ESA CONTRACT REPORT

Contract Report to the European Space Agency

Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction

WP-5000 report: Feasibility demonstration of 4D-Var assimilation system using CloudSat and CALIPSO observations

M. Janisková and M. Fielding

ESA ESTEC contract 4000116891/16/NL/LvH

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

ECMWF

Series: ECMWF ESA Project Report Series

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Contract Report to the European Space Agency

Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction

WP-5000 report: Feasibility demonstration of 4D-Var assimilation system using CloudSat and CALIPSO observations

Authors: M. Janisková and M. Fielding

ESA ESTEC contract 4000116891/16/NL/LvH

May 2018

ABSTRACT

This report describes the data assimilation system at European Centre for Medium Range Weather Forecasts (ECMWF) prepared for feasibility demonstration of direct inclusion of cloud radar and lidar observations into the Four-Dimensional Variational (4D-Var) system. It provides information on observations, their errors and the used observation operators. The necessary data handling for observations in the data assimilation system, such as quality control and bias correction is also briefly summarized.

4D-Var assimilation experiments have been performed using CloudSat (NASA's cloud radar mission) cloud radar reflectivity and CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) lidar backscatter observations, either separately or in combination. Obtained results indicate that 4D-Var analyses get closer to assimilated observations. However, impact of the cloud radar reflectivity is larger than that of the lidar backscatter. Impact on the first-guess (FG) and analysis (AN) departure statistics when verified against other observation types assimilated in 4D-Var is most pronounced and positive for conventional observations, especially for wind. For all other observations the observed impact is small. An impact of the new observations on the subsequent forecast is largest for zonal wind, relative to both operational analysis and radiosonde observations. Results also indicate an improved forecast of rain rates in the tropics by assimilating cloud radar and lidar observations.

Contents

1 Introduction

Active satellite instruments provide a vertically resolved characterization of clouds for the benefit of numerical weather prediction (NWP). Observations of clouds from almost coincident space-borne radar and lidar are already available with the launch of the CloudSat (Stephens et al., 2002) and the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO, Winker et al., 2009) missions in 2006. In the future, such type of data should also be provided by the Earth, Clouds, Aerosol and Radiation Explorer (EarthCARE, Illingworth et al., 2015) mission (to be launched in 2021). Whether information on clouds extracted from such data can be beneficial for NWP analyses and forecasts has been studied at the European Centre for Medium-Range Weather Forecasts (ECMWF).

In the past, assimilation experiments for cloud radar and lidar observations have been performed at ECMWF using a two-step technique which combines one-dimensional variational (1D-Var) assimilation with four-dimensional variational (4D-Var) data assimilation (Janisková et al., 2012; Janisková, 2015). In those studies performed during the previous ESA funded projects (Janisková et al., 2010, 2014), pseudo-observations of temperature and specific humidity retrieved from 1D-Var using observations of cloud radar reflectivity and lidar backscatter were assimilated in the ECMWF 4D-Var system to assess the impact of radar and lidar observations on the analyses and subsequent forecasts. The results from these experiments have shown that 1D-Var analyses fit both assimilated and independent observations better that the first guess, suggesting that the assimilation is able to produce a more realistic state of atmosphere and clouds when radar and lidar observations are available. The 1D+4D-Var assimilation experiments have indicated a positive impact of the new observations also on the subsequent forecasts. Thus performed studies demonstrated the potential offered by the assimilation of cloud observations from space-borne radar and lidar instruments. However, they also indicated that for any future operational implementation, the direct 4D-Var assimilation of cloud related observations needs to be considered. One of the reason for the direct assimilation is that the 1D+4D-Var assimilation method also needs to calculate 1D-Var retrieval errors that serve as observation errors in the second stage when retrieved temperature and humidity profiles are assimilated in 4D-Var. Their computation from the 1D-Var analysis covariance matrix is expensive for profiling observations and only affordable in non-operational applications. Thus in the current project, necessary developments have been done (Fielding and Janisková, 2017; Janisková et al., 2017) to the assimilation system at ECMWF in preparation for the inclusion of cloud radar and lidar data directly into the 4D-Var data processing.

Section 2 describes the general methodology of the 4D-Var system and provides information on observations, their errors and the used observation operators. It also briefly summarizes the necessary data handling for cloud radar and lidar observations, such as quality control and bias correction, in the data assimilation system. Experimental setup for the performed study is described in Section 3. Results from 4D-Var experiments using CloudSat cloud radar reflectivity and CALIPSO lidar backscatter observations are presented in Section 4. Finally, Section 5 provides a summary and draws the conclusions of this study.

2 4D-Var assimilation system

2.1 Methodology

The operational data assimilation system at ECMWF is 4D-Var (Rabier et al., 2000; Mahfouf and Rabier, 2000). This system is based on the incremental formulation proposed by Courtier et al. (1994).

The goal of 4D-Var is to define the atmospheric state such that the distance between model trajectory and observations is minimum, in a least-square sense, over a given time period $[t_0, t_n]$. The model trajectory is constrained to fit the observations by adjusting its initial conditions. Using the incremental approach, which allows to take into account some nonlinearities by updating the trajectory iteratively in the vicinity of the state around which the model is linearized, 4D-Var searches for the analysis increment $\delta\mathbf{x}_0$ at initial time t_0 that minimizes the following cost function:

$$\mathscr{J}(\delta\mathbf{x}_0) = \underbrace{\frac{1}{2}(\delta\mathbf{x}_0)^T\mathbf{B}^{-1}(\delta\mathbf{x}_0)}_{\mathscr{J}^b} + \underbrace{\frac{1}{2}\sum_{i=0}^{n}\left(H_i'\delta\mathbf{x}_i - \mathbf{d}_i\right)^T\mathbf{R}_i^{-1}\left(H_i'\delta\mathbf{x}_i - \mathbf{d}_i\right)}_{\mathscr{J}^o} + \mathscr{J}^c \tag{2.1}$$

where $\mathscr{J}^b$ is the background cost function, $\mathscr{J}^o$ is the observation cost function and $\mathscr{J}^c$ is the constraint cost function used in 4D-Var to control fast gravity waves using the digital filter technique developed by Gauthier and Thépaut (2001). At any time t_i, $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^b$ is the analysis increment representing the departure of the model state ($\mathbf{x}$) with respect to the background ($\mathbf{x}^b$) which consists of temperature, humidity, vorticity, divergence and surface pressure in the current 4D-Var system. $\mathbf{d}_i = \mathbf{y}_i^o - H_i(\mathbf{x}_i^b)$ is the so-called innovation vector providing the departure of the model background equivalent ($H_i(\mathbf{x}_i^b)$) to the observation ($\mathbf{y}_i^o$). H_i' is the linearized observation operator which also includes the spatial interpolations to the observation locations as well as the forecast model execution propagating the initial state $\mathbf{x}_0$ to the time of observation. $\mathbf{R}_i$ is the observation error covariance matrix (including measurement and representativeness errors) and $\mathbf{B}$ is the background error covariance matrix of the state $\mathbf{x}^b$ based on a wavelet formulation (Fisher, 2004) to introduce regime-dependent error statistics.

The incremental approach reduces the computational cost of 4D-Var since the perturbations $\delta\mathbf{x}_i$ and the gradient of the cost function are computed at lower resolution using the simplified model during successive minimizations. After each minimization, the model trajectory is recomputed at high resolution. The minimizations are performed with the model containing simplified physical parametrizations of cloud and large-scale moist processes, convection, radiation, vertical diffusion, orographic and non-orographic gravity-wave drag (Janisková and Lopez, 2013).

2.2 Observational part of the system using cloud radar and lidar observations

2.2.1 Observations

In our study, measurements of cloud radar reflectivity (in dBZ), from the CloudSat 94 GHz radar and/or lidar backscatter (in km^{-1} sr^{-1} using logarithmic scale) due to clouds at 532 nm from CALIPSO are assimilated in the 4D-Var system. Cloudy-only lidar observations have been obtained using the Vertical Feature Mask product (Liu et al., 2009; Vaughan et al., 2009) to remove range bins containing only molecular backscatter or affected by aerosols. The observation horizontal resolution is approximately 1.4 km for CloudSat radar and it varies between 300 m to 1.7 km for CALIPSO lidar depending on altitude. However, observations from these

instruments are averaged over each model grid-box at horizontal resolution approximately 36 km. This is twice coarser than experimental horizontal resolution of TCo639 (spectral truncation corresponding approximately to 18 km on a cubic octahedral grid) in our experiments.

2.2.2 Observation operator

The role of observation operator is to provide model equivalent to observations. In the case of radar/lidar assimilation, parameterization of cloud radar reflectivity and lidar backscatter due to clouds are required to convert model fields to reflectivity and backscatter, respectively. The radar and lidar observation operators are designed to meet the requirements of data assimilation system, i.e. to be computationally efficient and to allow coding of its adjoint counterpart. The observation operators developed in the previous studies (Janisková et al., 2010, 2014; Di Michele et al., 2014a,b) have been further refined and updated in the current project (Fielding and Janisková, 2017).

To save the computational cost of exact calculations on every profile, some calculations are made off-line. Specifically, the bulk microwave scattering properties (backscatter, extinction, single scattering albedo and asymmetry factor) are pre-computed at discrete intervals of water content and temperature by integrating individual particle single scattering properties over a given particle size distribution (PSD). The bulk scattering properties are computed for the six different prognostic hydrometeor types used within the ECMWF Integrated Forecasting System (IFS): stratiform and convective rain, stratiform and convective snow and liquid cloud and ice cloud. These are then stored in lookup tables for the radar and the lidar wavelengths. For each model atmospheric profile and level where the model-equivalent radar reflectivity or lidar backscatter are required, a bilinear interpolation is performed using the model temperature, T, and in-cloud water content, w, to obtain the unattenuated radar reflectivity and lidar backscatter, respectively. To reduce the computational cost of searching a pre-computed look-up table, the relationship between radar reflectivity or lidar backscatter and T and w can be parametrized using a two-variable two-degree polynomial with fitted coefficients. A detail description can be found in the WP-2000 report, Part 2 (Fielding and Janisková, 2017). The below description provides just brief summary of operator specifics:

i *Radar observation operator.* The total unattenuated radar reflectivity is computed as a sum of the unattenuated reflectivity for each hydrometeor type obtained from pre-computed look-up table or using the fitting method. Attenuation of the signal along the radar beam is also taken into account. Clear sky attenuation due to gases is calculated using the models of Liebe (1985) and Liebe et al. (1992). Cloudy sky attenuation is given by the sum of clear sky attenuation and the hydrometeor extinction. Due to computational cost, multiple scattering (MS) is not considered in data assimilation system and situations where MS is likely (like heavy precipitation) are excluded from the analysis.

ii *Lidar observation operator.* The lidar simulator for clouds has been developed as an extension of the reflectivity model. Since the typical wavelength of lidar ($\sim 1\ \mu$m) is much shorter than radar (~ 1 cm), lidar observation operator must account for two additional processes: scattering from gases (known as Rayleigh scattering or molecular backscatter) and multiple scattering. MS effects are not negligible for lidar since the lidar photons are much more likely undergo multiple scattering within clouds that radar due to greater extinction and strong forward peak of the phase function. For affordability reasons, a simple MS approach is used in the data assimilation. It is based on Platt approximation (Platt, 1973) and uses a multiplier which is applied to the optical depth of each cloudy layer. The value of multiplier can vary between 1 (the single-scattering limit) and 1/2 (the wide field-of-view limit).

2.2.3 Observation errors

The errors assigned to observations determine the weight to be given to the observation during the assimilation process. If the assumed error is large, this weight is small and so is the potential impact of the corresponding observation. An underestimation of the error may have a detrimental impact on the assimilation process as the analysis is drawn closer to the observation than the actual quality (reliability) of the measured values would otherwise suggest. However, under-weighting of observations will not lead to larger analysis errors. The assumed observation errors take into account the instrumental errors, together with forward modelling and representativity errors due to the narrow field of view (Fielding and Janisková, 2017):

i The instrument error is the random error in the measurement due to noise. If not directly provided as Level 1B product, it needs to be specified based on known instrument noise characteristics.

ii Forward modelling errors for both CloudSat and CALIPSO are based on evaluation of uncertainty in the microphysical assumptions. To obtain these errors, hundred different, physically plausible realisations of the scattering look up table are generated. The uncertainty in the bulk scattering properties is calculated by finding the standard deviation as a function of water content and temperature.

iii Representativity error describes the degree of mis-match between the observation scale and model scale in both space and time. For observations with narrow field of view (FOV), as satellite observations of cloud using radar and lidar, the representativity of the observations represent a significant issue and dominates the observation error in the majority of the cases. This error is obtained by a method based upon the assumption that the local variability of measurements along the satellite track is representative of the model gridbox variability, and the spatial correlation between measurements can be approximated using climatological correlation. To estimate representativity error, the standard deviation of measurements within each gridbox is computed during the assimilation and then multiplied by the appropriate scale factor based on the pre-computed correlation function.

2.3 Data handling for cloud radar and lidar observations

To handle observations appropriately in the data assimilations system, their quality control and screening as well as bias correction scheme are required.

2.3.1 Data quality control and screening

The aim of quality control and screening is to discard measurements flagged as of poor quality or leading to excessive first guess (FG) departures (i.e. very large differences between observed and corresponding simulated values). They are needed to avoid the assimilation system to perform sub-optimally in cases when *(i)* measurements are unphysical, *(ii)* forward model is not capable of representing the observations or *(iii)* there are excessive non-linear effects in the observation operator. Whereas screening blacklists observations for an objective reason, quality control reduces the weight of observations that are outliers to their expected error.

The quality control of CloudSat radar and CALIPSO lidar observations is based on statistics of the FG departures between observed reflectivities/backscatters and corresponding simulated observations generated by the forward operator. This lead to the choice of initial screening indicators and thresholds. Although any ground return should be removed by the cloud mask during pre-processing, observations with 1 km of the surface are blacklisted for safety. Cloud fraction is used as another indicator for screening. Both the model and observed values of cloud fraction use a minimum threshold screening of 0.2 to avoid non-linearity and representativity issues. Other screening thresholds considered:

i For CloudSat *radar observations*, the minimum detectable signal around -30 dBZ and the maximum signal 20 dBZ from hydrometeors (Kollias et al., 2007) are used as limits. The same thresholds are applied to the simulated reflectivities to avoid bias occurrance from assymetric sampling. Observation when the FG departures are smaller than -20 dB or larger than 20 dB are also screened out. The final screening applied to radar reflectivity is to remove profiles where multiple scattering is suspected. Following Battaglia et al. (2011), using a threshold of the integrated reflectivity, 41.3 dBZ_{int}, provides an optimum detection of cases where the magnitude of multiple scattering exceeds 3 dB.

ii For *lidar observations*, the final screening task is to remove observations that are likely to be severely attenuated. The integrated backscatter is used as a proxy for the amount of signal attenuation since the actual attenuation in the observations is unknown. To prevent the introduction of biases, the mean integrated backscatter of the actual and simulated observations is calculated. By setting the threshold in integrated backscatter at 0.02 sr^{-1}, observations having a transmission of less than 0.1 (being equivalent to an attenuation of 10 dB) will be excluded.

2.3.2 Bias correction

Since assimilation relies on the assumption that both observations and model background are unbiased quantities, it is necessary to quantify these systematic errors and to remove them. This is achieved by applying a bias correction scheme which is able to guarantee zero-mean statistics. Though observations themselves can be biased (for instance due to calibration errors of instruments or scanning angle errors), the main sources for biases are likely to be linked to the model. Model systematic errors can be result of the forecast model inaccuracies as well as deficiencies in the observation operators which are used to provide the model equivalent to the observations. The success of bias correction scheme depends on the choice of the indicators used. Indicators allow a different bias correction to be applied to the observations in different situations. In the case of satellite observations, while some calibration error is likely, the dominant source of bias is likely due to microphysical assumptions made in the observation operators and requires several different indicators to achieve reliable performance.

Selection of the most appropriate bias correction indicators is based on the the statistics of FG departures (Fielding and Janisková, 2017). The selected indicators for cloud radar and lidar observations used in experiments are hight (altitude), temperature, model hydrometeor type and mean radar reflectivity or mean lidar backscatter, respectively. The bias correction scheme produces look-up tables for the climatological bias based on all or a subset of the selected indicators.

3 Experimental framework

Using the full system of regularly assimilated observations at ECMWF, several assimilation experiments have been performed either only over the 12-hour period (i.e. the current length of 4D-Var assimilation window at ECMWF) between 31 July 2007 21:00 UTC and 1 August 2007 09:00 UTC (situation **2007080100**) or over 10 days covering 1-10 August 2007 period. 4D-Var experimentation has been done using a horizontal resolution of TCo639 spectral truncation (corresponding to approximately 18 km on a cubic octahedral grid) and 137 vertical levels.

(a) Experimental set-up

The following experiments have been run:

- **REF**: reference run, i.e run with all regularly assimilated observations, but without new cloud radar and lidar observations included in the 4D-Var system;
- **ALL_RAD_LID**: experiment assimilating cloud radar and lidar observations on top of all other normally assimilated observations;
- **ALL_RAD**: experiment assimilating only cloud radar observations in combination with all other assimilated observations;
- **ALL_LID**: as **ALL_RAD**, but using cloud lidar observations;
- **RAD_LID**: 4D-Var experiment assimilating cloud radar and lidar observations only;
- **RAD**: experiment using cloud radar observations only;
- **LID**: 4D-Var run with cloud lidar observations only.

