ESA Contract Report

PEARL Cloud contract 4000128669/19/NL/CT

Contract Report to the European Space Agency

Optimising impact of radar reflectivity and lidar backscatter observations on analysis

WP-2000 report

Preparations for EarthCARE Assimilation - Radar and Lidar Cloud Observations (PEARL Cloud)

Authors: Mark Fielding and Marta Janisková

Contract officer: Tobias Wehr

October 2020

ESA Contract Report ESA Contract Report ESA Contract Report

Series: ECMWF ESA Contract Report Series

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

Contact: library@ecmwf.int

© Copyright 2023

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

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. The content of this document is available for use under a Creative Commons Attribution 4.0 International Public License.
See the terms at *https://creativecommons.org/licenses/by/4.0/*.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.

ABSTRACT

This report provides a description of updates to the ECMWF data assimilation system to improve the impact of cloud radar reflectivity and lidar backscatter on the analysis. Updates include a new triple-column approach for accounting for sub-grid cloud inhomogeneity and a model for observation error correlations. The report begins with a brief description of the data assimilation system at ECMWF to be used for direct assimilation of cloud radar and lidar observations. Next, to study and optimize the impact of cloud radar and lidar observations on the analysis, 4D-Var assimilation experiments have been performed using Cloudsat cloud radar reflectivity and CALIPSO lidar backscatter observations. The first experiments were performed using the ECMWF model cycle CY43R1, for which the system of cloud radar and lidar observations was prepared at first. The results obtained indicate that 4D-Var analyses get closer to these observations and the impact on other assimilated observations is mainly neutral. Generally, the impact of the cloud radar reflectivity is larger than of the lidar backscatter. The report then summarizes updates of the data assimilation system for cloud radar and lidar observations in the more recent model cycle CY46R1. As part of the updates, the observation operator is adapted to model sub-grid condensate variability and a framework for accounting for the vertical correlation in observation error is developed. The impact of these observations in the updated system is positive; the updated observation operator improves the analysis fit to the observations. The fit to AMSU-A observations is also improved compared to both CY43R1 and an experiment using the old single-column version of the observation operator.

Contents

1 Introduction

The previous assimilation project at European Centre for Medium-Range Weather Forecasts (ECMWF) (Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction, Janisková and Fielding, 2018b) focused on developments towards direct assimilation and monitoring to exploit cloud radar and lidar data for their assimilation in NWP models. Including CloudSat (Stephens et al., 2002) and CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations Winker et al., 2009) data, the project demonstrated the feasibility to assimilate such observations using Four-Dimensional Variational (4D-Var) data assimilation of system. It also prepared grounds for assimilation of such observational types from the future Earth, Clouds, Aerosols and Radiation Explorer (EarthCARE) mission. The successful real-time assimilation of EarthCARE observations relies on a number of further tasks. One of such tasks is optimisation of impact coming from these observations on analysis and this is investigated in this work package.

In this work package, impact of cloud radar and lidar observations on analysis have been first investigated using the model cycle CY43R1 of the ECMWF Integrated Forecasting System (IFS). This is the cycle where all developments required for assimilation of these observations was included for the first during the previous ESA founded project (Janisková and Fielding, 2018b). In that project, assimilation experiments were only run for shorter period, maximum of 10 days, just to demonstrate feasibility of 4D-Var assimilation system to use cloud radar and lidar observations (Janisková and Fielding, 2018a). This study found that including cloud radar reflectivity and lidar backscatter in the assimilation system indicated a positive impact on both the analysis for to observations, and the subsequent short-term forecast. These first results were extremely promising. However, gaining more benefit on forecast skill by including new observations into a well-established observing system as used at ECMWF is extremely difficult and further research and studies need to be done to optimise impact of these observations on analysis and on forecast. Therefore assimilation experiments have been extended up to three months to investigate impact of these observations for longer period and to get indications where further research should be concentrated. Although, the first experiments were done using the model cycle CY43R1, the cycle used at ECMWF at the beginning of this project in October 2019, once adaptation of the assimilation system for cloud radar and lidar observations to higher model cycle, CY46R1, was concluded, experiments started to be carried out using this new cycle. Therefore results using two model cycles, CY43R1 and CY46R1, will be presented in this report.

This report summarizes outcomes from the experiments carried out and updates done in order to optimise impact of cloud radar and lidar observations on the model analysis. Section 2 provides a brief description of the used data assimilation system and also an information on observations and their handling in the system. Section 3 discusses the results from analysis studies carried out using the model cycle CY43R1. Updates related to cloud radar and lidar assimilation are discussed in Section 4. The results from experiments done in the updated cycle, CY46R1, are presented in Section 5. A brief summary and perspectives are provided in Section 6.

2 Methodology

2.1 The ECMWF operational 4D-Var

As described in WP-1000 report, the current operational data assimilation system used at ECMWF is 4D-Var based on incremental formulation (Courtier et al., 1994; Rabier et al., 2000). The cost function, $\mathscr{J}$, in the incremental approach is formulated as follows:

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

In this formulation, $\mathscr{J}^b(\delta\mathbf{x}_0)$ is the background cost function which measures the distance between the initial state of the model $\mathbf{x}_0$ and the background $\mathbf{x}_0^b$ obtained from a short-range forecast valid at the initial time of the assimilation period. $\mathscr{J}^o(\delta\mathbf{x}_0)$ is the observation cost function measuring the distance between the model trajectory and corresponding observations. A constraint cost function $\mathscr{J}^c(\delta\mathbf{x}_0)$ is used in 4D-Var to control fast gravity waves using the digital filter approach developed by Gauthier and Thépaut (2001). $\mathbf{R}_i$ is the observation error covariance matrix (including measurement and representativeness errors) and $\mathbf{B}$ is the background error covariance matrix of the state $\mathbf{x}^b$ and is based on a wavelet formulation (Fisher, 2004) to introduce regime-dependent error statistics. H_i' is the linearized observation operator providing the model equivalent to the observations and it also includes the spatial interpolations to observation locations as well as the propagation of the initial state to each observation time using the forecast model. $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^b$ is the analysis increment and represents 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. Finally, $\mathbf{d}_i = \mathbf{y}_i^o - H_i(\mathbf{x}^b{}_i)$ is the so-called innovation vector providing the departure of the model background equivalent from the observation ($\mathbf{y}_i^o$).

2.2 Experimental framework

Experiments were performed using cycles 43r1 and 46r1 of ECMWF's operational assimilation and forecasting system. Global analyses and forecasts of variables such as pressure, temperature, humidity, cloud and precipitation are produced two times a day using an incremental formulation of 4D-Var (Courtier et al., 1994; Rabier et al., 2000) with a 12 h time window. A brief description of the data assimilation system is provided in WP-1000, Assimilation system adaptation and maintenance (Janisková and Fielding, 2022).

The atmospheric model has a semi-Lagrangian, spectral formulation with 137 levels from the surface to an altitude of $\sim$ 75 km. In the configuration used for our experiments, the high resolution trajectory has a horizontal resolution of TCo639 (corresponding to approximately 18 km on a cubic octahedral grid). To account for some non-linearity in the model and observation operators, the minimization consists of three inner loops all at a reduced resolution: the first at spectral resolution with truncation T95 (corresponding to $\sim$ 200 km), the second at T159 ($\sim$ 125 km) and the third at T255 ($\sim$ 80 km). 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). A total of 40 million observations are processed and used daily at

ECMWF. The vast majority of these are satellite measurements (around 90 satellite data products, but ECMWF Integrated Forecasting System (IFS) also benefits from observations from non-satellite sources, including surface-based and aircraft reports.

2.3 Observations and their handling for cloud radar and lidar assimilation

2.3.1 Cloud radar and lidar observations

In our study, measurements of cloud radar reflectivity (in dBZ), from the CloudSat 94 GHz radar and 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 data is obtained using the level-2 (L2) CloudSat geometric profile product (Marchand et al., 2008) and the Cloud-Aerosol Lidar with Orthogonal Polarisation (CALIOP) L2 lidar 5km cloud profile data version 3 (Winker et al., 2009; Young et al., 2013). 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, in our experiments the observations from these instruments are horizontally averaged to model gridboxes with resolution corresponding to approximately 36 km or 72 km.

2.3.2 Observation operator, errors and quality control

A brief summary of the observation operators providing model equivalent to the observations is only provided here. Unattenuated lidar backscatter and radar reflectivity are determined as a function of model hydrometeor water content and temperature in off-line computation which leads to the creation of look-up table. This table is searched, using a bilinear interpolation, for each model atmospheric profile and level where the model-equivalent radar reflectivity or lidar backscatter are required. To reduce computational cost, a parameterized version of the look-up table using a two-variable two-degree polynomial fit was also developed. Attenuation is then accounted for by integrating from the top model level to the bottom level and summing the extinction coefficients which are also a function of water content and temperature. For radar, profiles of radar reflectivity where multiple scattering is likely to be significant are detected using the integrated reflectivity (Battaglia et al., 2011) and excluded. For lidar, multiple scattering is approximated using a tuned version of the Platt approximation.

