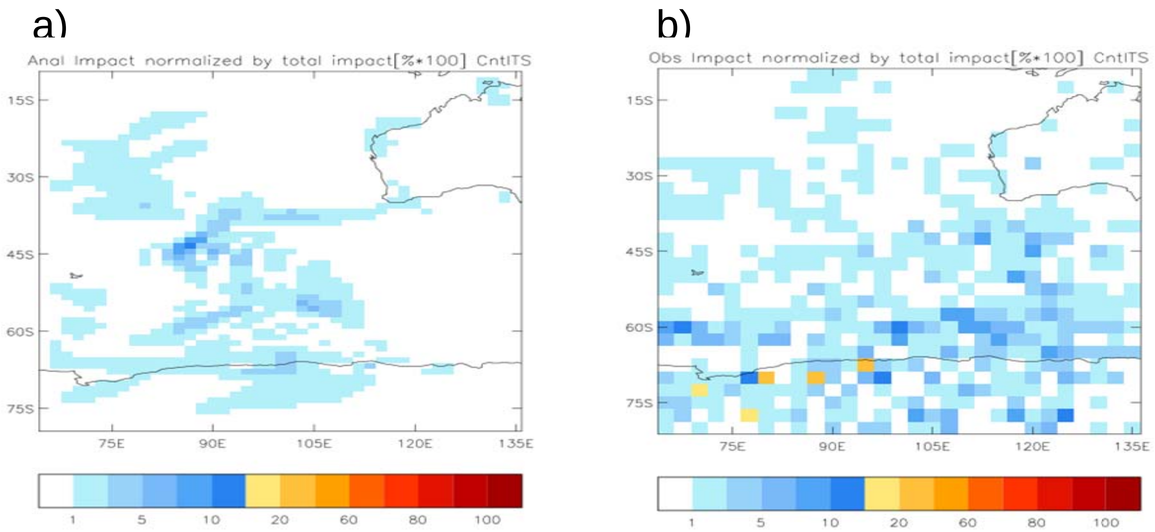

Figure 7. Percentage values of the analysis sensitivity for theta field(a) and observation impact of ATOVS(b) normalised by the total global sum at 18UTC 5 June 2010 on the coast of Antarctica with adjoint integral timestep of 15 minutes. The scales are multiplied by 100.