Observations averaged over model grid-box at approximately 36 km have been used with:

- **1err** - single observation errors as defined in Section 2.2.3;
- **2err** - observation errors twice as large as defined;
- **0.5err** - observation errors half of the defined value;

10-day forecasts have been run from the analyses to study the impact of the new observations not only on 4D-Var analyses, but also on subsequent forecasts.

(b) Description of selected cases

Majority of the experiments have been performed for the situation **2007080100** in order to provide a detail investigation of the impact coming from including new observations into the system. Figure 3.1 illustrates a typical observation coverage, multiple tracks of CloudSat and CALIPSO satellites for the 12h period from 21:00 UTC on 31 July 2007 to 09:00 UTC on 1 August 2007.

Situation **2007080100** cover a variety of meteorological conditions as shown on Fig. 3.2 for two selected tracks in the 12-hour period. Figure 3.2 a,b displays global satellite imagery from both geostationary and polar orbiting satellites under the CloudSat orbit track. The satellite imagery is a combination of geostationary data provided by the National Center for Environmental Prediction (NCEP) and Advanced Very High Resolution Radiometer (AVHRR) polar) imagery provided by the Comprehensive Large Array-data Stewardship System (CLASS) of the National Oceanic and Atmospheric Administration (NOAA). Figure 3.2 c shows the model total water content for the same situation at analysis time, i.e. 1 August 2007 00:00 UTC. On top of that there is the observation coverage for the 12-hour period from 21:00 UTC on 31 July 2007 to 09:00 UTC on 1 August 2007.

For the situation **2007080100**, there is remarkable good agreement in cloud position between observations and model. Even tropical cyclone, such as Typhoon Usagi located southeast of Japan on 1 August 2007, or Tropical Storm Erick (eastern Pacific), are well captured by the model. Category 4 Typhoon Usagi was at its maximum intensity on the morning of 1 August. At the time, the storm packed sustained winds of 222 km/h with gusts to 269 km/hr, but it weakened slightly in the hours that followed before striking Japan near Nobeoka, Miyazaki as a rapidly weakening typhoon late on 2 August. For Tropical Storm Erick, with a peak intensity winds of 65 km/h and a minimum barometric pressure of 1004 hPa at 0600 UTC on 1 August, continued wind shear caused Erick to gradually weaken to a tropical depression. In the selected period, there was also a short-lived Tropical Storm Chantal in the northern Atlantic Ocean. It developed on 31 July 2007 between Bermuda and Cape Cod and began undergoing extratropical transition early on 1 August 2007.

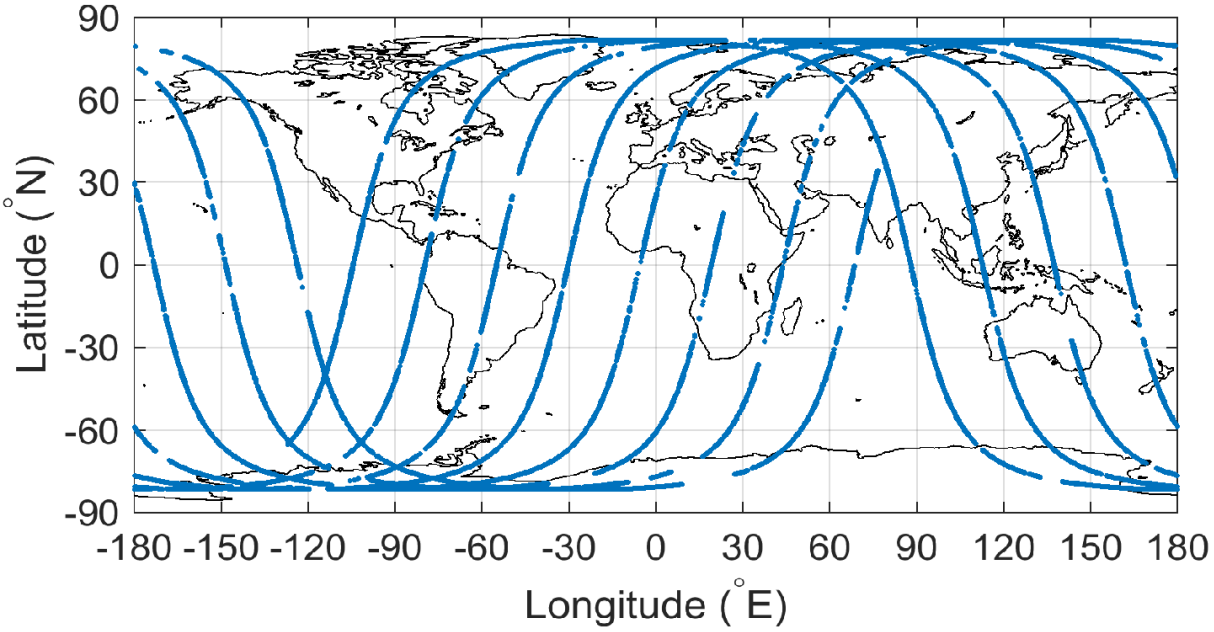

Figure 3.1: CloudSat/CALIPSO data coverage for the 12h period between 31 July 2007 21:00 UTC and 1 August 2007 09:00 UTC.

Not only tropical cyclones, but also plenty of extratropical cloud formations are well captured by the model. A good agreement between the model and the satellite imagery is, for instance, obvious south of Africa, south-east of Australia or over eastern part of Russia.

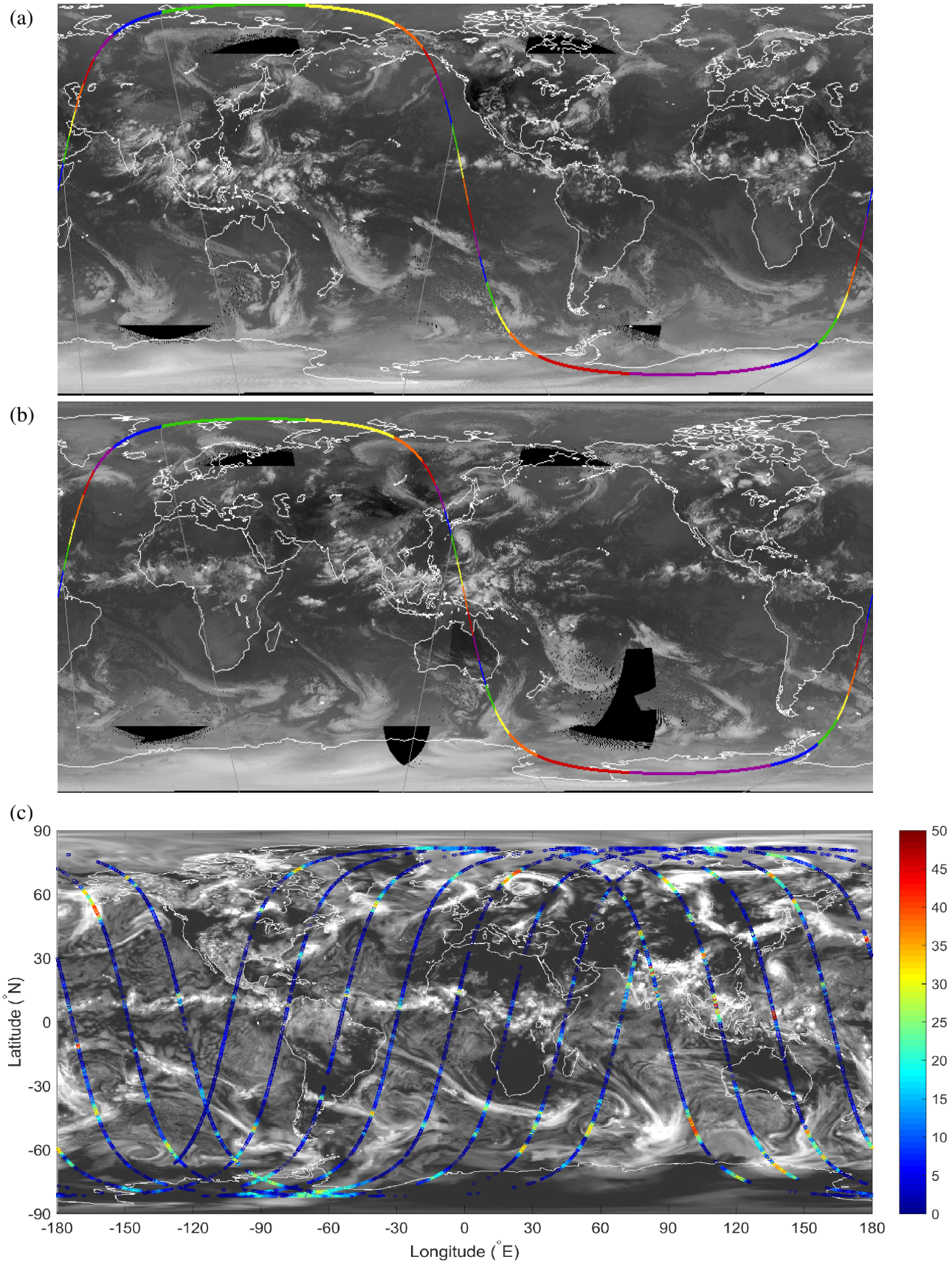

Figure 3.2: Satellite imagery and CloudSat track for the period (a) 21:09 - 22:48 UTC on 31 July 2007 and (b) 03:44 - 05:20 UTC on 1 August 2007. (c) displays the model total water content at analysis time, 1 August 2007 00:00 UTC.

4 Results of 4D-Var experiments using cloud radar and lidar observations

4.1 Impact on 4D-Var analysis

4.1.1 Comparison against own observations

Figure 4.1 shows probability distribution function (PDF) of the FG and analysis (AN) departures for the assimilated cloud radar reflectivity from the different assimilation experiments (**REF**, **ALL_RAD_LID**, **RAD_LID**) and for the different domains as specified in figure caption. The PDFs of first-guess departures get peak more centralised around zero after applying bias correction. This indicates that built correction (Fielding and Janisková, 2017) by correcting for systematic differences, as intended, provides better PDF distributions. As the analyses are getting closer to the observations, the PDFs of analysis departures become even more narrow. The most symmetric PDF shapes are achieved by **RAD_LID** experiment assimilating just cloud radar reflectivity and lidar backscatter. This is expected as 4D-Var minimisation does not need to adjust to any other different types of observations when assimilating only cloud radar and lidar data comparing to the experiments using also all other observations regularly assimilated by the system. The PDF of both the FG and AN departures is wider and also less smooth for the Tropics than for any other domains. This is due to prevailing convective situations there which are more difficult to capture by both the forecast model and the observation operator. This is improved when looking at departures normalised by observation error (Fig. 4.2). The normalised departures are generally more narrow and smoother for all the domains.

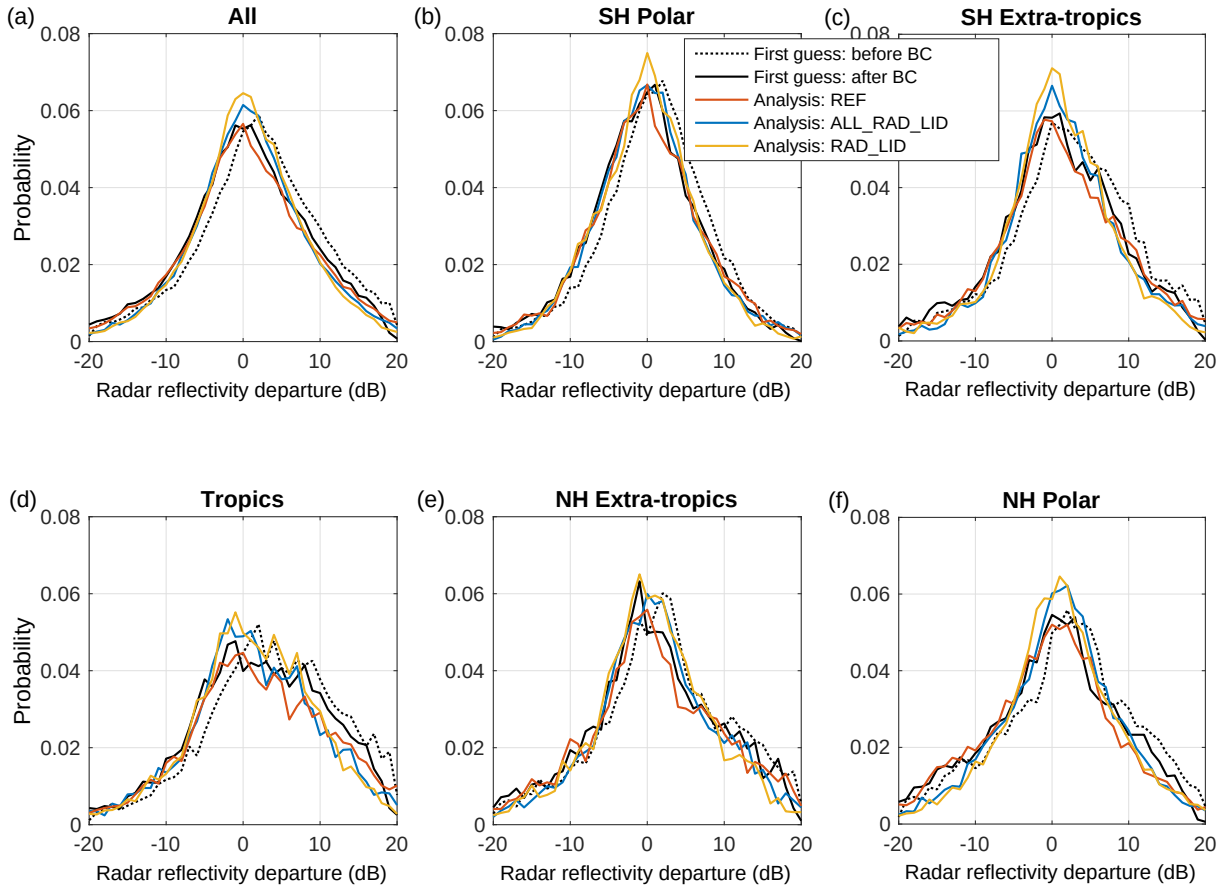

*Figure 4.1: Probability distribution functions of the first guess departures before (black dotted line) and after (black solid line) bias correction applied combined with the analysis departures for the reference experiment (***REF***, red solid line), experiment assimilating cloud radar and lidar observations with all other normally assimilated observations (***ALL_RAD_LID***, blue solid line) and experiment assimilating cloud radar and lidar observations only (***RAD_LID***, orange solid line). Results are presented for radar reflectivity departures (dB) over (a) the whole globe, (b) South Hemisphere (SH) Polar area, (c) SH Extra-tropics, (d) Tropics, (e) North Hemisphere (NH) Extra-Tropics and (f) NH Polar area. Situation* **2007080100** *with 12-hour assimilation period 31 July 2007 21:00 UTC - 1 August 2007 09:00 UTC.*

The PDFs of the FG and AN departures for cloud lidar backscatter are shown in Fig. 4.3 for the different experiments and domains. Firstly, considering observations from all domains (Fig. 4.3a), the bias correction again performs as expected, producing a smoother, more gaussian-like curve centred on zero. Secondly, only assimilating routine observations (**REF** experiment), the PDF of analysis departures contains the least observations between -10 and 10 dB, with lower peak than the bias corrected first-guess PDF. This is a natural consequence of the screening we apply to the observations; observations are only assimilated where there are clouds in both model and observations at the initial screening. If the model adds cloud to areas that passed screening, it is unlikely that the departures would fall out of the -10 dB to 10 dB range. However, if the model removes clouds (partly due to the log scale) it is possible for large negative departures to be created, thus reducing the total number of analysis departures in the -10 to 10 dB range.