The observation errors due to instrument error, forward model error and representativity error are all calculated separately using an 'error inventory' approach. Under the hypothesis of uncorrelated errors, the individual components are added in quadrature to define the total observation error. Due to the narrow field of view of the space-borne cloud radar and lidar observations, an appropriate definition of the representativity error is required to account for the mismatch of scales between the observations and the model as described in Fielding and Stiller (2019).

The quality control for the observations is based on first-guess (FG) departures, thresholds for indicators of signal strength, estimated total attenuation and, for the radar, expected multiple scattering. The bias is based on a climatology of FG departures when the observations were passively monitored. To provide an implicit regime dependence for the correction scheme, indicators of temperature and heigh are used by the scheme.

Full details of the observation operators and observation handling are described in Fielding and Janisková (2020).

2.3.3 *Observation coverage*

Typical data coverage for cloud radar and lidar observations for the 12h period corresponding to the length of the assimilation window is shown in Fig. 2.1 displaying multiple orbits of CloudSat and CALIPSO satellites. Despite the narrow footprint of the observations, for each assimilation window around half of the model gridboxes are within 500 km of an A-train overpass. Over the period of one month (Fig. 2.2) nearly the whole globe is densely cover by cloud radar and lidar observations.

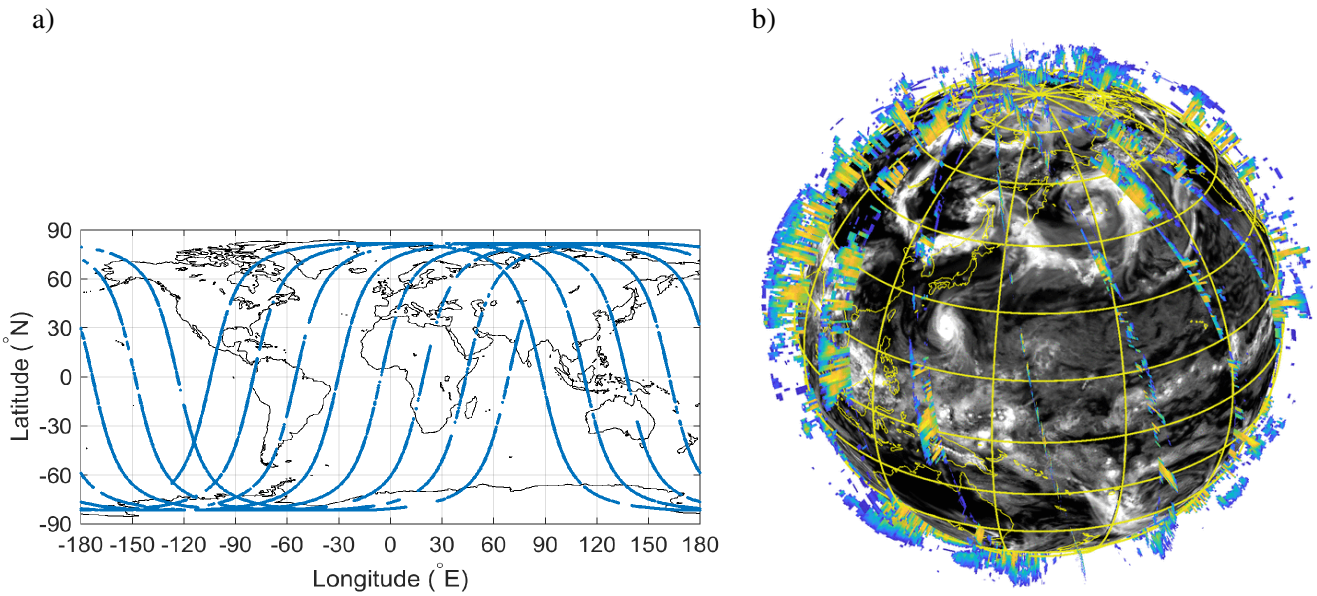

Figure 2.1: Example of CloudSat/CALIPSO data coverage for the 12h assimilation window shown for the period between 31 July 2007 21:00 UTC and 1 August 2007 09:00 UTC: (a) orbits of CloudSat and CALIPSO satellites, (b) model cloud cover in grey shading and cloud radar reflectivity profiles in colour.

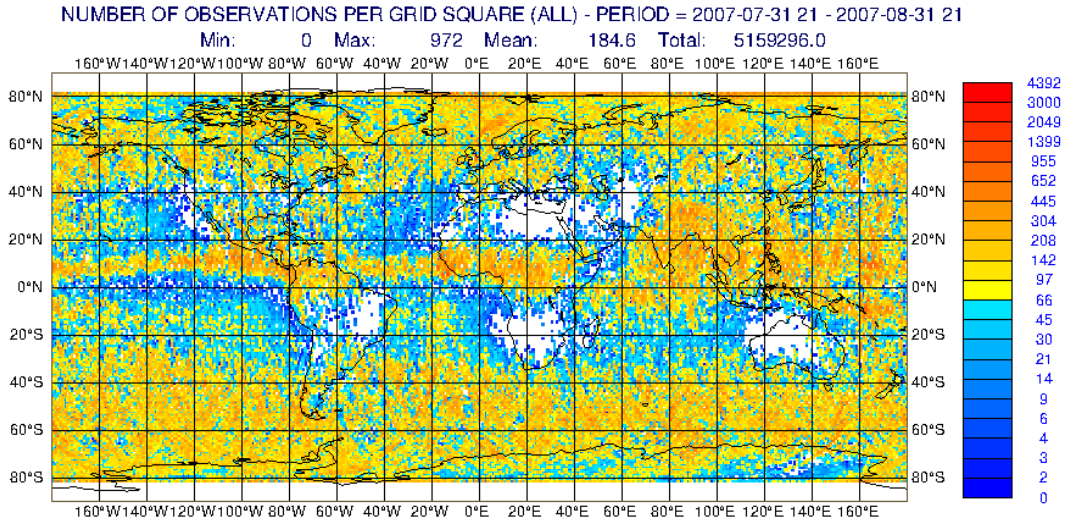

Figure 2.2: Same as Fig. 2.1, but over one-month period of August 2007.

3 Observation impact in CY43R1

3.1 Experimental setup

Several assimilation experiments have been run using the full system of regularly assimilated observations at ECMWF. Experiments have been performed over the period of one to three months between 1 August 2007 and 31 October 2007. For that period, to investigate impact of cloud radar and lidar observations on analysis, many different 4D-Var experiments have been done using different:

i observations (radar, lidar only or both observations without or in combination with all other assimilated observations),

ii observation errors (different degree of the error inflation),

iii observation reduction (increased horizontal averaging or vertical thinning).

The results from only the following experiments, which could provide more inside how the new cloud observations could impact 4D-Var analysis in the future, will be presented:

- **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;
- **RADLID**: experiment assimilating cloud radar and lidar observations on top of all other normally assimilated observations.

Cloud radar and lidar observations averaged over model grid-box at approximately 36 km have been used with:

- **1err** - single observation errors as defined in Fielding and Janisková (2017a, 2020);
- **1.5err** - observation errors one and half of the defined value;
- **2err** - observation errors twice as large as defined.

Horizontal averaging of cloud radar and lidar observations to the coarser resolution of approximately 72 km (**Ha**) have been also tested.

3.2 Comparison of experiments against new assimilated observations

In order to check whether the model is able to represent cloud radar reflectivity and lidar backscatter observations well and whether the analysis is getting closer to these observations, the performed experiments are evaluated against the newly assimilated observations. For such first step evaluation, experimental results are compared against these observations themselves.

Firstly, the results from the 4D-Var experiment assimilating cloud radar and lidar observations, with double observation errors, in combination with all other routinely assimilated observations (**RADLID_2err**) are displayed for CloudSat cloud radar reflectivity in Fig. 3.1 and CALIPSO cloud lidar backscatter in Fig. 3.2. An evaluation is presented in term of time series for the first guess and analysis (AN) departures, i.e. differences between observations and first guess or analysis, respectively, together with their standard deviations.

For cloud radar reflectivity (Fig. 3.1), there is a positive bias before bias correction. After applying the bias correction described in Fielding and Janisková (2017a, 2020), the magnitude of the bias is reduced

generally and the departures between observations and model are distributed around the zero line for the whole 3-month period. This indicates that the bias correction scheme is efficient in removing the bias and even more that there is no significant drift in the bias corrected departures.