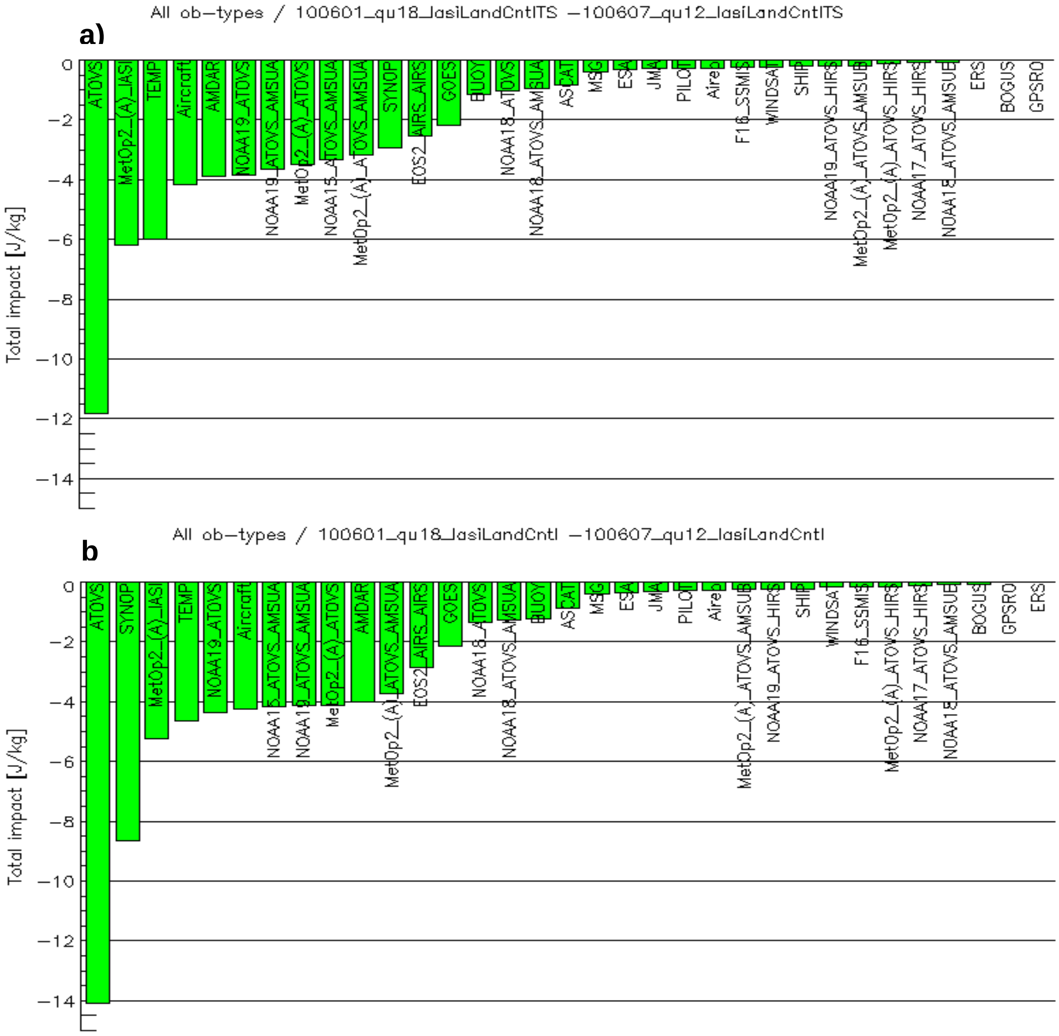

Figure 8. Global total observation impact of each observation type for the period from 18UTC 1 to 12UTC 7 June 2010 with adjoint PF-model timesteps of 15 minutes (a) and 30 minutes (b).

4. Summary and future work

We studied the cause of exaggerated observation impacts produced by the adjoint-based observation sensitivities tool at the Met Office. The calculation of impacts on 24-hour forecast error for the 18UTC 5 June 2010 analysis was investigated. This case showed extremely large observation impacts along the coast of Antarctica which accounted for ~20% of the total global impact. The large observation impact is different from that found in cases of so-called “super-sensitivity” and we found its cause lay in problems with the adjoint PF-model’s time-integration scheme. Similar problems have been seen in 4D-Var at the Met Office and a modified version of the scheme has since been developed. The problem can be avoided in this case by increasing temporal resolution making the adjoint integration computationally stable.

We calculated observation impact with the reduced model timestep of 15 minutes, which makes the adjoint integration stable in this study, and compared it with the routine observation impact results using 30-minute timestep for the period from 18UTC 1 to 12UTC 7 Jun 2010. SYNOP observations were measured as the 2nd most important observation type in the routine observation impact but became only the 11th in the revised impact study. IASI impacts were under-represented and ATOVS impacts were over-represented in the routine observation study. Observation impacts in the region of the adjoint model instability were so large that the order of significance of observation types was largely affected leading to incorrect interpretation of results.

It would be useful to routinely monitor for excessively large observation impacts during the application of the adjoint sensitivity tool to avoid the possibility of misleading results. As a preliminary attempt, it is recommended that each component of the analysis impact be monitored such as in Figure 2. Analysis sensitivity fields for all variables can be made available for monitoring over the global domain with no additional costs in the adjoint sensitivity calculations other than storage.

Although reductions in the timestep-length of the adjoint PF-model have shown to be effective in removing computational instabilities it would be prudent to first test the modified version of Var, mentioned in section 3, using the original PF-model settings. This would avoid any unnecessary cost in running the observation sensitivities system. Even then it is likely that a reduced timestep length may be required, if not for measuring the impact on 24-hour forecasts then for longer ones.

5. References

Andersson E., M. Fisher, R. Munro, and A. McNally, 2000: Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case poor convergence. *Quart. J. Roy. Met. Soc.* **126**: 1455-1472.

Baker N. L., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation targeting problem. *Quart. J. Roy. Met. Soc.* **128**: 1431-1454.

Joo S., 2011. Adjoint sensitivity tool applied to satellite observation over land. *Presentation at DAOS Working Group 4.*

Lorenc A. C., and R. Marriott, 2012. Observation impacts in the Met Office global NWP system. In preparation.

Gelaro R., R. H. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX observation impact intercomparison experiment, *Mon. Wea. Rev*., **138**, 4009-4025.

Met Office
FitzRoy Road, Exeter
Devon EX1 3PB
United Kingdom

Tel (UK): 0870 900 0100 (Int) : +44 1392 885680
Fax (UK): 0870 900 5050 (Int) :+44 1392 885681
enquiries@metoffice.gov.uk
www.metoffice.gov.uk