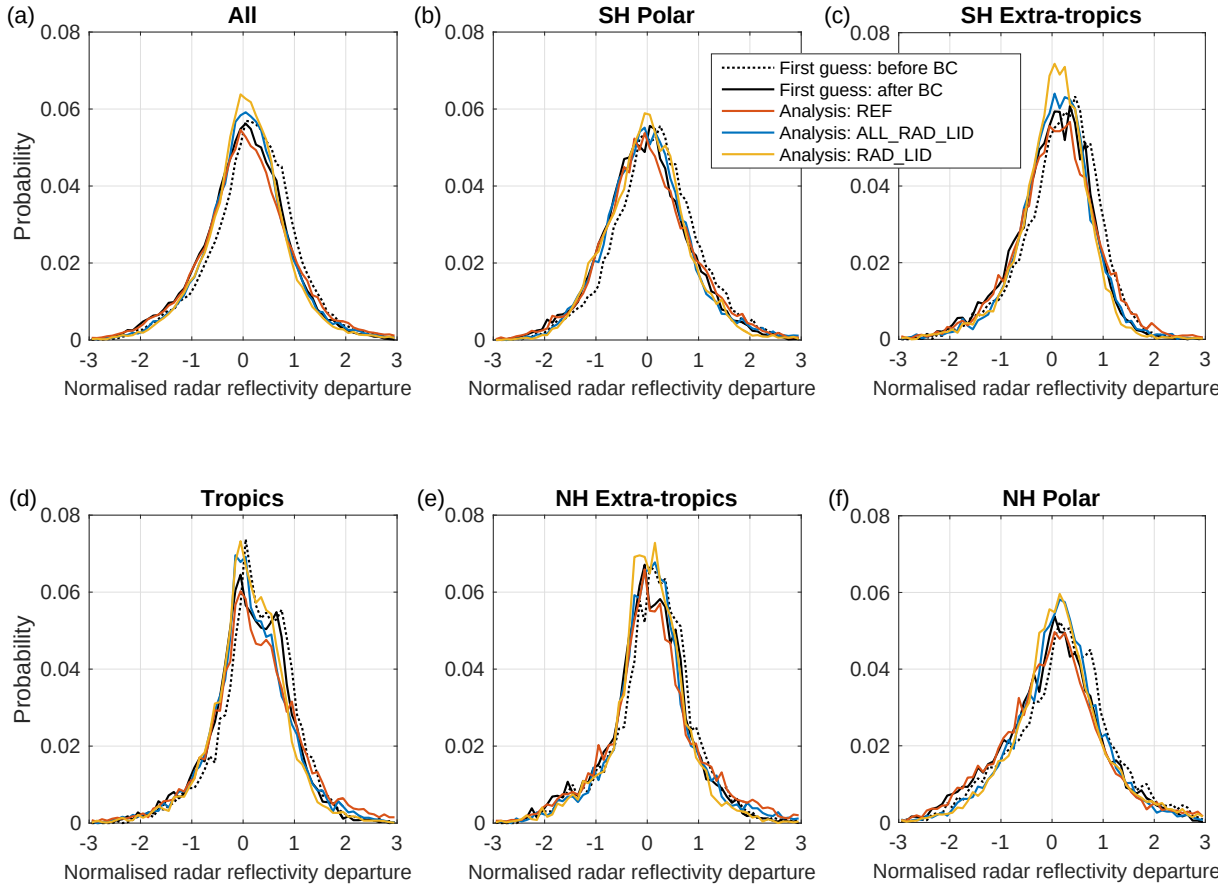

Figure 4.2: Same as Fig. 4.1, but for normalised radar reflectivity departures.

In general, all assimilation experiments that include assimilating the lidar observations increase the peak in the analysis PDF around zero, showing that the AN is closer to the lidar observations than the FG. For the **ALL_RAD_LID** experiment the fit to observations is improved slightly, while using only radar and lidar (**RAD_LID**) increases the fit more. The greatest fit to the lidar observations is found when assimilating lidar only (**LID**), which is not surprising as there are no other observations that could influence the analysis and draw the lidar AN values away from the observations. Similar to the radar, the regionally stratified departures are smallest for the southern hemisphere. Unlike the radar, there is less difference between the tropics and elsewhere, presumably due to the abudance of high ice cloud that is prevalent in the tropics, which is well observed by the lidar. Again, similar to the radar, the normalised PDFs (Fig. 4.4) are smoother and more gaussian-like, suggesting our observation error model is proficient at differentiating regimes with variable expected error.

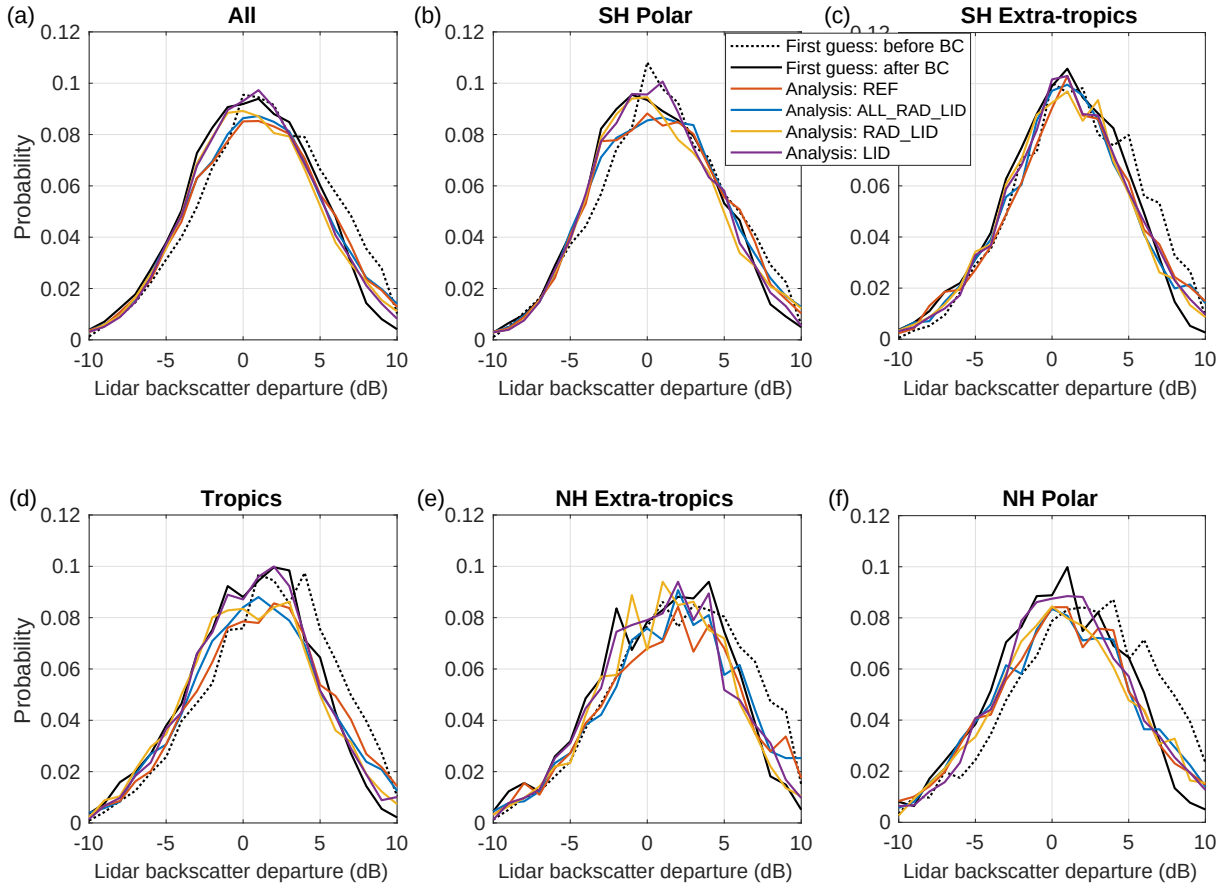

*Figure 4.3: Probability distribution functions of the first guess departures before (black dotted line) and after (black solid line) bias correction applied combined with the analysis departures for the reference experiment (***REF***, red solid line), experiment assimilating cloud radar and lidar observations with all other normally assimilated observations (***ALL_RAD_LID***, blue solid line), experiment assimilating cloud radar and lidar observations only (***RAD_LID***, orange solid line) and assimilation using lidar observations only (***LID***, violet solid line). Results are presented for lidar backscatter departures (dB) for (a) the whole globe, (b) South Hemisphere (SH) Polar area, (c) SH Extra-tropics, (d) Tropics, (e) North Hemisphere (NH) Extra-Tropics and (f) NH Polar area. Situation* **2007080100** *with 12-hour assimilation period between 31 July 2007 21:00 UTC and 1 August 2007 09:00 UTC.*

Model level dependent values of FG and AN departures for radar and lidar for 10-day assimilation cycling are shown in Figs. 4.5 and 4.6 respectively. Note that models levels decrease in height from the model top to the surface. Only the reference experiment (**REF**, plotted in red) and the experiment assimilating all observations, including radar and lidar (**ALL_RAD_LID**, plotted in black) are shown. To ensure meaningful statistics when considering variables in logarithmic space, an additional screening step is applied to the results to reject absolute AN departures greater than 20 dB for radar reflectivity and 10 dB for lidar backscatter. This screening is applied to both AN departures and FG departures. In total over 350000 superobs (i.e. observations averaged over the defined grid) of radar reflectivity and 160000 superobs of lidar backscatter are considered during this period. The largest amount of observations assimilated tends to be in the polar and extra-tropical regions for the radar, while the greatest amount of observations assimilated for the lidar tends to be in the tropics. For both the radar and the lidar, the smallest root mean square (RMS) errors in FG and AN departures can be seen in the Southern Hemisphere (Figs. 4.5c and 4.6c).

Firstly considering the radar reflectivity statistics in more detail, we can see that the RMS in AN departures is consistently less when radar and lidar observations are assimilated. However, the FG departure is greater when

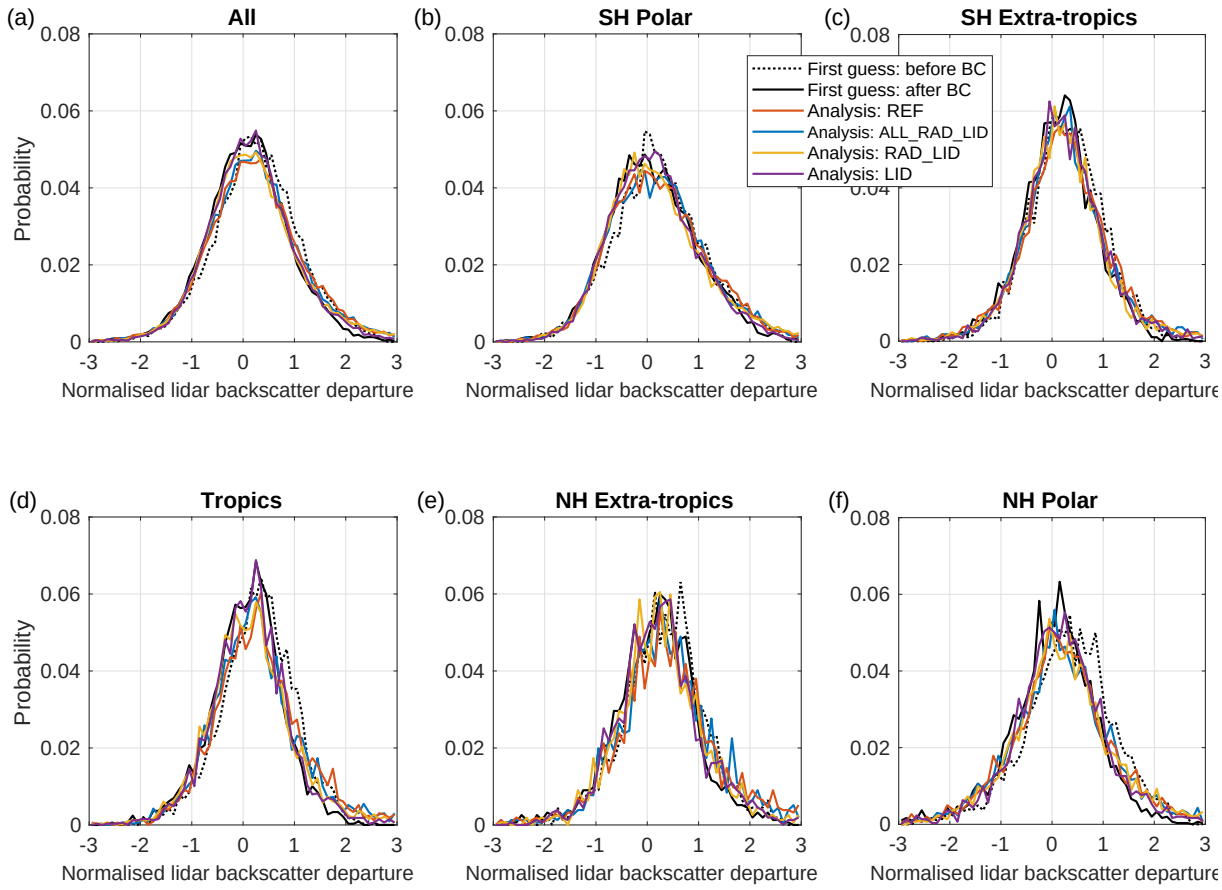

Figure 4.4: Same as Fig. 4.3, but for normalised lidar backscatter departures.

radar and lidar observations are assimilated. This is due to the additional screening that is applied to both the FG and AN; because the RMS in AN departures are generally smaller for the **ALL_RAD_LID**, observations with greater FG departures will be kept and used to generate the statistics. The same results can be found when considering the bias in radar reflectivity departures; the bias in AN departures is significantly smaller in magnitude at all height levels and domains for the **ALL_RAD_LID** experiment. The bias in FG departures for the **ALL_RAD_LID** experiment tend to be smaller than for the **REF** experiment, but there are occasions when they are greater, for example in lower model levels in the North Hemisphere, again due to the additional screening applied to the results.

Next, comparing to the CALIPSO data, the AN RMS is generally smaller when all observations including radar and lidar are assimilated. The greatest differences can be found in the upper troposphere in the tropics and at heights typical of the boundary layer top in the southern hemisphere. In all regions RMS differences are smallest in the upper troposphere. There is a trend for the magnitude of bias in AN departures to be reduced in the **ALL_RAD_LID** experiment relative to the **REF** experiment for the majority of levels, but a longer period of assimilation cycling is required for this to be significant. It is likely that the strong attenuation of the lidar signal makes interpreting the bias difficult, particularly for model layers closer to the surface.

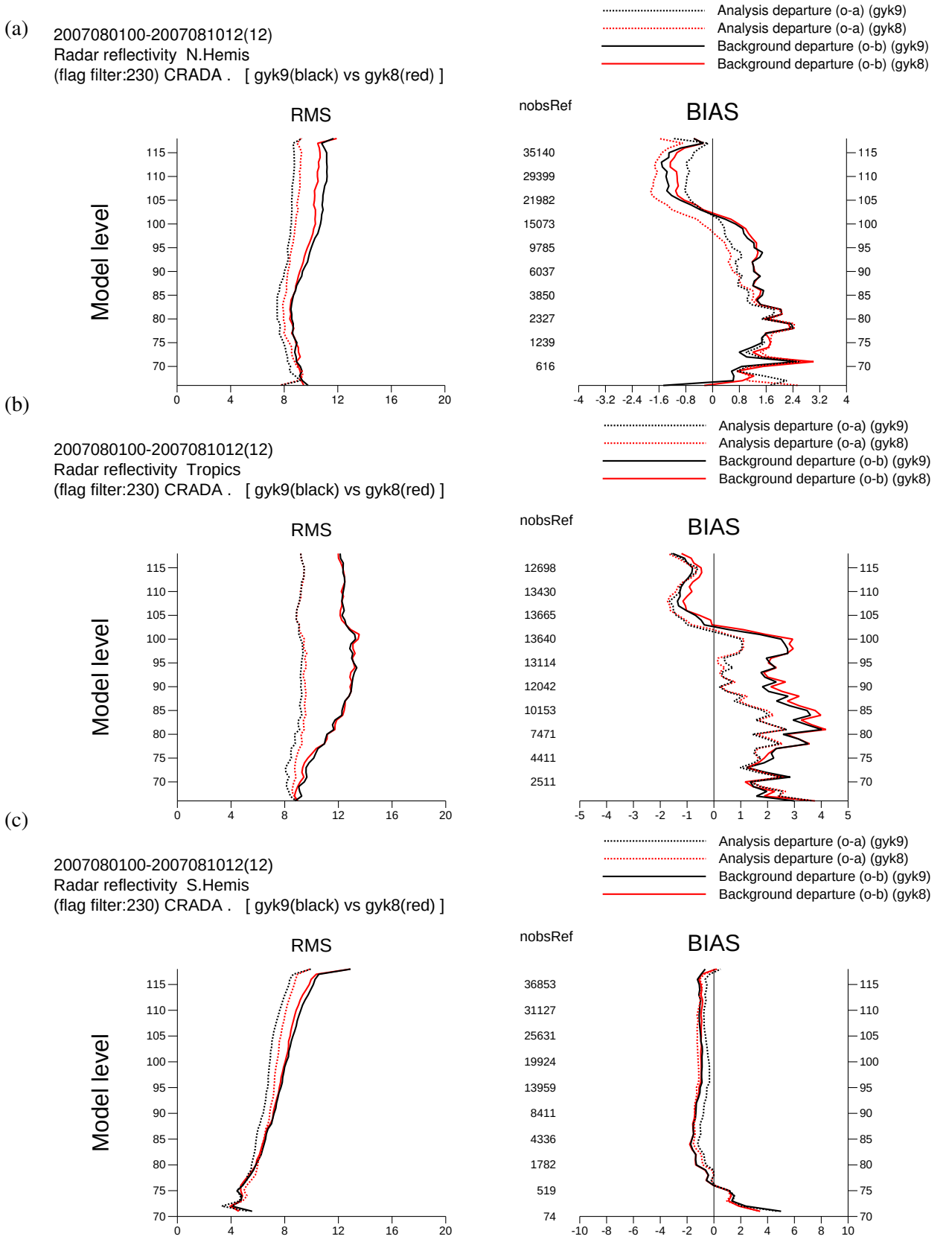

*Figure 4.5: Root mean square error (left) and bias (right) of background departures (solid line) and analysis departures (dotted line) with respect to cloud radar reflectivity observations for the reference experiment (***REF***, red) and experiment assimilating in addition cloud radar reflectivity and lidar backscatter observations using double observation errors (***ALL_RAD_LID_2err***, black). The number of observations for the period from 1 August 2007 00:00 UTC to 10 August 2007 12:00 UTC is displayed in the middle. Results are shown for (a) Norther Hemisphere, (b) Tropics and (c) Southern Hemisphere. The larger the model level number, the closer it is to the surface (i.e., level 117 is approximately 1km above the surface and level 70 is approximately 13 km above the surface).*