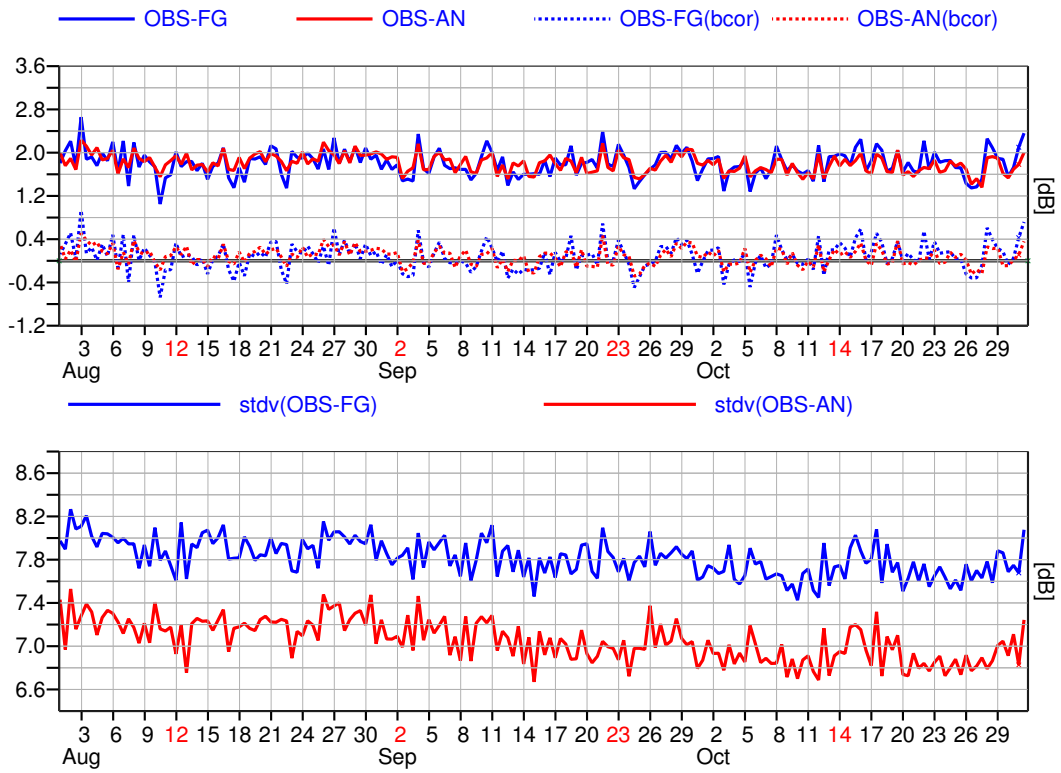

*Figure 3.1: Globally averaged time series of: (top) bias before (solid line) and after applying bias correction (dotted line) and (bottom) standard deviation for the first guess (FG, blue) and analysis (AN, red) departures with respect to CloudSat cloud radar reflectivity observations (OBS). The results shown are from the 4D-Var experiment assimilating cloud radar and lidar observations, with double observation errors, in combination with all other routinely assimilated observations (***RADLID_2err***), for three months in 2007.*

For cloud lidar backscatter (Fig. 3.2), although the global mean bias is reduced by applying the bias correction, some of it is still retained. This might be due to the fact that correcting the biases in a strongly attenuated signal is not straightforward. These remaining biases could be due to cloud microphysics assumptions or perhaps our parametrization of multiple scattering (Fielding and Janisková, 2017a). However, further research is required to investigate the possible source of those biases.

Overall, as illustrated on the time series of standard deviation for the FG and AN departures with respect to CloudSat cloud radar reflectivity (Fig. 3.1) and CALIPSO cloud lidar backscatter (Fig. 3.2), the performed experiments have shown that the bias correction is efficient to bring the model closer to the observations as intended. The impact of the new observations on the 4D-Var analysis when compared against own observations based on statistics for three months of assimilation cycling confirmed that the analysis is getting closer to these observations. However, the experiments revealed that the model is not drawn as strongly to cloud lidar observations, perhaps due to ambiguities in matching the observations by increasing cloud amount at the top of cloud. This results in corrections of the departures at the base of the cloud due to the increased attenuation of the modelled signal and that might lead to an excess of cloud in the model. Whether assimilating the whole profile rather than just when there is cloud in both model and observations might help to solve this problem could be investigated in the future.

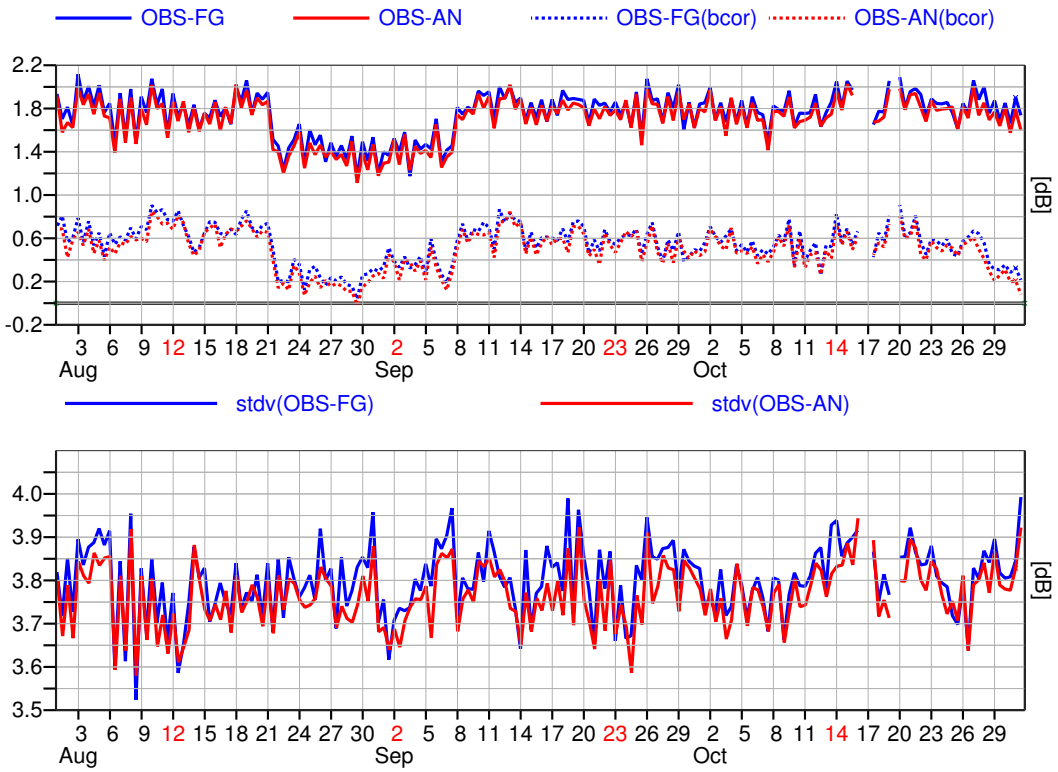

Figure 3.2: Same as Fig. 3.1, but for departures with respect to CALIPSO cloud lidar backscatter.

Secondly, experiments have been performed where additional horizontal averaging has been applied to both the cloud radar reflectivity and the cloud lidar backscatter (experiment **RADLID_2err_Ha**) as a part of investigation of the impact coming from the different observation processing described in sec. 3.1 on the first-guess and analysis departures. Figures 3.3 and 3.4 show that averaging these observations to the coarser resolution of approximately 72 km leads to additional decreasing of the departures to these observations. This impact is more significant for the radar than the lidar observations.

To illustrate how impact of the new observations is distributed over the globe, horizontal distributions of the model first guess and analysis departures with respect to CloudSat cloud radar reflectivity observations are shown in Fig. 3.5 and 3.6, respectively. The results are shown for the 4D-Var experiment assimilating cloud radar and lidar observations using double observation errors without applying additional horizontal averaging (**RADLID_2err**) run for three months in 2007. Figures clearly show that the analysis gets closer to the observations comparing to the model first guess values. This is also reflected on mean standard deviation of the departures being decreased from 10.6 dB for the FG to 9.8 dB for the analysis.

In order to further investigate the impact of the new observations, the high dependent performance of the system has been assessed. To summarize outcomes of that investigation, the vertical profiles of FG and AN departures for radar and lidar for 3-month assimilation cycling are shown for Tropics, Northern and Southern Hemisphere in Figs. 3.7 and 3.8 respectively. The results from the two experiments assimilating cloud radar and lidar observations on top of all other routinely assimilated observations as used in Figs. 3.3 and 3.4 are presented here. One should note that:

i the number of used observations is obviously decreased in the experiment **RADLID_2err_Ha** (plotted in black) compared to the experiment **RADLID_2err** (plotted in red) because of applying

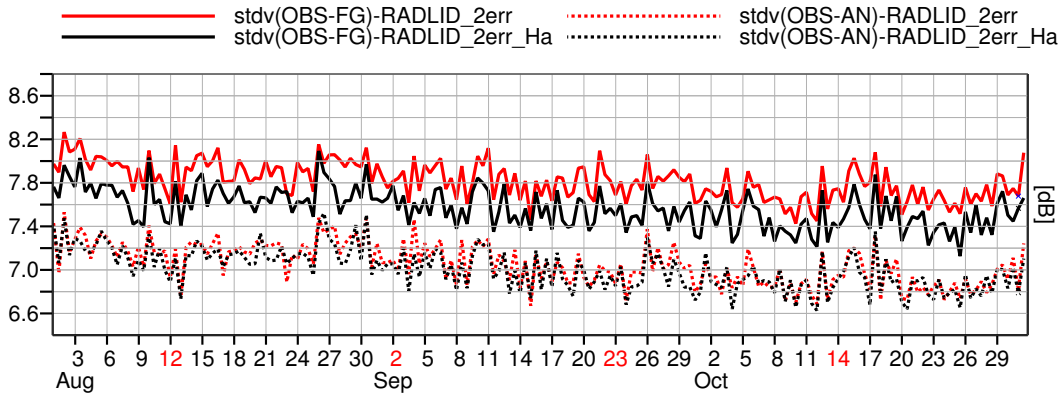

*Figure 3.3: Globally averaged time series of standard deviation for the first guess (FG, solid) and analysis (AN, dotted) departures with respect to CloudSat cloud radar reflectivity observations (OBS). The results shown are from the 4D-Var experiment assimilating, in combination with all other routinely assimilated observations, cloud radar and lidar observations using double observation errors without (***RADLID_2err****, in red) or with (***RADLID_2err_Ha****, in black) applying additional horizontal averaging to these observations for three months in 2007.*

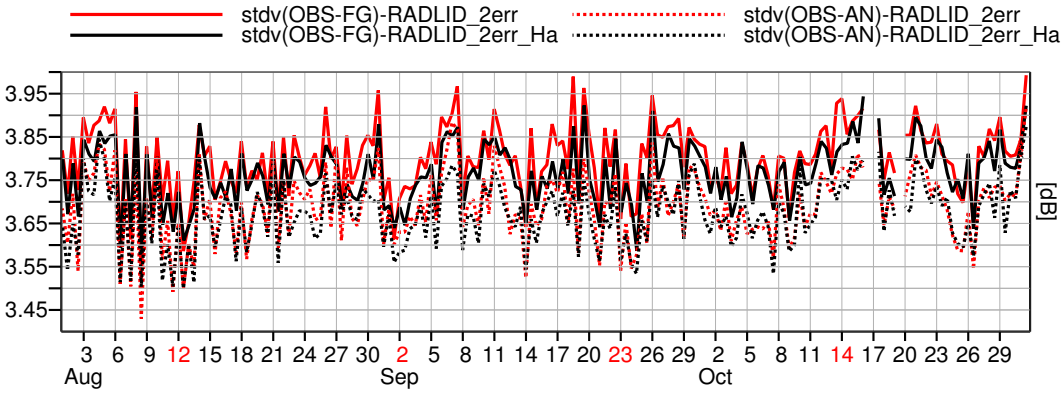

Figure 3.4: Same as Fig. 3.3, but for departures with respect to CALIPSO cloud lidar backscatter.

additional horizontal averaging to cloud radar and lidar observations,

ii an additional screening step is applied to both AN departures and FG departures to reject absolute AN departures greater than 20 dB for radar reflectivity and 10 dB for lidar backscatter in order to ensure meaningful statistics when considering variables in logarithmic space.

Assessment of height dependent values of FG and AN departures revealed that the smallest root mean square (rms) errors in these departures are observed in the Southern Hemisphere, especially for the radar (Figure 3.7). In the Tropics, the rms errors above 800 hPa are the largest because of prevailing convective clouds which are of smaller scale therefore more difficult to represent by the model with the coarser resolution than the convective scale.

For the cloud radar reflectivity (Figure 3.7), root mean square (rms) errors in the AN departures are clearly reduced with respect to the FG departures for all domains. In the levels below 500 hPa, reduction is larger in the extra-tropics than in the Tropics.

For the lidar backscatter (Fig. 3.8), AN rms errors are reduced compared to the FG errors above 400 hPa in all regions. This indicates the cloud lidar backscatter observations have a greater positive impact on the

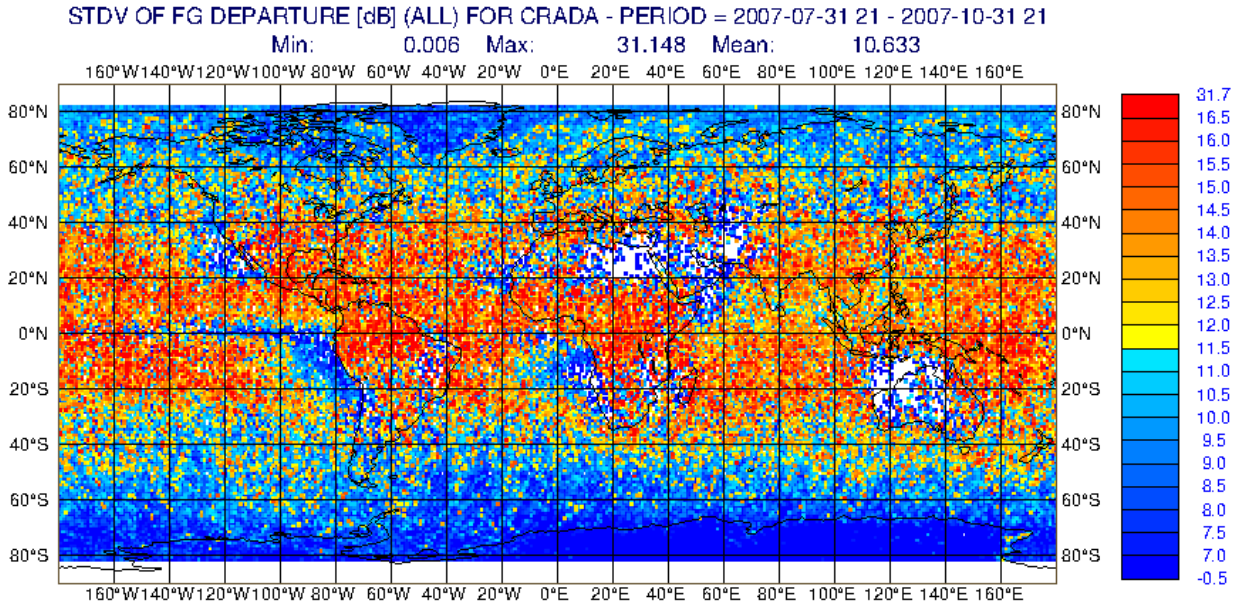

*Figure 3.5: Standard deviation of the first guess (FG) departures with respect to CloudSat cloud radar reflectivity observations. The results shown are from the 4D-Var experiment assimilating, in combination with all other routinely assimilated observations, cloud radar and lidar observations using double observation errors without applying additional horizontal averaging (***RADLID_2err***) for three months in 2007. Mean value of the standard deviation is provided in the legend.*

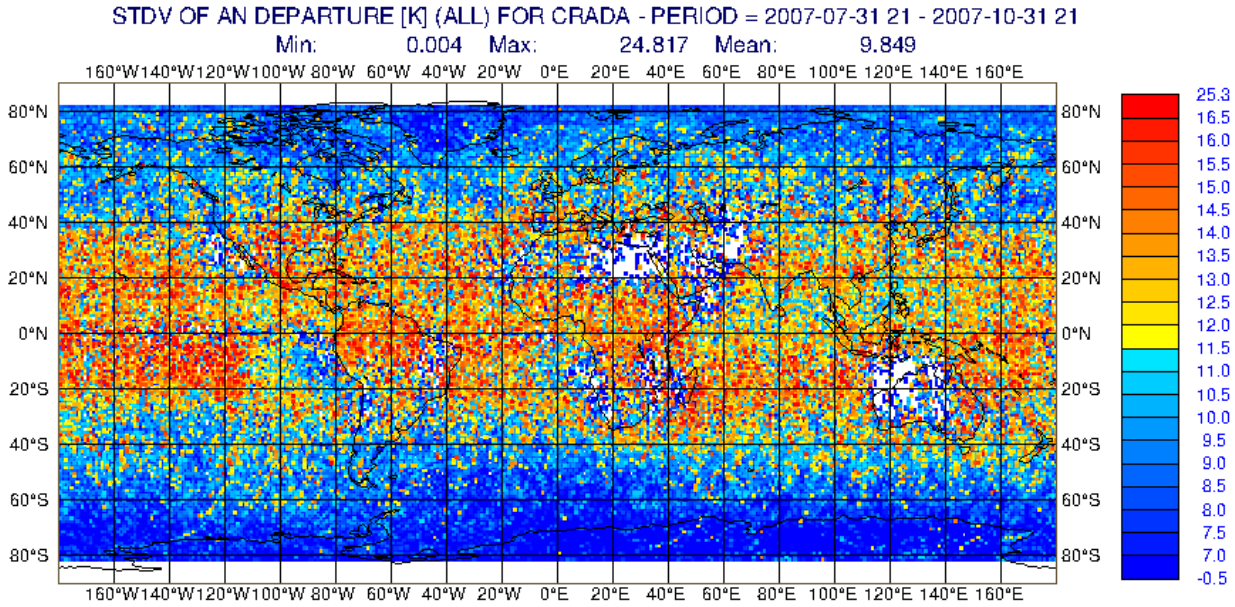

Figure 3.6: AN departures - radar

analysis at these levels. However, the AN rms errors for the cloud lidar backscatter are slightly larger than the FG ones below 400 hPa, especially in the Tropics and in the Northern Hemisphere. This suggests that adjusting AN departures in those levels is driven more by other observations than the lidar one. Larger observation errors in the areas of convective clouds, mainly because of greater representativity issues of