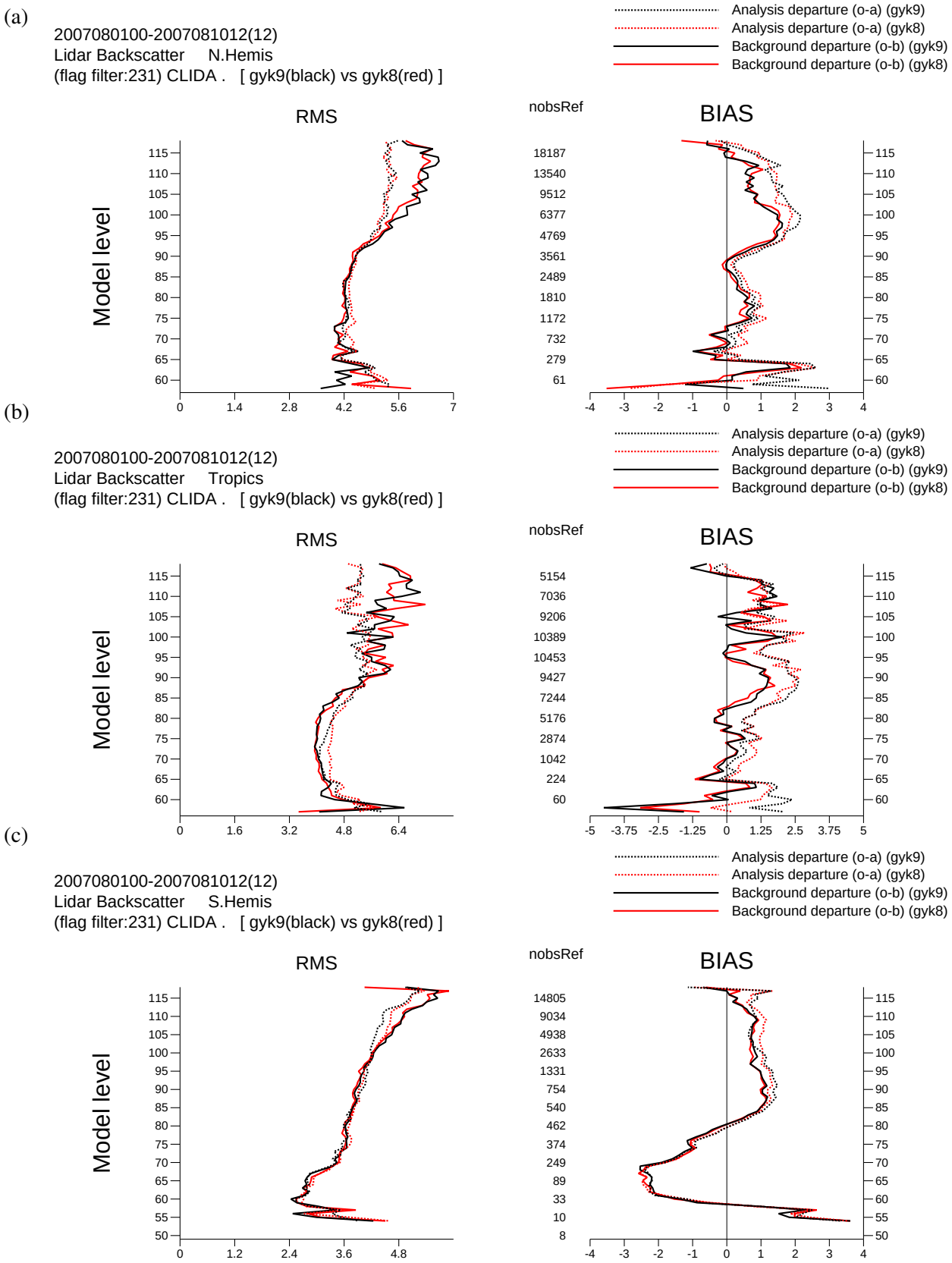

Figure 4.6: Same as Fig. 4.5, but for cloud lidar backscatter.

4.1.2 Comparison of experiments against conventional observations

In this section the impact of assimilating radar and lidar observations on the analysis is assessed by comparing the FG departures of conventional observations (such as radiosondes or aircraft observations) between the **REF** and **ALL_RAD_LID** experiments. The experiments used 10 days of assimilation cycling, which allows the assimilation system time to adapt ('spin-up') to the new observations and provides more robust statistics than only using one day. All statistics shown are averaged to pressure levels and subset by three geographical regions: North Hemisphere, Southern Hemisphere and the Tropics. The impact of scaling the observation error is also evaluated.

First, comparing against radiosonde wind profiles (Fig. 4.7) we can see that assimilating radar and lidar using the original error model tends to degrade the model background wind field in both the North Hemisphere and the Tropics in terms of standard deviation of FG departures. However, increasing the observation errors by a factor of two (**2err**) to account for the known correlations reverses this finding, with an overall reduction in standard deviation of fit to radiosonde wind observations relative to the reference experiment. For both experiments using different error models, the bias in wind FG departures compared to the **REF** experiment is similar; in both cases the magnitude of the bias is marginally smaller. As the the double error model produces the most promising results, we will only consider this experiment for the rest of this section.

Figure 4.8 shows an evaluation with respect to wind observations from aircraft. While radiosonde observations provide detailed vertical structure of the atmosphere locally, aircraft observations contain a much greater quantity of information on the global atmospheric state, particularly in the north hemisphere. Given the reasonably short length of assimilation cycling performed, they are a particularly useful conventional observation type to verify with. Comparing the figure with Fig. 4.7a, a similar but smoother signal of neutral impact of radar and lidar assimilation on winds in the boundary layer, to an increasingly better fit higher up in the troposphere. The bias in background departures is slightly greater than the reference for most height levels, but the differences are typically much less than 0.01 % of the mean.

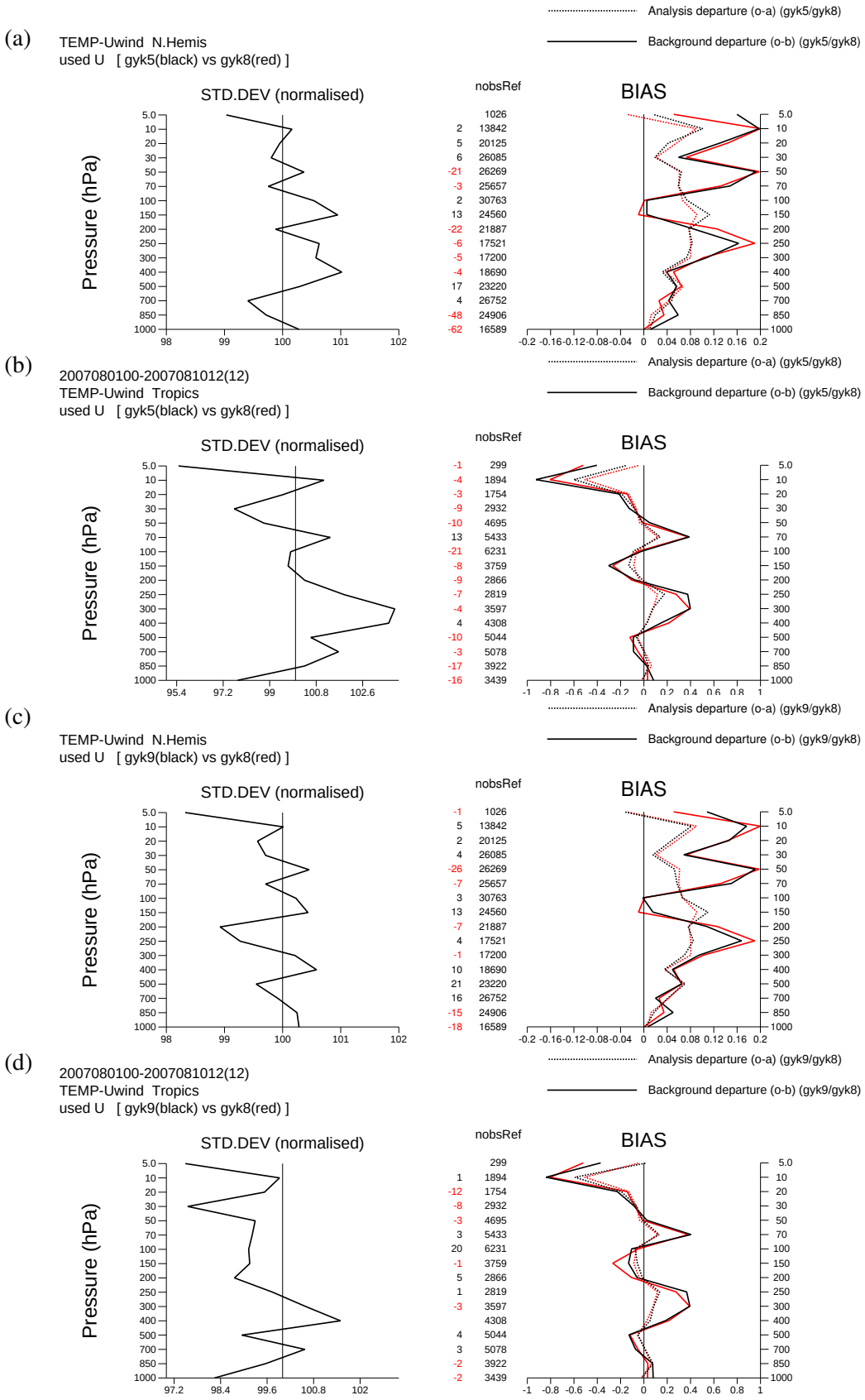

*Figure 4.7: Normalised standard deviation (left) and bias (right) of background (solid line) and analysis departures (dotted line) with respect to TEMP zonal wind observations for the reference run (**REF**, red) and experiment assimilating in addition cloud radar reflectivity and lidar backscatter observations (**ALL_RAD_LID**, black) using either (a, b) single or (c, d) double observation errors. The number of observations for **REF** experiment (nobsRef) for the period from 1 August 2007 00:00 UTC to 10 August 2007 12:00 UTC is displayed in the middle together with negative red and positive black numbers indicating how many less or more, respectively, observations are used by the **ALL_RAD_LID** run. Results are shown for (a, c) Northern Hemisphere and (b, d) Tropics.*

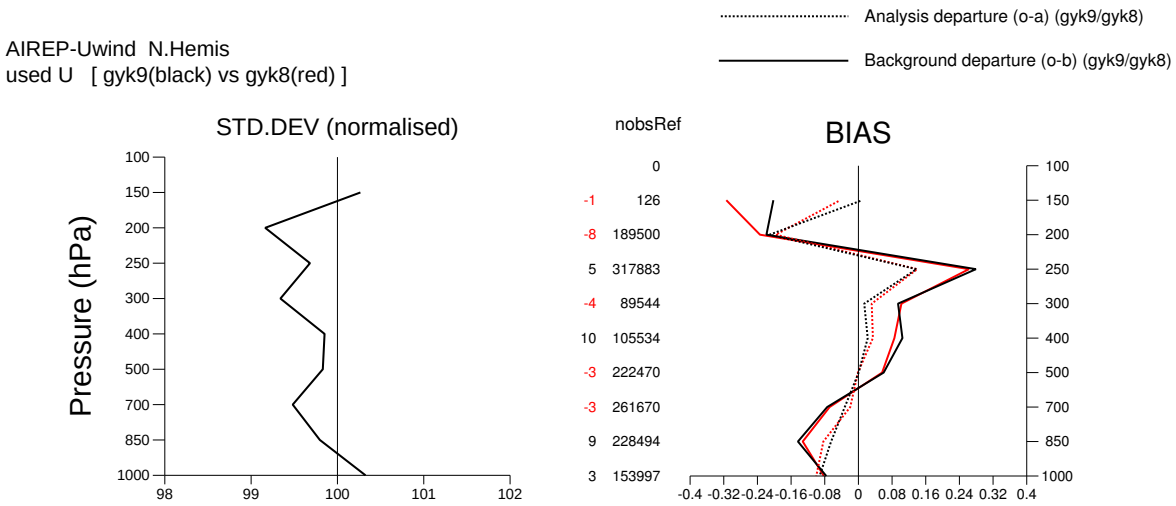

*Figure 4.8: Normalised standard deviation (left) and bias (right) of background (solid line) and analysis departures (dotted line) with respect to AIREP zonal wind observations for the reference run (***REF***, red) and experiment assimilating in addition cloud radar reflectivity and lidar backscatter observations (***ALL_RAD_LID***, black) using double observation errors. The number of observations for* **REF** *experiment (nobsRef) for the period from 1 August 2007 00:00 UTC to 10 August 2007 12:00 UTC is displayed in the middle together with negative red and positive black numbers indicating how many less or more, respectively, observations are used by the* **ALL_RAD_LID** *run. Results are shown over the Northern Hemisphere.*

Next, the impact of assimilating radar and lidar on conventional observations of temperature is shown in Fig. 4.9. For radiosonde measurements, the fit to observation background departures is improved at levels above 300 hPa, whereas it is slightly degraded below in both the North hemisphere and in the tropics. Conversely, for aircraft observations, the standard deviation is reduced by up to 1% throughout the troposphere in the tropics and no clear signal over the North hemisphere. This contradictory finding could be due to the fact radiosonde observations are more likely to be over land, while aircraft observations have a greater volume over ocean. Because the IFS model (and other NWP forecast models) tend to be better at representing oceanic cloud, a greater benefit can be extracted from the cloud radar and lidar observations in marine situations. Differences in the temperature bias between **REF** and **ALL_RAD_LID** are negligible for radiosondes and slightly improved at all levels for aircraft observations.

Finally, the difference in FG departures of specific humidity are shown in Fig. 4.10. In the Northern hemisphere troposphere, specific humidity tends to be closer to radiosonde observations when assimilating radar and lidar, but further cycling is required to remove noise from the signal. Similarly in the tropics the observations are limited so should be treated with caution, but show a better fit in the lower troposphere and a worse fit above. While differences in the bias are minimal in the North Hemisphere, the bias in the tropics is reduced, particularly around 850 hPa, which could be related to an improved representation of precipitation in the tropics, explore later in Sec. 4.2.3.

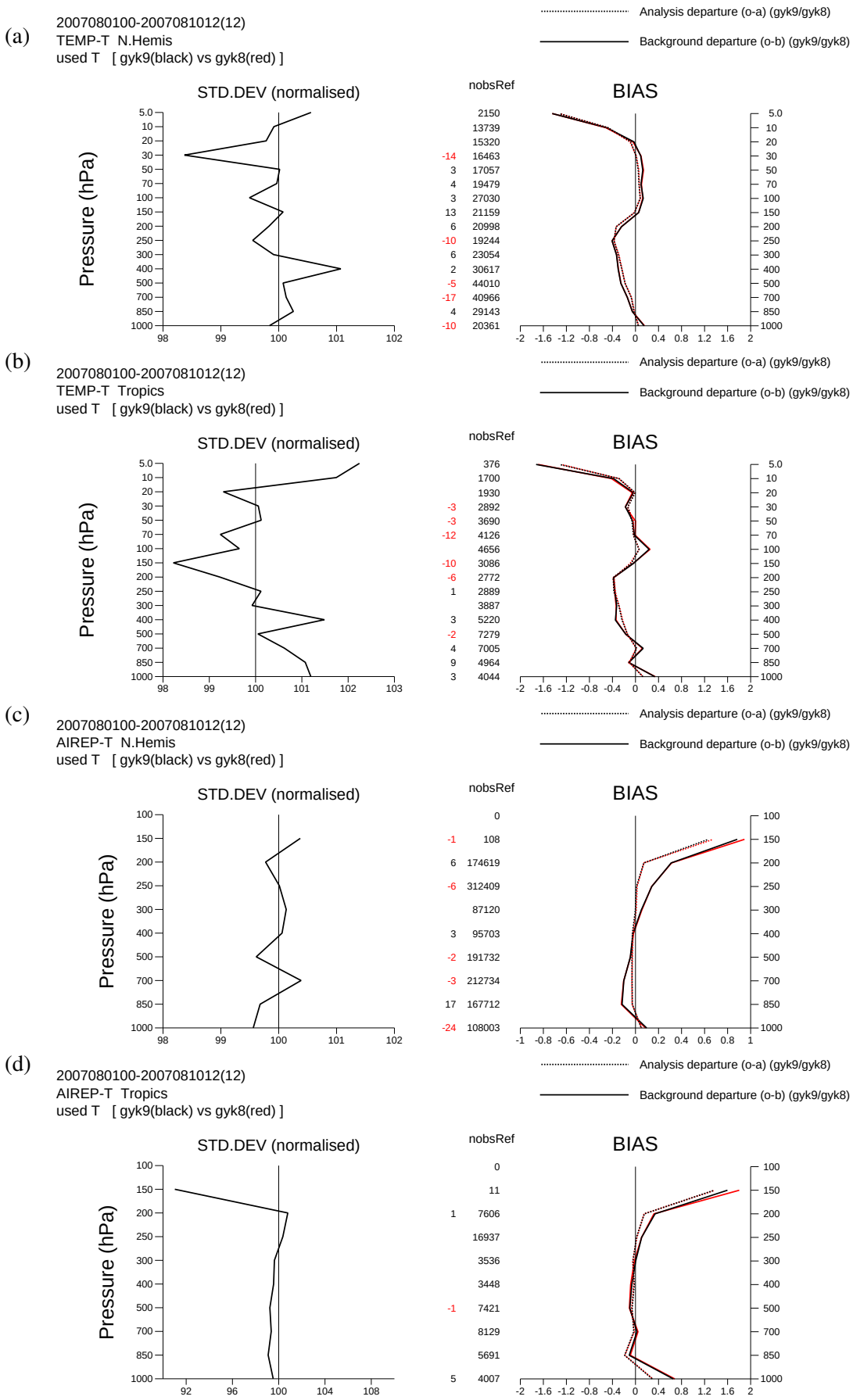

Figure 4.9: Same as Fig. 4.8, but for (a, b) TEMP and (c, d) AIREP (b, d) temperature observations over (a, c) the Northern Hemisphere and (b, d) Tropics..