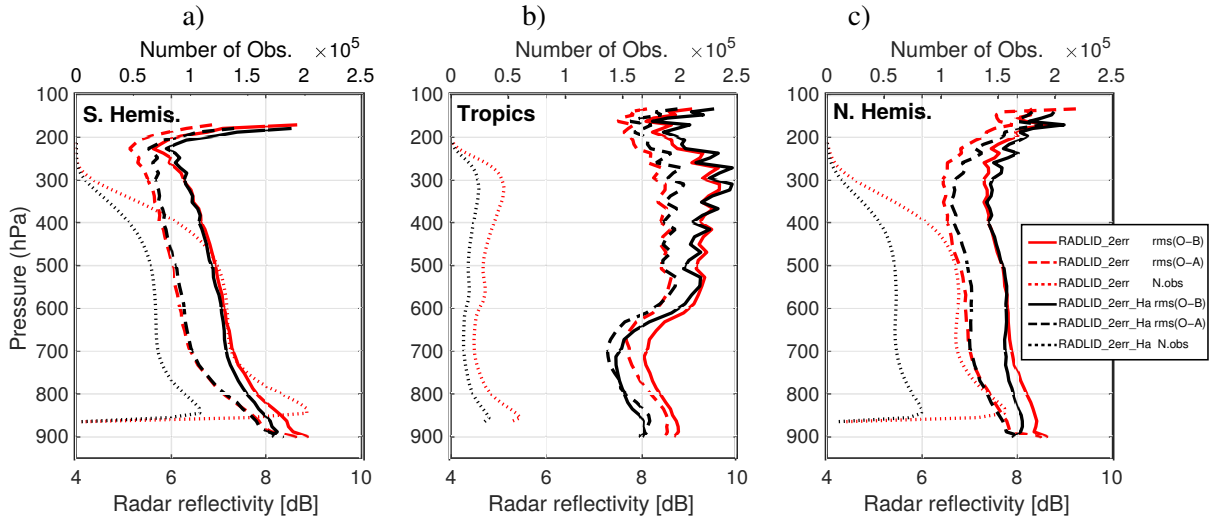

*Figure 3.7: Root mean square error (rms) of background departures (solid line) and analysis departures (dashed lines) with respect to cloud radar reflectivity observations for experiments assimilating, in addition to other assimilated observations, cloud radar reflectivity and lidar backscatter using double observation errors without (***RADLID_2err***, in red) or with (***RADLID_2err_Ha***, in black) applying additional horizontal averaging to these observations. Number of cloud radar and lidar observations used in the experiments is displayed by dotted lines. Results are shown for (a) Southern Hemisphere, (b) Tropics and (c) Northern Hemisphere over the period from 1 August to 31 October 2007.*

such clouds, might be also contributing to the above problem.

Figure 3.7 shows that the smallest root mean square (rms) errors in FG and AN departures are observed in the Southern Hemisphere, especially for the radar. In the Tropics, the rms errors above 800 hPa are the largest because of prevailing convective clouds which are of smaller scale therefore more difficult to represent by the model with the coarser resolution than the convective scale. Generally, the AN departures are significantly more reduced for the radar (Fig. 3.7) than lidar (Fig. 3.8) observations indicating lower potential impact of the cloud lidar backscatter on the analysis.

Generally, the AN departures are significantly more reduced for the radar (Fig. 3.7) than lidar (Fig. 3.8) observations indicating lower potential impact of the cloud lidar backscatter on the analysis.

3.3 Comparison of experiments against other assimilated observations

The next step in evaluating the impact of the assimilated cloud radar reflectivity and lidar backscatter is to compare results against other assimilated observations. This has been done by comparing the first-guess fits to other routinely assimilated observations for the period up to three months. The presented results are based on statistics using averaged values over the whole globe.

Firstly, the impact of scaling the observation error is assessed (Fig. 3.9 - 3.10) over a one-month period. Observations of cloud radar reflectivity and lidar backscatter were used with the different observation errors: **1err** - errors as defined in Fielding and Janisková (2020), **1.5err** - errors 1.5 of the defined values and **2err** - errors twice as large as the defined values. The results indicate that the relatively better FG fits to other assimilated observations is obtained when observation errors for cloud radar and lidar data are inflated to account to account for the known correlations. This is mostly pronounced when increasing the observation errors by a factor of two (**2err** experiment) for AMSU-A tropospheric channels 5-8 and

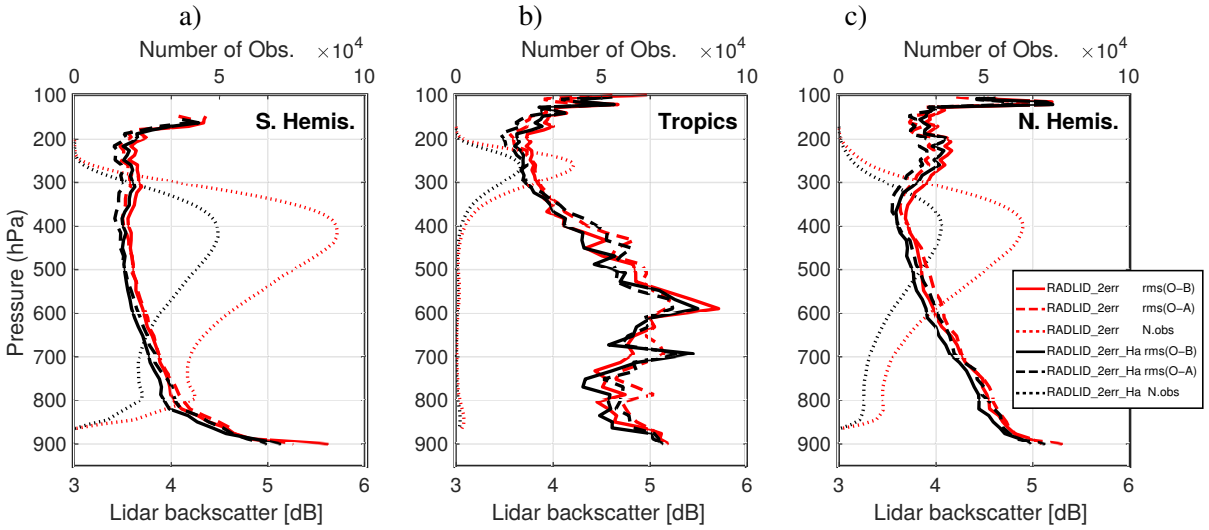

Figure 3.8: Same as Fig. 3.7, but for cloud lidar backscatter.

stratospheric channels 9-14 (Fig. 3.9a). It is also the case for High-resolution Infrared Radiation Sounder (HIRS) temperature channels 5-8 and 13-15 (Fig. 3.9b). Increasing the observation error by a factor of two also removes the degradation seen for the FG departures of HIRS water vapour channels 10-12 when using the default error values. For the SATOB wind observations (Fig. 3.9c), there is no statistically significant differences between the different error scalings for this observation type. Although, the first guess gets closest to the observations when observation errors for cloud radar and lidar are increased by one and half with respect to the estimated errors.

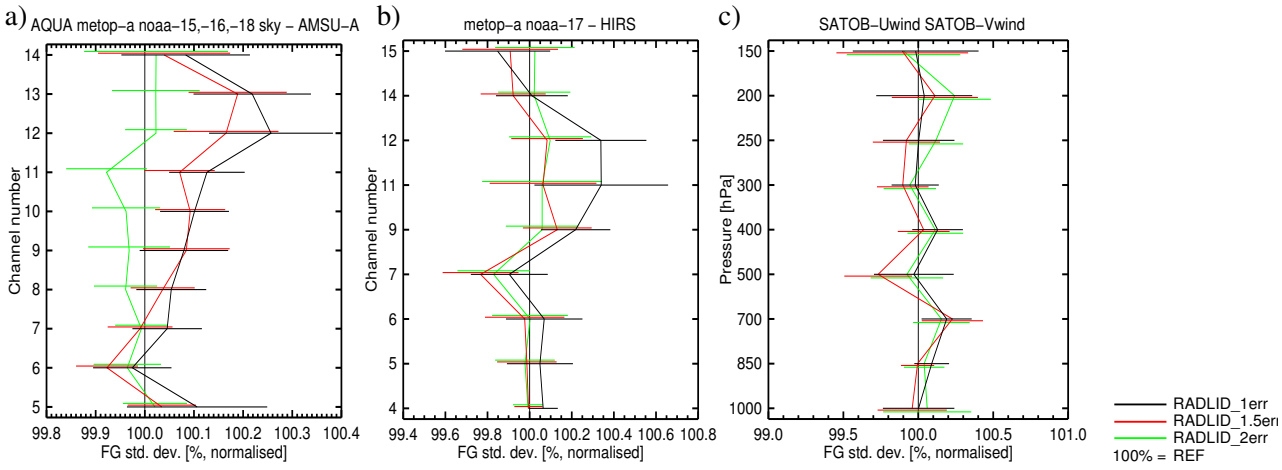

*Figure 3.9: Difference between experimental runs and the reference run (REF) in standard deviation normalised by the standard deviation of REF for the first guess fits to different observations: (a) AMSU-A tropospheric (channels 5-8) and stratospheric (channels 9-14) temperature, (b) HIRS temperature (channels 5-8, 13-15), water vapour (channels 10-12) and (c) SATOB wind. The reference run is represented by the 100% line and negative values correspond to improvements in the experimental runs. Horizontal bars indicate statistical significance at the 95% level. The results for experiments assimilating, in addition to other observations, cloud radar reflectivity and lidar backscatter using different multiple of the computed observation errors, single (**RADLID_1err**, in black), 1.5 (**RADLID_1.5err**, in red) and double (**RADLID_2err**, in green), are shown for the whole globe over the period from 1 August to 31 August 2007.*

The impact of error inflation is not significant for the conventional observations as shown in (Fig. 3.10).

However, as overall the double error version produces the most promising results, this experiment was extended to run over the period of three months.

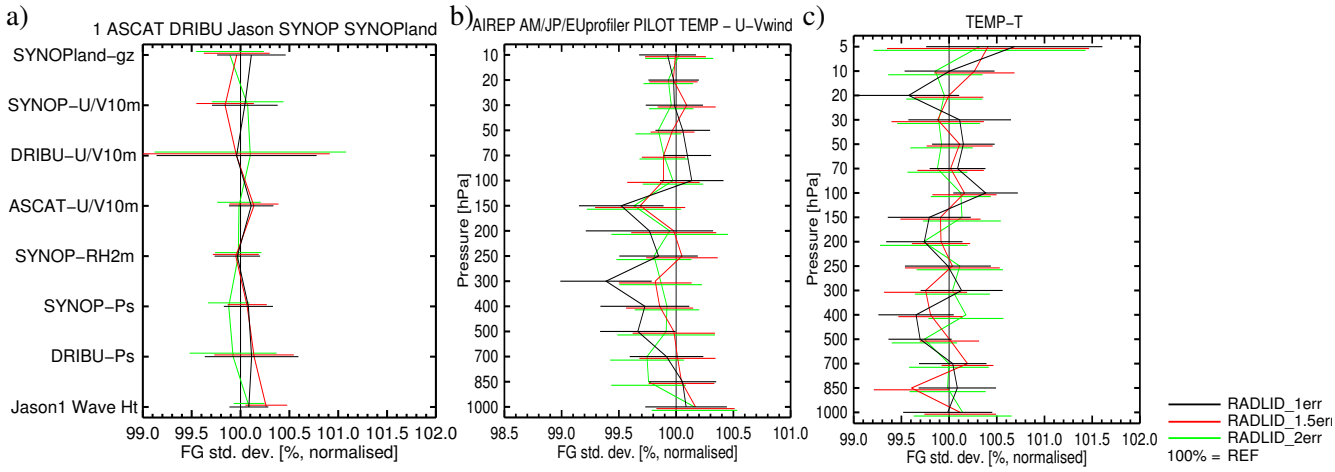

Figure 3.10: Same as Fig. 3.9, but for the first guess fits to different conventional observations: (a) surface pressure, relative humidity and 10-meter wind, (b) wind and (c) temperature profiles.

Secondly, the influence of observation reduction by using increased horizontal averaging has been evaluated over the period of three months (Fig. 3.11 - 3.12). The FG departures from two experiments, one without (**RADLID_2err**) and another with (**RADLID_2err_Ha**) applying increased horizontal averaging (from approximately 36 km to 72 km), are compared. When increased horizontal averaging is applied, there is an improvement with respect to the temperature observations as seen for both tropospheric and stratospheric channels of AMSU-A (Fig. 3.11 a), as well as HIRS channels 5 and 12 which are sensitive to temperature (Fig. 3.11 b). When evaluated with respect to SATOB wind observations (Fig. 3.11 c), the impact of resolution used for the cloud radar and lidar observations is rather neutral.

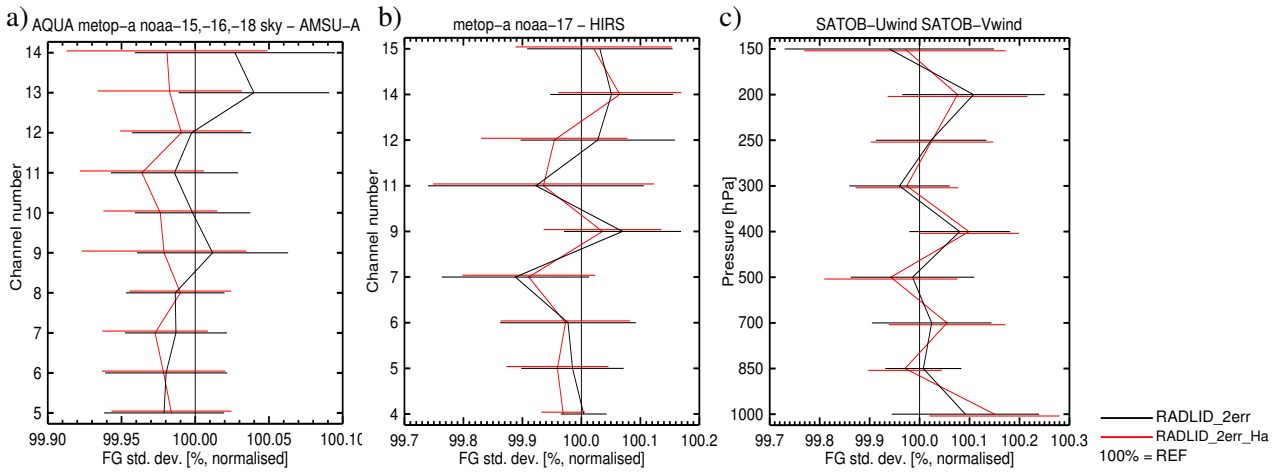

*Figure 3.11: Same as Fig. 3.9, but for experiments assimilating, in addition to other assimilated observations, cloud radar reflectivity and lidar backscatter using double observation errors without (**RADLID_2err**, in black) or with (**RADLID_2err_Ha**, in red) applying additional horizontal averaging to these observations. The results are shown for the whole globe over the period from 1 August to 31 October 2007.*

The impact of using coarser resolution for cloud radar and lidar observations on the conventional surface observations of pressure, humidity and wind (Fig. 3.12a) varies, but overall it is close to neutral. This is

also the case for TEMP temperature observations (Fig. 3.12c). As seen in Fig. 3.12b, the impact of new observations on the analysis is generally positive for conventional wind observations (such as TEMP, PILOT, AIREP and wind profilers).

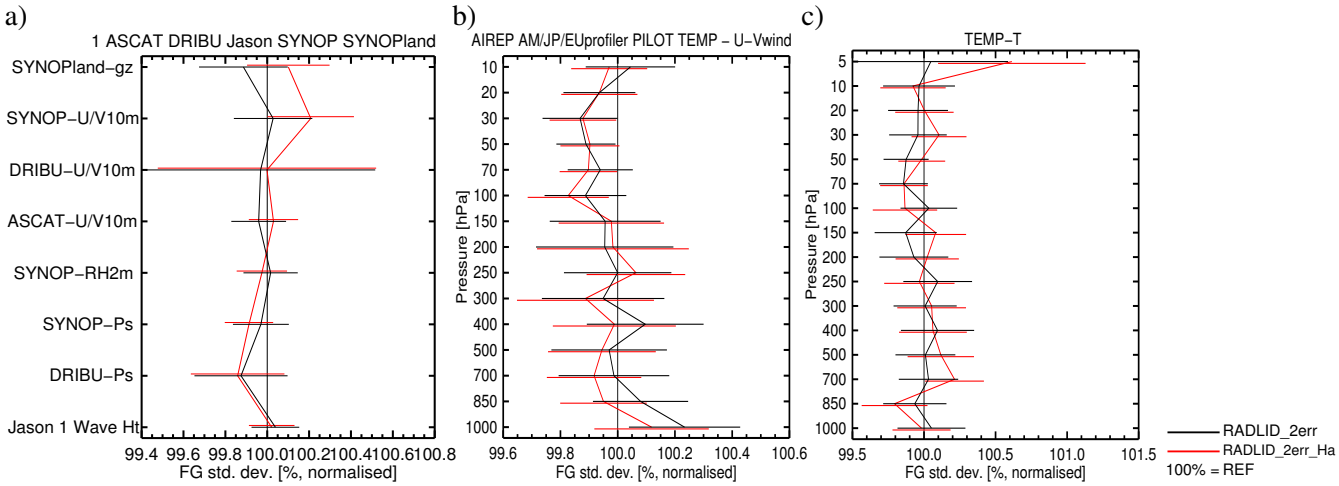

Figure 3.12: Same as Fig. 3.11, but for the first guess fits to different conventional observations: (a) surface pressure, relative humidity and 10-meter wind, (b) wind and (c) temperature profiles.

Altogether, verification against other assimilated observations has shown that their background departures are either neutral or slightly reduced when assimilating the new observations. Generally the best results have been obtained when cloud radar reflectivity and lidar backscatter were used with double observation errors in combination with applying additional horizontal averaging to these observations.

4 Improving the impact of cloud radar and lidar observations

There are many ways to improve the impact of cloud radar and lidar observations within a data assimilation system, several of which have already begun to be explored. The most basic is through quality control and screening. From the model side, observations of clouds that are not well represented by the model must be discarded or have less weight assigned to them. On the other hand observations often contain artefacts or biases that must be corrected if they are to be used without degrading the analysis.