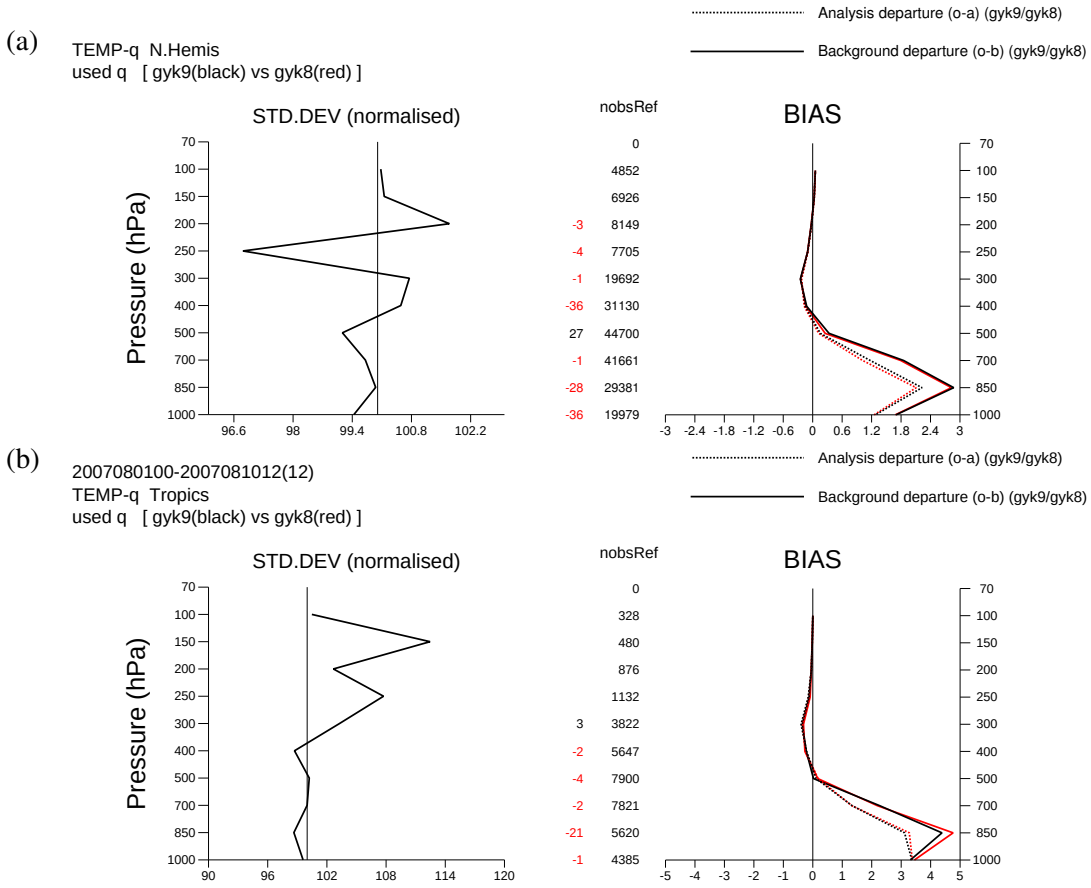

Figure 4.10: Same as Fig. 4.8, but for TEMP specific humidity observations over (a) the Norther Hemisphere and (b) Tropics.

4.1.3 Along-track evaluation and analysis increments

In addition to evaluating the impact of radar and lidar on the assimilation system via large-scale statistics, it is useful to investigate the performance of the system on individual clouds. To do this, we will consider the first 1000 model profiles containing radar and lidar observations assimilated in the data assimilation window between 21:00 UTC 31st July 2007 and 09:00 UTC 1st August 2007, which roughly corresponds to one complete orbit (Fig. 3.2c). The selected profiles cover a wide-range of regimes from boundary layer clouds (between 100-200), ice-cloud (between 300-500) and deep-convective clouds (between 750 and 1000). For convenience, we will refer to the profiles as 'the cross-section'.

On first inspection, the first-guess radar reflectivity generated from the model background (Fig. 4.11b) is remarkably similar to the superobbed CloudSat radar reflectivity for the cross-section (Fig. 4.11a) and thus gives confidence that the model's performance is sufficient to extract information from the observations. However, there are discrepancies and in particular there is: an over-estimation of precipitation in areas around the 300th and 600th profiles and a lack of precipitation around the 100th, 400th and 700th profiles. By considering the first guess departures (Fig. 4.11d), the regions with larger discrepancies become more obvious; blue (red) areas show that radar reflectivity is too large (small) in the model compared to observations.

The analysis radar reflectivity from the **ALL_RAD_LID** experiment is shown in Fig. 4.11c. As is expected, the analysis radar reflectivity provides a closer fit to the observations. For example, comparing the first guess and analysis panels, the areas of overestimated precipitation around the 300th and 600th profiles are reduced and the structure of the precipitation is much better resolved in relation to the observations. This is corroborated by the reduction in the analysis departures (Fig. 4.11e). Quantitatively, the magnitude of the bias and standard

deviation are reduced by 8 % and 5 % respectively.

The magnitude of the analysis departures relative to the first guess departures is partly dependent on the magnitude of the observation errors assigned to the radar observations (Fig. 4.11f), but also dependent on all the other observations assimilated and their respective observation errors. For the cross-section, the observation errors tend to be lowest in regions of ice cloud and stratiform precipitation, due to the expectation of lower representativity error. In more convective regions, the observation error is greater and therefore the analysis tends to be drawn less to the observations. However, there are also convective situations where the analysis is closer to the observations, such as the shallow convection seen around the 700th profile.

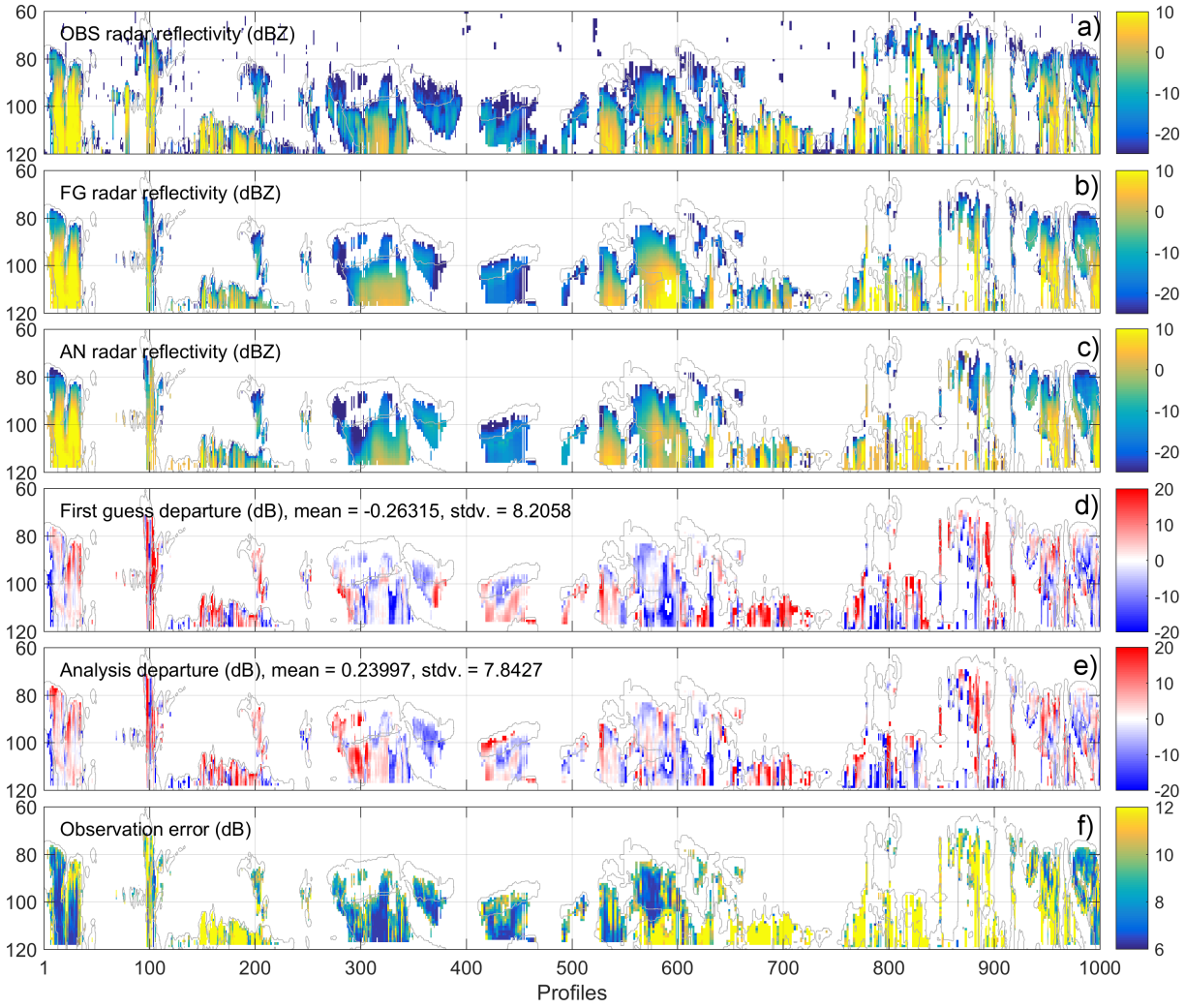

*Figure 4.11: Cross-sections of radar reflectivity related variables corresponding to the orbital track shown in Fig. 3.2c (21:00 UTC 31st July 2007). Panels show (a) observed CloudSat radar reflectivity (dBZ), (b) model equivalent (FG) radar reflectivity using the model background (dBZ), (c) model equivalent (AN) radar reflectivity using the model analysis from the assimilation experiment using all observations and radar and lidar (**ALL_RAD_LID**), (d) shows the first guess departure (panel (a) minus (b)), (e) shows the analysis departure (panel (a) minus (c)) and (f) shows the observation error (dB) assigned to the observations. Note that the first guess radar reflectivity is only displayed where there are hydrometeors detected in both model and observations. To elucidate the position of the model clouds, the first-guess model cloud boundaries are shown in grey.*

A visualisation of the cross-section for lidar backscatter is shown in Fig. 4.12. Whereas the initial fit of the first guess radar reflectivity to the observations was impressive, there are probably greater difference between the

superobbed CALIPSO lidar backscatter (Fig. 4.12a) and the first guess model equivalent (Fig. 4.12b). Because the lidar is more sensitive to cloud than the radar, it is apparent when the model cloud top boundary disagrees with CALIPSO, such as around the 100th and 350th profiles. However in other places the agreement in cloud top is good such as around the 475th and 850th profiles.

Where cloud top in the observations is much higher than in the model, a two-tone pattern can be seen in the first guess departures (Fig. 4.12d), for example around the 350th profile. Here the first guess departures tends to be positive at the top of the cloud and negative lower down; the additional cloud encountered by the CALIPSO lidar leads to a greater signal attenuation in the observations than in the model. This leads to difficulties in reducing the analysis departures (Fig. 4.12e) as although the cloud amount can be increased at the top of the cloud to match the observations, this will also correct the departures at the base of the cloud due to the increased attenuation, which could leave an excess of cloud in the model (as suggested by the larger negative radar reflectivity analysis departures in Fig. 4.11e). Assimilating the whole profile rather than just when there is cloud in both model and observations might help solve this problem.

There are situations where the lidar analysis departures are smaller than the first guess departures, such as around the 500th and 750th profiles. In total, for this transect, the mean difference in lidar backscatter compared to observations increases by 6% and the standard deviation increases by 4%. Further analysis of different situations where the lidar assimilation does not converge to the observations would help to diagnose issues with either the lidar forward model, error specification or screening criteria.

To understand how the information from radar reflectivity and lidar backscatter gets converted to more tangible cloud related variables such as temperature and relative humidity, Figs. 4.13 and 4.14 show analysis increments for four different assimilation experiments (all standard observations, **REF**; All observations plus radar and lidar, **ALL_RAD_LID**; radar only, **RAD**; and lidar only, **LID**). Comparing different assimilation experiments also allows us to infer whether information on these variables comes from the radar, lidar or other instruments in the observing system. The main differences **ALL_RAD_LID** experiment relative to the **REF** experiment can be summarised as:

- **ALL_RAD_LID** has decreased ice water and snow water content around the 300th profile compared to the **REF**, generated through a decrease in cloud fraction, relative humidity and slight relative increase in temperature.
- **ALL_RAD_LID** has an increase in ice cloud and snow water content around the 375th profile via an increase in cloud fraction, and relative decrease in temperature.
- **ALL_RAD_LID** tends to decrease the ice and snow water content of the cloud around the 450th profile, but increase the ice water content below the cloud, effectively lowering the cloud in the atmosphere. This is achieved via a strong decrease in temperature and increase in relative humidity below cloud base.

By considering the radar only and lidar only experiments, we can ascertain where the two measurement types are in agreement over the placement and magnitude of clouds and also the spread of information through the atmosphere. Increments in temperature and relative humidity within clouds tend to be greatest for the lidar only experiment, but at cloud base and below tend to be greatest for the radar only experiment. This is expected, as due to their different wavelengths, radar tends to detect more precipitation and lidar tends to detect more cloud. The increments broadly agree, apart from around the 450th profile where cloud fraction is decreased in the lidar only experiment while it is increased in the radar only experiment. This is due to differences in the cloud top height in the model background versus the CALIPSO data, whereas the increment in cloud fraction in the radar only experiment is generated through background error correlations rather than a direct measurement.

Non-local analysis increments in the vertical structure of temperature and relative humidity can be seen elsewhere, particular between the 300th and 400th profiles. It is not clear whether these are real and beneficial, or

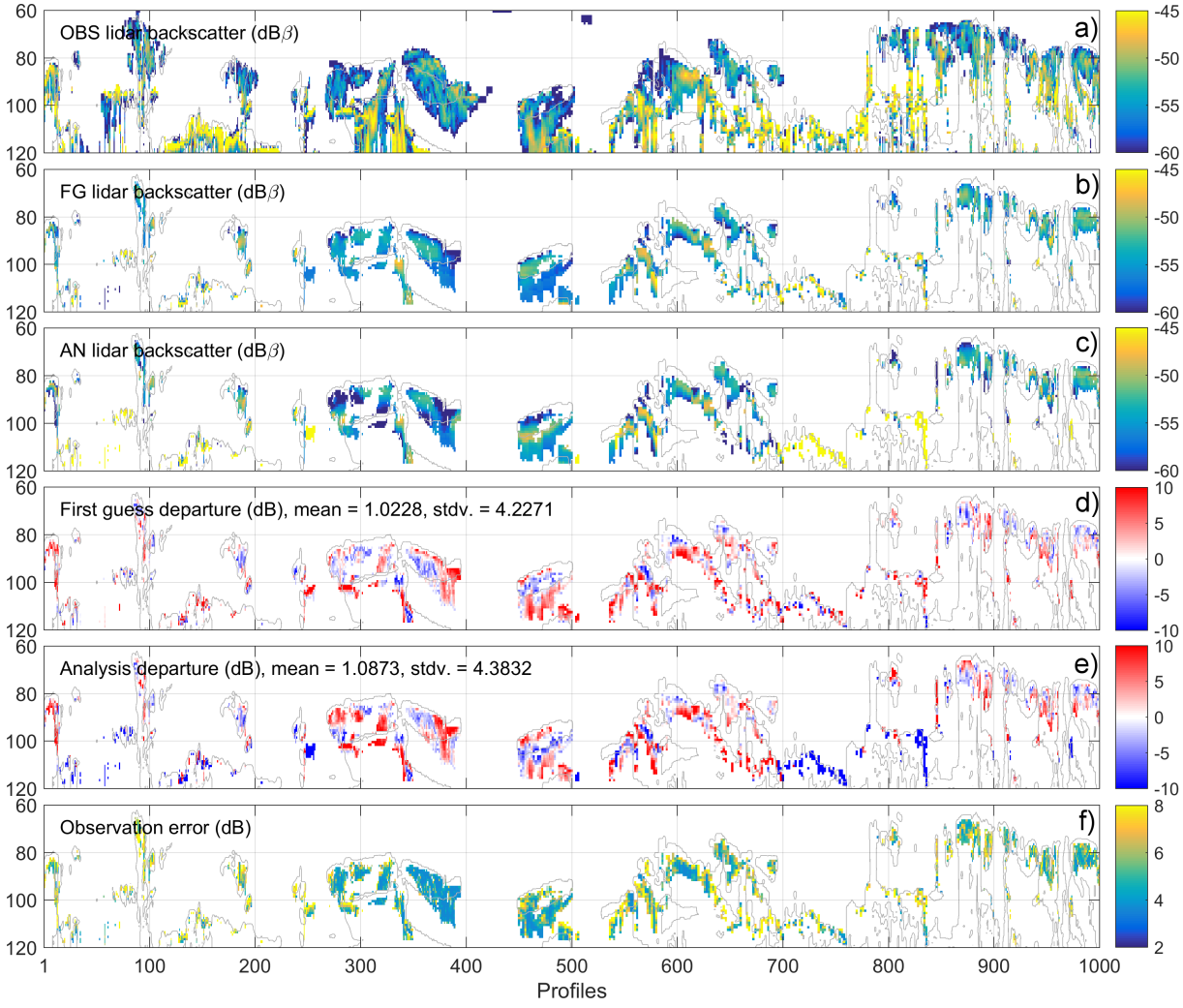

Figure 4.12: Same as Fig. 4.11, but for CALIPSO lidar backscatter

spurious artefacts of the background error correlation matrix required for 4D-Var. In general the increments above cloud in the **ALL_RAD_LID** experiment agree in sign with those of the radar only and lidar only experiments, indicating that they come from assimilated radar and lidar observations rather than just created by other assimilated observations.