Related to the quality control and screening is the definition of the observation errors that control the weight of the observations in the assimilation system. The definition of flow-dependent observation error variances for radar and lidar observations was extensively explored in Janisková and Fielding (2018b), and the scaling of these was investigated in Sec. 4.2. The correlations of the observation errors have not yet been considered and are explored in Sec. 4.2 of this report.

Finally, the observation operators that provide the model equivalent to the observations are critical components in the assimilation system. Making the operators more realistic should improve the impact of the observations by allowing more observations to be assimilated and potentially allow the observation operator component of the observation error to be reduced thus increasing the weight of the observations. Improving the quality of the observation operators also can reduce noise and uncertainty in model evaluation studies.

The greatest source of uncertainty in the observation operators is probably in the choice of microphysical assumptions, such as the shape and size of the particle size distribution. Given the huge natural variability

in these microphysical assumptions, it is difficult to define a one-size-fits-all approach, particularly if the choices are shared with other instrument types. Nevertheless, the microphysical assumptions for the observation operators were explored in Fielding and Janisková (2017a). Instead, in Sec. 4.1 we improve the macrophysical cloud assumptions, such as cloud overlap and subgrid condensate variability, made in the observation operators.

4.1 A triple-column approach for subgrid condensate variability

In a numerical weather prediction model, parameterizations are required to represent physical processes that are not represented explicitly by the model's large-scale equations. For example the cloud physics scheme uses many different parameterizations for cloud processes, such as auto-conversion of cloud droplets to rain or evaporation of precipitation. These parameterizations use the grid-box mean cloud properties to compute the mean process rates. However, we know that in reality clouds are highly inhomogeneous, particularly at scales 1–10 km. Further, the parameterizations of cloud processes are often very non-linear; any non-linear process acting on a grid-box mean variable will be biased if sub-grid variability exists and is not taken into account. The same is true for any non-linear observation operator using grid-box mean variables as its input.

To account for subgrid condensate variability in our radar and lidar observation operator, the double-column approach (that accounts for cloud overlap between model layers) has been extended to a 'triple-column' approach. The approach is inspired by the Tripleclouds (Shonk and Hogan, 2008) radiative transfer algorithm. Tripleclouds uses three separate uniformly filled columns, one clear and two cloudy. The cloud water content is increased in one-column (the 'thick' column) and decreased in the other column (the 'thin' column) so that the mean radiation across the column is consistent with an assumed probability distribution function (PDF) of subgrid water content variability. While a broadband radiative transfer algorithm has to represent both direct and diffuse radiation, for our radar and lidar observation operator we make a single scattering approximation (radiation does not scatter more than once in the radar or lidar beam) so we only need consider the 'direct beam' radiation.

Numerically, the triple-column approach is very similar to the double-column method and computes the transmission of radiation in the two cloudy columns, T_{w1} and T_{w2} and the clear-sky transmission, T_c, iteratively through the atmosphere from the top model level to the lowest model level. For each layer, the transmittance is equal to the incoming signal multiplied by the transmissivity of the current layer. The proportion of signal that travels in and out of either the cloudy or clear layer is governed by the assumed overlap between adjacent layers, i.e.:

$$\begin{pmatrix} T_c^{l+1} \\ T_{w1}^{l+1} \\ T_{w2}^{l+1} \end{pmatrix} = O \begin{pmatrix} \frac{1}{1-C_{l+1}} e^{-2\alpha_c \Delta h_l} \\ \frac{2}{C_{l+1}} e^{-2\alpha_{w1} \Delta h_l} \\ \frac{2}{C_{l+1}} e^{-2\alpha_{w2} \Delta h_l} \end{pmatrix} \cdot \begin{pmatrix} T_c^{l} \\ T_{w1}^{l} \\ T_{w2}^{l} \end{pmatrix}, \tag{4.1}$$

where α is the particle extinction at a given layer, Δh is the depth of a given layer and the overlap matrix, O, is defined as

$$O = \begin{pmatrix} 1 - C_{ij} & 0.5(C_{ij} - C_{l+1}) & 0.5(C_{ij} - 0.5C_{l+1}) \\ 0.5(C_{ij} - C_l) & 0.25\phi_1(C_l + C_{l+1} - C_{ij}) & 0.25\phi_2(C_l + C_{l+1} - C_{ij}) \\ 0.5(C_{ij} - C_l) & 0.25\phi_2(C_l + C_{l+1} - C_{ij}) & 0.25\phi_1(C_l + C_{l+1} - C_{ij}) \end{pmatrix}, \tag{4.2}$$

where each row in the overlap matrix corresponds to one of the three ways that the fraction of signal can

travel from each column at a given model layer to the columns in the model layer below. By definition, these fractions can be calculated from the two adjacent cloud layers, C_l and C_{l+1}, and the joint areal fraction between two layers, C_{ij}. For maximum overlap (see Hogan and Bozzo, 2018, for other flavours of overlap), $C_{ij} = max(C_l, C_{l+1})$. The vertical decorrelation in subgrid condensate variability, ρ, is treated in the same way as cloud overlap, with $\phi_1 = 1 + \rho^{\frac{1}{2}}$ and $\phi_2 = 1 - \rho^{\frac{1}{2}}$ (following the same approach as Hogan and Illingworth, 2000). In all our experiments we set $\rho = 1$, i.e., the overlap in subgrid condensate variability is assumed to be maximal.

Finally, the grid-box average attenuated signal, Z_{at} at any level is simply:

$$Z^l_{at} = T^l_c Z^l_c + T^l_{w1} Z^l_{w1} + T^l_{w2} Z^l_{w2}. \tag{4.3}$$

where Z is the unattenuated backscatter. For the computation of lidar attenuated backscatter, the so-called Platt-coefficient is included in the extinction term in Eq. 4.1 to represent (albeit crudely) multiple scattering (see Eq. 10 in Fielding and Janisková, 2020).

To compute the unattenuated backscatter for the thin and thick cloud regions, the grid-box mean water content, w needs to be scaled. Following Shonk and Hogan (2008), the fractional standard deviation of subgrid water content variability, f, is used to determine the magnitude of the scaling. If we assume a Gamma distribution for the size distribution of w then to a good approximation:

$$\begin{aligned} w_1 &= e^{-f(1+0.5f(1+0.5f)))} w \\ w_2 &= (2 - \frac{w_1}{w}) w, \end{aligned} \tag{4.4}$$

can be used to find the 16th percentile (and 84th percentile) of the subgrid water content distribution, where w_1 is the water content for a given layer in the thin column and w_2 is the water content for a given layer in the thick column. See Shonk and Hogan (2008) for the justification for using the 16th percentile of the water content distribution.

The scaled water contents are then used for determining the extinction in each column and the same scaling is applied to all six different hydrometeor types. The determination of f is a research topic in itself; many studies have attempted to quantify the parameter either globally or for different regimes (e.g., Shonk et al., 2010; Hill et al., 2015; Ahlgrimm and Forbes, 2016). For now, we use a globally constant value of $f = 0.75$ to be consistent with the cloud and radiation schemes at ECMWF.

To evaluate the new triple-column approach, the multi-column approach (see Di Michele et al., 2012, for full description) has been adapted to include the representation of subgrid condensate variability. The multi-column approach works by creating a set of sub-columns, where each level is assigned to be either totally cloudy or totally clear, such that the sum of sub-columns at each level is equivalent to the grid cloud fraction, and their distribution is consistent with an assumed overlap scheme. To include condensate variability, a Monte Carlo approach is taken such that the assumed size distribution of condensate variability is sampled using a random number generator. Note that the triple-column method has several advantages over the multi-column approach as it is easier much cheaper to run and, importantly for data assimilation, is differentiable.

4.1.1 Impact on radar

To assess the impact of the triple-column method on the model equivalent radar reflectivity, the forward operator has been run on model data matched to a single CloudSat orbit (shown in Fig. 4.1a). The orbit

spans a variety of cloud regimes including a stratiform frontal regime for model profiles 1100-1300, an area of deep convection for model profiles between 1350-1400 and ice cloud above precipitating boundary-layer clouds for model profiles between 1500-1700.