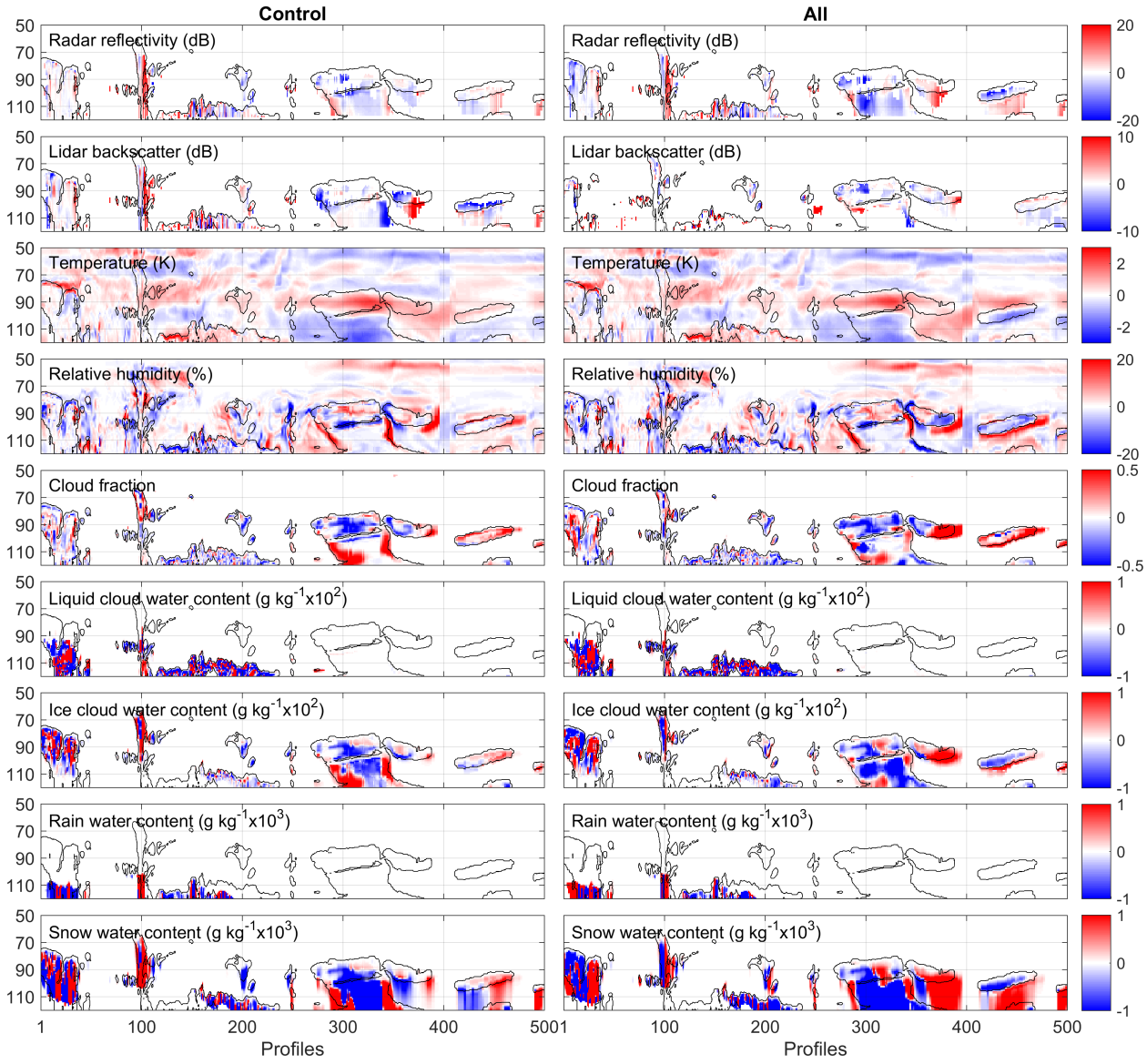

*Figure 4.13: Analysis increments for the first 500 profiles shown in Figs. 4.11 and 4.12 in the reference experiment (**REF**; left column) and all observations experiment (**ALL_RAD_LID**; right column). Panels from top to bottom show model-level increments of: radar reflectivity (dB), lidar backscatter (dB), temperature (K), relative humidity (%), cloud fraction, liquid cloud water content multiplied by 100 (gkg^{-1}), ice cloud water content multiplied by 100 (gkg^{-1}), rain water content multiplied by 1000 (gkg^{-1}) and snow water content multiplied by 1000 (gkg^{-1})*

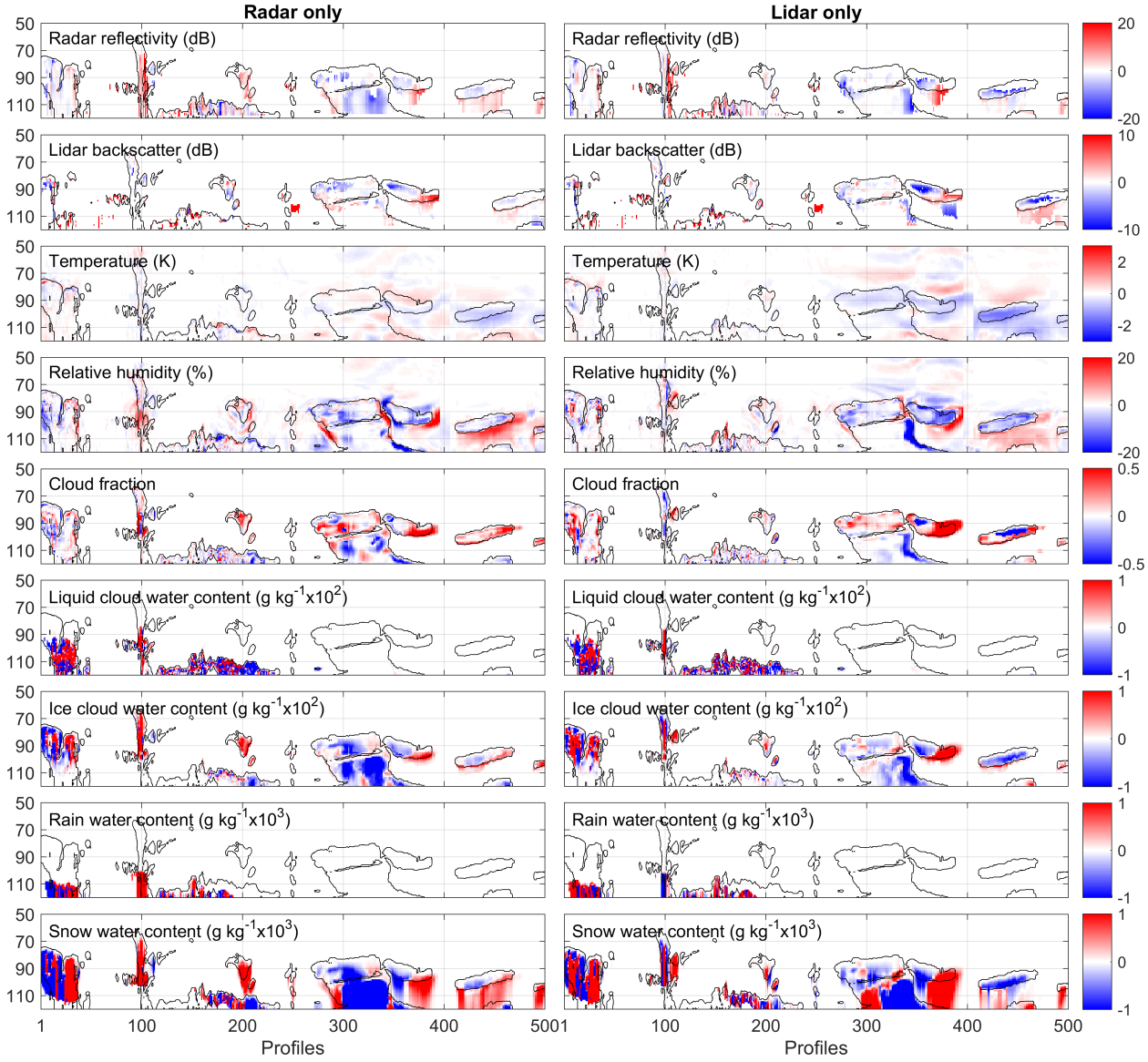

*Figure 4.14: Same as Fig. 4.13, but for the experiment where only radar observations are assimilated (***RAD***; left column) and the experiment where only lidar observations are assimilated (***LID***; right column).*

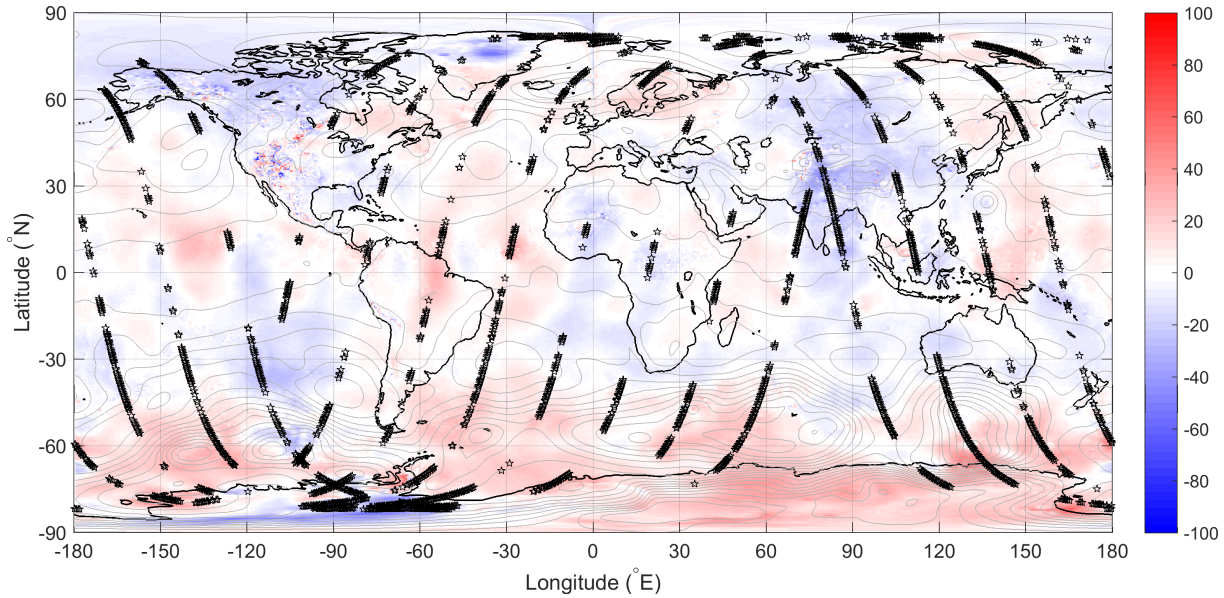

*Figure 4.15: Difference in analysis increments of mean sea level pressure (MSLP; Pa) between the reference experiment (***REF***) and the experiment where radar and lidar are also assimilated (***ALL_RAD_LID***). Black stars show gridboxes where at least one radar observation in the column has actively been assimilated. Isobars of MSLP are also plotted.*

In addition to the forecast model within 4D-Var, the background error correlation matrix spreads information from observations in the horizontal. To assess the structure of this horizontal information spread, Fig. 4.15 shows the difference in increments of mean sea level pressure (MSLP) between the **REF** experiment and **ALL_RAD_LID**. The greatest differences tend to be seen where there are the greatest density of radar and lidar observations, and a tilted pattern in the increments relative to the latitudinally ascending or descending tracks of the polar-orbiting CloudSat and CALIOP satellites can be seen.

4.2 Impact on the subsequent forecast

4.2.1 Verification of forecast against operational analyses

In this section the impact of the assimilation of cloud radar and lidar observations on the skill of subsequent forecasts is evaluated. The evaluation considers differences in the rms forecast error between the forecasts initiated by the **ALL_RAD_LID** experiment and the forecast initiated by the reference experiment for one assimilation cycle with the analysis time 00:00 UTC 1 August 2007. Both experiment's forecast errors are with respect to the operational analysis interpolated to the resolution used in the experiments (TCo639). Zonal means of these rms error differences are shown for specific humidity, temperature and zonal wind in Fig. 4.16 for different forecast lead times. Two key features stand out when analysing the impact of radar and lidar assimilation on the zonal means:

- Firstly the forecast generated using the **ALL_RAD_LID** analysis shows an increase in forecast skill of temperature and wind, with the greatest increases in three regions; one in the NH extra-tropics, one just north of the equator corresponding to the Inter-Tropical Convergence Zone (ITCZ), and a third in the SH extra-tropics. These three regions correspond to locations with the greatest quantity of cloud and hence radar and lidar observations (see Fig. 3.2). Any forecast improvement in specific humidity is difficult to detect, possibly due to the inherent noisy nature of the field. It could also be related with difficulties in the correct specification of the model background, which spreads the information of humidity away

from the cloud radar and lidar observations. If the correlations between different model layers are not representative of the situation, it could lead to spurious increments of specific humidity elsewhere in the model.

- Secondly, improvement in the forecast (indicated by the blue shaded areas) tends to increase with forecast lead time, particularly for wind and temperature. One reason is that using the operational analysis as a reference could favour the reference forecast, which was generated with a similar observation set that did not contain observations from cloud radar and lidar. Another possible reason is that the difference in forecast errors could be influenced by the fact the operational analysis was generated using a different model cycle with different physics, horizontal and vertical resolutions. Using the **ALL_RAD_LID** analyses as the 'truth' would mitigate this and will be explored in the future. Using more assimilation cycles for the analysis would also make the findings more robust.

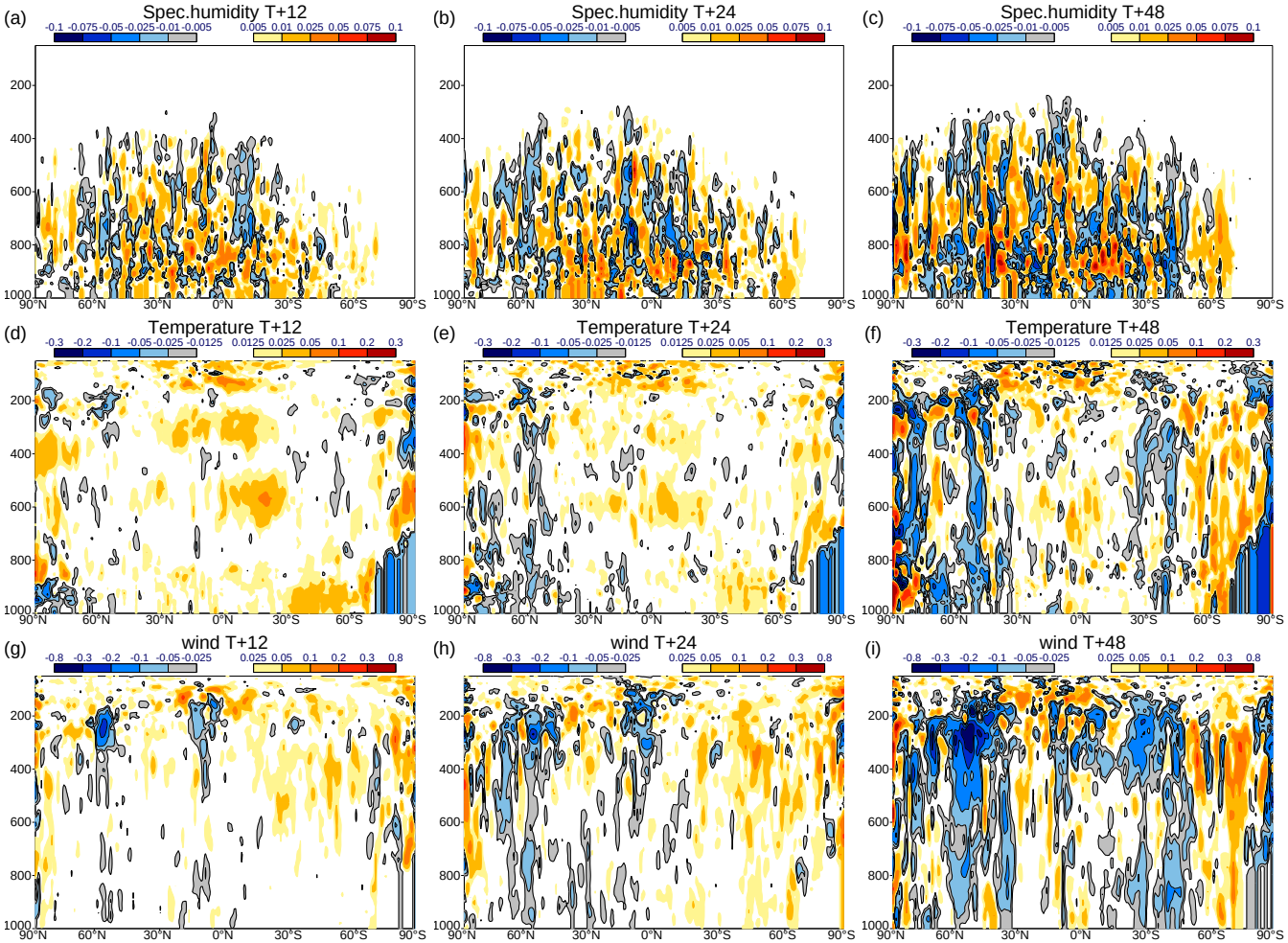

Figure 4.16: Zonal mean of differences of (a,b,c) specific humidity ($g.kg^{-1}$), (d,e,f) temperature (K) and (g,h,i) wind ($m\ s^{-1}$) rms errors for the differences between the forecasts starting from analysis created by 4D-Var assimilation of cloud radar reflectivity and lidar backscatter observations with doubled errors and the operational analysis and between the forecast starting from the reference analysis and the operational analysis. (a,d,g) 12-hour, (b,e,h) 24-hour and (c,f,i) 48-hour forecasts. Situation **2007080100** *with 12-hour assimilation period between 31 July 2007 21:00 UTC and 1 August 2007 09:00 UTC. Reduction (resp. increase) of rms errors for the experimental run is shown with blue (resp. red) shadings.*

Quantitatively, by averaging across all latitudes, Fig. 4.17 shows global profiles of specific humidity, temperature and zonal wind as considered in Fig. 4.16. As discussed, the greatest impact in improvements of forecast

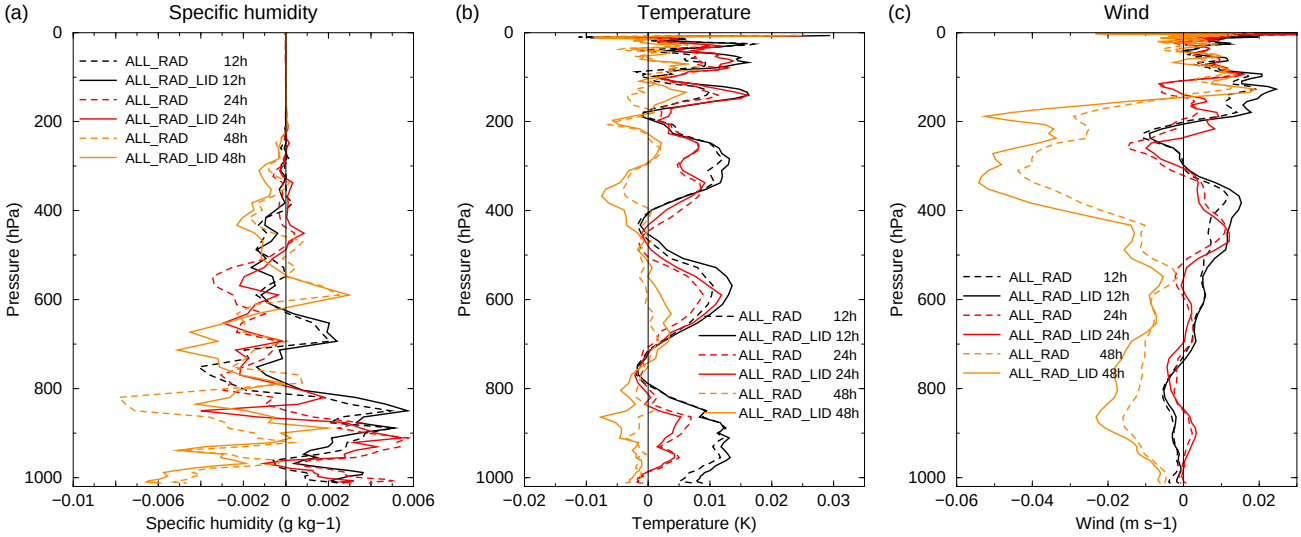

*Figure 4.17: Differences of rms errors for (a) specific humidity ($g.kg^{-1}$), (b) temperature(K) and (c) wind ($m\ s^{-1}$) as in Fig. 4.16, presented as global values for 12, 24 and 48h forecasts (see legend) and for experiments with analysis created by 4D-Var assimilation of either cloud radar reflectivity in combination with all other observations (***ALL_RAD****, dashed lines) or both the cloud radar reflectivity and lidar backscatter with other observations (***ALL_RAD_LID****, solid lines).*

errors by radar and lidar assimilation is seen for the longer 48 hour lead time. While not necessarily apparent in Fig. 4.16, the forecast skill in predicting specific humidity also improves with forecast lead time for the **ALL_RAD_LID** forecast. Figure 4.17 also reveals that assimilating both radar and lidar has greater benefit to the forecast than radar only. The greatest benefit of the lidar can be seen in the 48 hour forecast for zonal wind.

Concentrating on a 24 hour lead time, Fig. 4.18 shows the effect of increasing or decreasing errors on the globally averaged profiles of specific humidity, temperature and zonal wind. In agreement with the degradation of the analysis fit to wind observations when decreasing radar and lidar observation errors seen in Fig. 4.7a, a clear decrease in forecast skill is apparent for all three variables. One reason for the degradation is that by decreasing the observation errors for radar and lidar, the model is forced to drawn away from other assimilated observations elsewhere in the model that contain useful information. Another reason is that fitting the model too closely to the observations can actually pull the model analysis further from the truth, particularly for profiling cloud radar and lidar observations that can have large representativity errors and uncertainties in the forward models. This range in forecast skills produced by a crude doubling or halving of errors highlights how much potential improvement could be gained from careful adjustments to the observation error and will be a priority for future investigations.

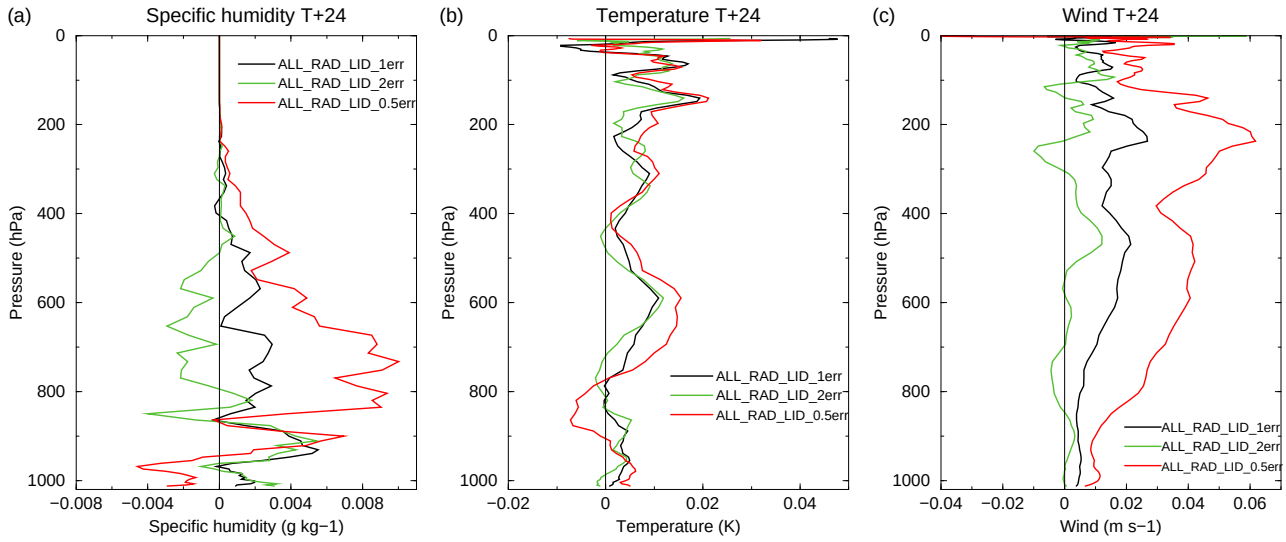

*Figure 4.18: Differences of rms errors for (a) specific humidity ($g.kg^{-1}$), (b) temperature(K) and (c) wind ($m\ s^{-1}$) as in Fig. 4.16, presented as global values for 24h forecasts and for experiments with analysis created by 4D-Var assimilation of cloud radar reflectivity and lidar backscatter observations with single (***1err***, black solid line), double (***1err***, green solid line) and (***0.5err***, red solid line).*

4.2.2 Verification against assimilated observations

An alternative to verifying forecasts against model analyses is to verify them against observations. While at analysis time and for very short forecasts, the observations might be correlated to the analysis, for longer forecast lead times the observations used in the verification will be increasingly independent. Therefore, the observations should not suffer from the potential problems with verifying against analyses discussed in the previous section. However, one downside is that the observation coverage is limited compared to verifying against an analysis, so any signals are likely to require longer cycling times to remove noise.

Figure 4.19 shows a comparison of the standard deviation in differences between zonal wind observations from radiosonde observations and the corresponding forecast zonal wind forecast from the **ALL_RAD_LID** experiment normalised by the reference experiment. In agreement with the findings from previous sections, the double error experiment (Fig. 4.19b) produces a better fit to observations than the single error experiment (Fig. 4.19a). Considering the radiosonde observations in the Northern hemisphere, where the observations are densest, a signal for increased skill in predicting tropospheric winds compared to the reference is seen for the 24 hour forecasts. Contrary to the results of Fig. 4.18, the skill in 48 hour forecasts is reduced, but still comparable to the reference forecast. Statistics for the tropics and the Southern hemisphere are noisy, suggesting significantly more assimilation cycling is required to draw any significant conclusions.

A similar pattern of increased skill in the shorter term forecasts relative to the reference can be seen in radiosonde temperature and humidity observations (Fig. 4.20). Although the signals are again noisy, the greatest improvements in forecast skill are seen in the Southern hemisphere for both temperature and humidity. Improvements are also seen in specific humidity in the lower troposphere for 24 hour forecasts.

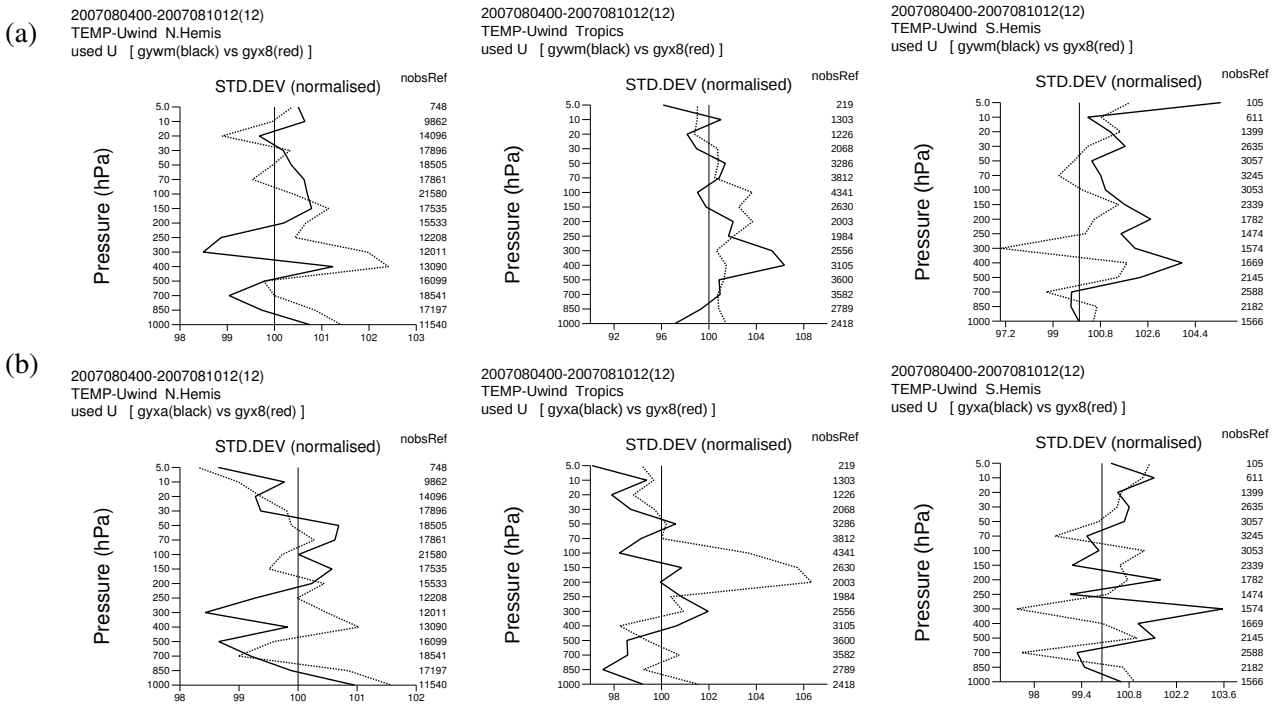

Figure 4.19: Normalised standard deviation of forecast departures with respect to TEMP zonal wind observations for 24h (solid line) and 48h (dotted line) forecasts shown for the experiments assimilating cloud radar reflectivity and lidar backscatter observations using (a) single and (b) double observation errors and combined with all other routinely assimilated observations. Results are shown for (left) Northern Hemisphere, (middle) Tropics and (right) Southern Hemisphere. The number of observations (nobsRef) for the period from 4 August 2007 00:00 UTC to 10 August 2007 12:00 UTC is displayed at the right side of each profile.

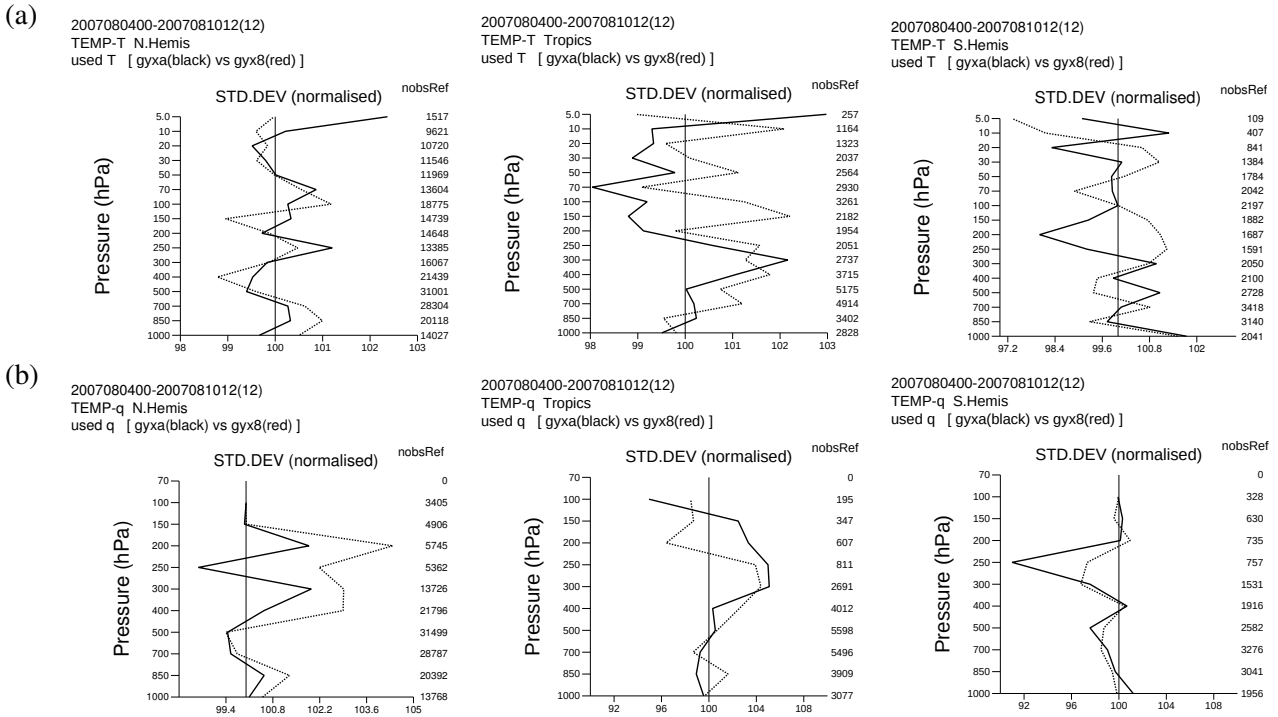

Figure 4.20: Same as Fig. 4.19, but for the forecast departures with respect (a) temperature an (b) specific humidity shown for the experiments assimilating cloud radar reflectivity and lidar backscatter observations using double observation errors.

4.2.3 Assessment of rain rates in the tropics using independent observations

Due to the vast quantity and diversity of observations assimilated in the ECWMF 4D-Var system, it can be difficult to find independent datasets to verify the analysis and forecast performance. However the Tropical Rainfall Measurement Mission (TRMM) provides a useful dataset that was not assimilated at ECMWF in the satellite's lifetime. For simplicity we have chosen to use the TRMM level 2 near-surface rain rate product as the model equivalent surface rain-rate is readily output from the IFS forecast model. In the future, a TRMM forward model of radar reflectivity could be used to take advantage of the vertically-resolved nature of TRMM's precipitation radar and also allow a more direct comparison between model and observations (in observation space). Using an in-house forward model would allow for greater control over the assumptions made in the evaluation and remove any dependence on the forward model used by the TRMM level-2 processing (in particular allowing consistency with the current CloudSat and CALIPSO forward models).

To compute the statistics, both the TRMM near-surface rainfall product and the model surface rainfall are averaged to a common 0.5x0.5 degree grid and three-hourly temporal resolution. The TRMM satellite has an orbital range of -35 to 35 degrees latitude as can be seen in Fig. 4.21 and the precipitation radar (PR) has a 250 km swath width. As the PR had a minimum detectable signal is around 18 dBZ, the evaluation is restricted to precipitation events greater than 0.4 mm h^{-1}. Forecasts are generated from the reference experiment, **ALL_RAD_LID** and **ALL_RAD_LID** with double errors up to a lead time of 72 hours and then compared to the TRMM observations in 3 hourly steps. The evaluation uses 14 forecasts initiated from the analyses in 12-hour intervals between 00:00 UTC 1st August 2007 to 12:00 UTC 7th August 2007.

Figure 4.22 shows that the RMS in short-term surface rain rate forecasts over the tropics compared to TRMM is reduced by around 2 % when assimilating CloudSat radar reflectivity and CALIPSO lidar backscatter. The RMS is reduced a further 1 % when using the forecasts initalised with the double error analysis, in agreement with the improved fit to other observation types (e.g., radiosonde) when using double errors. Interestingly, a diurnal cycle in the RMS can be observed due to the increased convection over land during the afternoon, predominantly over Africa. To smooth this signal, a 24 hour moving window is also used to compute the RMS (dashed line). From these smoothed lines, we can see the greatest reduction in RMS is seen between lead times

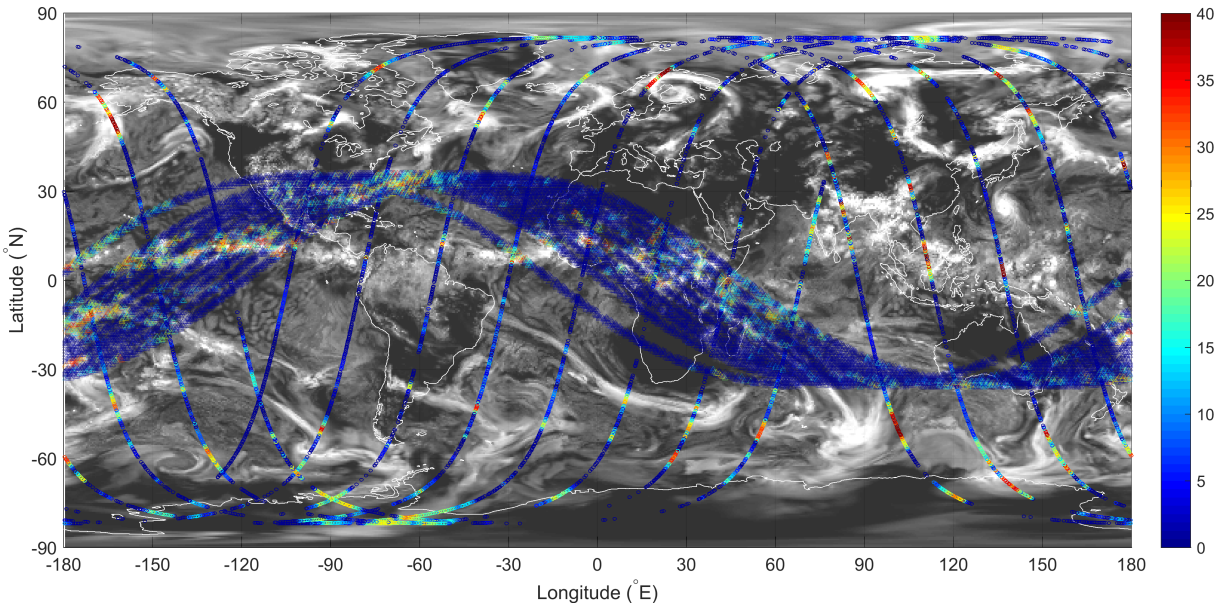

Figure 4.21: Same as Fig. 3.2, but with TRopical Measurement Mission (TRMM) observations of near-surface rain rate overlayed.

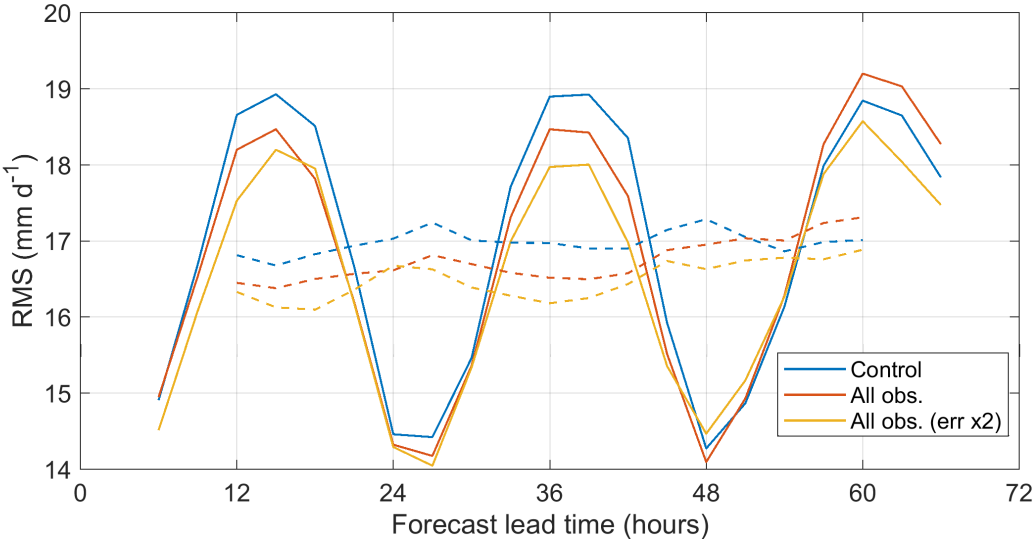

*Figure 4.22: Root-mean-square error (RMS) between TRMM and forecast near-surface rain rates for varying forecast lead times of up to 3 days. Statistics are generated from forecasts initialised at 00:00 UTC and 12:00 UTC using the reference analysis (***REF***; blue) and all observations analysis (***ALL_RAD_LID***; red). The solid lines indicate a 12 hour averaging window, while the dashed lines indicate a 24 hour averaging window.*

of 12 to 48 hours. The shallow increase in RMS for increasing forecast lead times confirms the difficulties in predicting rainfall in the tropics (e.g., Chakraborty, 2010).

5 Summary

In this report we have demonstrated, for the first time, the feasibility of assimilating CloudSat and CALIPSO observations into a global NWP model. In general, including cloud radar reflectivity and lidar backscatter in the assimilation system was found to have a positive impact on both the analysis fit to observations, and the subsequent short-term forecast. As found in Janisková (2015), instrument-by-instrument, the radar was shown to provide the largest impact on forecast errors, but a combination of radar and lidar gave the greatest total benefit. The analysis generated by including radar and lidar observations on top of routine observations was found to decrease the standard deviation of FG departures to radiosonde and aircraft observations relative to only assimilating routine observation. Improvements were also seen to the fits of temperature and specific humidity. The greatest impact of the new observations on forecast skill was also found on zonal wind, relative to both operational analyses and radiosonde observations. A further promising result that showed the forecast of rain rates in the tropics could be improved by assimilating cloud radar and lidar observations following an evaluation with independent TRMM radar measurements of near surface rainfall.

Gaining benefit on forecast skill by including new observations into a well-established observing system is extremely difficult, so the results presented here are extremely promising and warrant the opening of many avenues of further research that were not able to be explored here. The sensitivity of the results to the prescribed observation error was shown to be large and it is envisaged that relatively easy gains in forecast skill would be achievable through careful tuning. The behaviour of the assimilation system for different regimes, for example the effect of cloud radar and lidar on convective situations, requires further work and could result in improvements in the forward operator assumptions or screening criteria. Another line of potentially very fruitful research is to investigate the synergistic benefit of cloud radar and lidar observations to other observation types related to clouds, in particular the all-sky radiance assimilation framework in use at ECMWF. Finally, no attempt was made to optimise the pre-processing of the observations, so investigations of the averaging scale and possible thinning of the observations in both the horizontal and vertical would be beneficial.

Acknowledgments

The authors would like to thank Philippe Lopez, Robin Hogan, Alan Geer and Stephen English for invaluable scientific discussions relating to the project. Peter Lean, Gabor Radnoti and Tomáš Král are thanked for their technical help in adapting the 4D-Var assimilation system to process radar and lidar observations and perform verification experiments. Mohamed Dahoui is also thanked for his help with generating observation statistics and plotting. Finally Draško Vasiljević and Marijana Crepulja are thanked for their help with the offline observation processing.

The NASA CloudSat Project is kindly acknowledged for providing the CloudSat data. The authors are also grateful to the NASA Langley Research Center - Atmospheric Science Data Center for making the CALIPSO data available.

List of Acronyms

1D-Var/4D-Var	One-Dimensional/Four-Dimensional Variational Assimilation
AN	Analysis
AIREP	AIRcraft Weather REPort
AVHRR	Advanced Very High Resolution Radiometer
BC	Bias Correction
BUFR	Binary Universal Form for the Representation of meteorological data
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CLASS	Comprehensive Large Array-data Stewardship System
CloudSat	NASA's cloud radar mission
EarthCARE	Earth, Clouds, Aerosols and Radiation Explorer
ECMWF	European Centre for Medium Range Weather Forecasts
ESA	European Space Agency
ESTEC	European Space Research and Technology Centre
FG	First Guess
FOV	Field Of View
IFS	Integrated Forecasting System of ECMWF
ITCZ	Inter Tropical Convergence Zone
MS	Multiple Scattering
MSLP	Mean Sea Level Pressure
NASA	National Aeronautics and Space Administration
NCEP	National Center for Environmental Prediction
NH	North Hemisphere
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OBS	OBServations
PDF	Probability Distribution Function
PR	Precipitation Radar
PSD	Particle Size Distribution
rms/RMS	root mean square error
SH	South Hemisphere
STSE	Support-to-Science-Element
TCo639	Model cubic octahedral grid with spectral truncation T639
TRMM	TRopical Rainfall Measurement Mission
UTC	Universal Time Coordinated
WP	Work Package

References

Battaglia, A., T. Augustynek, S. Tanelli, and P. Kollias, 2011: Multiple scattering identification in space-borne w-band radar measurements of deep convective cores, *Journal of Geophysical Research: Atmospheres*, **116**(D19), n/a–n/a, D19201.

Chakraborty, A., 2010: The skill of ecmwf medium-range forecasts during the year of tropical convection 2008, *Monthly Weather Review*, **138**(10), 3787–3805.

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

Di Michele, S., E. Martins, and M. Janisková, 2014a: Observation operator and observation processing for cloud radar, WP-1100 report for the project Support-to-Science-Element STSE Study - EarthCARE Assimilation, AO/1-6441/10/NL/CB, 59 pp.

Di Michele, S., E. Martins, and M. Janisková, 2014b: Observation operator and observation processing for cloud lidar, WP-1200 report for the project Support-to-Science-Element STSE Study - EarthCARE Assimilation, AO/1-6441/10/NL/CB, 40 pp.

Fielding, M. and M. Janisková, 2017: Observation quality monitoring and pre-processing, WP-2000 report for the project Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction, ESA ESTEC contract 4000116891/16/NL/LvH, 61 pp.

Fisher, M., 2004: Generalized frames on the sphere, with application to the background error covariance modelling, *Proc. Seminar on Recent Developments in Numerical Methods for Atmospheric and Ocean Modelling, Reading, UK, ECMWF*, pp. 87–102.

Gauthier, P. and J.-N. Thépaut, 2001: Impact of the digital filters as a weak constraint in the preoperational 4D-Var assimilation system of Météo-France, *Mon. Weather Rev.*, **129**, 2089–2102.

Illingworth, A. et al., 2015: The earthcare satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation, *Bulletin of the American Meteorological Society*, **96**(8), 1311–1332.

Janisková, M., 2015: Assimilation of cloud information from space-borne radar and lidar: Experimental study using 1D+4D-Var technique, *Q. J. R. Meteorol. Soc.*, **141**, 27082725, doi:10.1002/qj.2558.

Janisková, M., S. Di Michele, and E. Martins, 2014: Support-to-Science-Elements (STSE) Study - EarthCARE Assimilation, ESA Contract Report on Project 4000102816/11/NL/CT, 225 pp.

Janisková, M., M. Fielding, M. Crepulja, D. Vasiljević, T. Král, and P. Lean, 2017: Assimilation system development for cloud radar and lidar observations, WP-3000 report for the project Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction, ESA ESTEC contract 4000116891/16/NL/LvH, 26 pp.

Janisková, M. and P. Lopez, 2013: Linearized physics for data assimilation at ECMWF, in *S.K. Park and L. Xu (Eds), Data Assimilation for Atmospheric, Ocean and Hydrological Applications (Vol II), Springer-Verlag Berline Heidelberg*, pp. 251–286, doi:10.1007/978–3–642–35088–7–11.

Janisková, M., P. Lopez, and P. Bauer, 2012: Experimental 1D+4D-Var assimilation of CloudSat observations, *Q. J. R. Meteorol. Soc.*, **138**, 1196–1220, doi:10.1002/qj.988.

Janisková, M., O. Stiller, S. Di Michele, R. Forbes, J.-J. Morcrette, M. Ahlgrimm, P. Bauer, and L. Jones, 2010: QuARL - Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles, ESA Contract Report on Project 21613/08/NL/CB, 329 pp.

Kollias, P., E. E. Clothiaux, M. A. Miller, B. A. Albrecht, G. L. Stephens, and T. P. Ackerman, 2007: Millimeter-wavelength radars: New frontier in atmospheric cloud and precipitation research, *Bulletin of the American Meteorological Society*, **88**(10), 1608–1624.

Liebe, H. J., 1985: An updated model for milimeter wave propagation in moist air, *Radio Science*, **20(5)**, 1069–1089.

Liebe, H. J., P. Rosenkranz, and G. A. Hufford, 1992: Atmospheric 60-GHz oxygen spectrum - new laboratory measurements and line parameters, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **48**, 629–643.

Liu, Z., M. Vaughan, D. Winker, C. Kittaka, B. Getzewich, R. Kuehn, A. Omar, K. Powell, C. Trepte, and C. Hostetler, 2009: The CALIPSO lidar cloud and aerosol discrimination: Version 2 algorithm and initial assessment of performance, *J. Atmos. and Ocean. Tech.*, **26(7)**, 1198–1213.

Mahfouf, J.-F. and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Part II: Experimental results with improved physics, *Q. J. R. Meteorol. Soc.*, **126**, 1171–1190.

Platt, C., 1973: Lidar and radioinetric observations of cirrus clouds, *Journal of the Atmospheric Sciences*, **30**(6), 1191–1204.

Rabier, H., F.and Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics, *Q. J. R. Meteorol. Soc.*, **126**, 1143–1170.

Stephens, G., D. Vane, R. Boain, G. Mace, K. Sassen, Z. Wang, A. Illingwort, E. O'Connor, W. Rossow, and S. Durden, 2002: The CloudSat mission and the A-train, *Bull. Am. Meteorol. Soc.*, **83(12)**, 1771–1790.

Vaughan, M., K. Powell, R. Kuchn, S. Young, D. Winker, C. Hostetler, W. Hunt, Z. Liu, M. McGill, and B. Getzewich, 2009: Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements, *J. Atmos. and Ocean. Tech.*, **26(10)**, 2034–2050.

Winker, D., M. Vaughan, A. Omar, Y. Hu, K. Powell, Z. Liu, W. Hunt, and S. Young, 2009: Overview of the CALIPSO mission and CALIOP data processing algorithms, *J. Atmos. and Ocean. Tech.*, **26(7)**, 2310–2323.