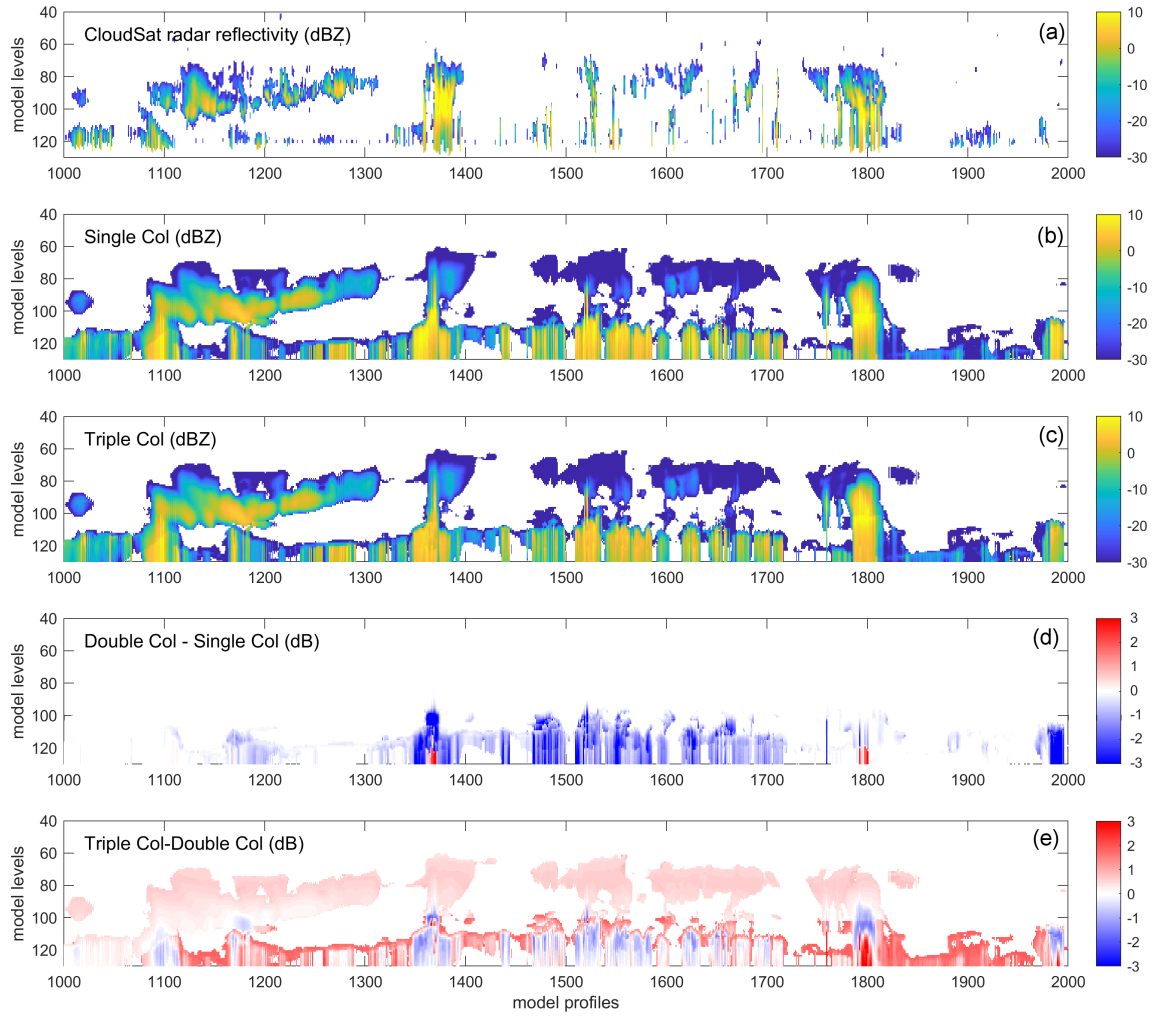

Figure 4.1: Effect of double-column and triple-column methods on forward modelled radar reflectivity. Panels show: (a) superobbed CloudSat radar reflectivity for an orbit during 20070801, (b) corresponding profiles of forward modelled radar reflectivity using the single column method, (c) same as (b), but for using the triple-column method, (d) the difference between forward modelled radar reflectivity (dB) using the double column method and the single column method, (e) the difference between forward modelled radar reflectivity (dB) using the triple column method and the double column method.

Before considering the impact of the triple-column method, it is useful to consider the effect of the double column method over the single column approach (Fig. 4.1d). This informs us of the effect of accounting for cloud overlap. Accounting for cloud overlap modifies the attenuation of the radar signal, so has the greatest effect in heavier precipitation. Generally, Fig. 4.1d shows that the radar reflectivity is decreased by between 1-3 dB where the single column reflectivity is greater than 0 dBZ (indicative of precipitation greater than 1 mm/hr in W-Band radar). In the heaviest precipitation (such as model profiles around 1375 and 1800) the Double column method can increase radar reflectivity relative to the single column method. This occurs when there is another cloud layer which increases the cloud fraction thus exposing the cloud top to previously unattenuated radar signal.

The effect of the triple column method is more complex than the double-column method because it affects both the attenuated signal and the unattenuated signal. To understand the behaviour of the triple-column method versus the double column method (Fig. 4.1e) it is helpful to consider a simplified version of the radar equation that converts model water content, w, to attenuated radar reflectivity, Z_{at}:

$$Z_{at} = Aw^b \exp(-Bw), \tag{4.5}$$

where A and B are coefficients and b is an exponent, which all depend on the underlying microphysical assumptions.

In the presence of attenuation (given by exponential function in 4.5), subgrid variability of w will tend to decrease the grid-box mean attenuation (i.e., let more of the signal pass through the clouds). This is due to the non-linearity in the exponential function. However, if the subgrid condensate is assumed to be maximally overlapped ($\rho = 1$ in eq. 4.2) then the thick cloudy region (corresponding to larger values of reflectivity) will be attenuated more strongly than the thin region and, in some cases, completely attenuated. The weighting of more attenuation in the thick region where the unattenuated reflectivity is greater leads to an overall reduction of radar reflectivity.

Conversely, for areas with little or no attenuation, subgrid variability of w will tend to increase radar reflectivity if b in eq. 4.5 is greater than 1. In areas of liquid cloud, the small droplet size relative to the radar wavelength means that we are in the Rayleigh regime, such that $b = 2$. As the size of the drops gets larger for precipitation b will tend to 1, which implies subgrid variability of w will not affect radar reflectivity. In Fig. 4.1d, we can see that the greatest increases in radar reflectivity occur at the tops of low level clouds, where we would expect b to be closer to 2. For ice and snow, $b = 1.7$, but is variable depending on the mass-size relationship. This leads to increases in radar reflectivity, but not as strongly as for liquid cloud.

The effect of the triple-column approach has also been compared statistically, and against CloudSat observations. Figure 4.2 shows the bias and standard deviation in forward modelled radar reflectivity compared with CloudSat radar reflectivity (O-B) for one data assimilation window (12 hours). There are several features to note. The first is that in general, the double- and triple-column methods perform very similarly to their multi-column equivalents. The second is that, due to the competing effects, the bias in radar reflectivity is generally quite small, with the greatest difference for temperatures greater than freezing. Similarly the standard deviation in radar reflectivity is very similar for cold-phase clouds, but is reduced by around 5% for temperatures greater than freezing.

Histograms of the radar reflectivity departures (Fig. 4.3a) show that the overall number of observations that would pass the FG departure screening check of ± 20 dB are broadly similar using either the single- double- or triple-column approaches. The overall bias in the FG observations is also slightly increased, but could easily be accounted for by compensating errors in the microphysical assumptions.

4.1.2 Impact on lidar

The impact of the new triple-column method on attenuated lidar backscatter has also been investigated. As was done for studying the impact on radar, it is useful to explain the impact of the double-column approach first. Figure 4.4 shows observed CALIPSO lidar backscatter and corresponding forward modelled lidar backscatter for the same transect as Fig. 4.1. Whereas in the profiles of radar reflectivity, attenuation was generally weak, attenuation is strong everywhere for the lidar except in thin ice clouds. Because contiguous layers of cloud are assumed to be maximally overlapped, accounting for cloud frac-

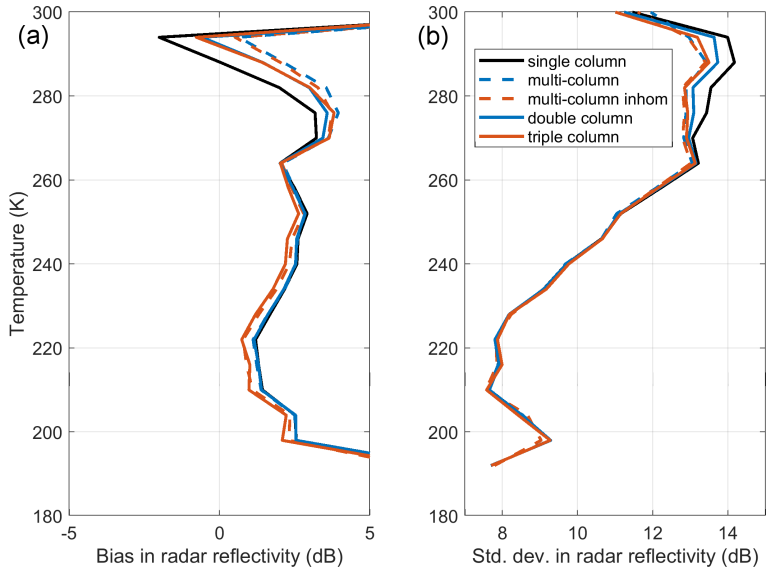

Figure 4.2: Comparison of radar reflectivity forward model solvers with CloudSat observations (O-B departures) for the 12 hour data assimilation windows of the 20070801 00Z analysis as a function of temperature (K). Panels s how (a) bias between observations and model using the single column (black line), multi-column with and without representing subgrid condensate variability (red and blue dashed lines respectively) using 50 subcolumns, double column (blue line) and triple column (red line), (b) same as (a) but for standard deviation of the O-B departures.

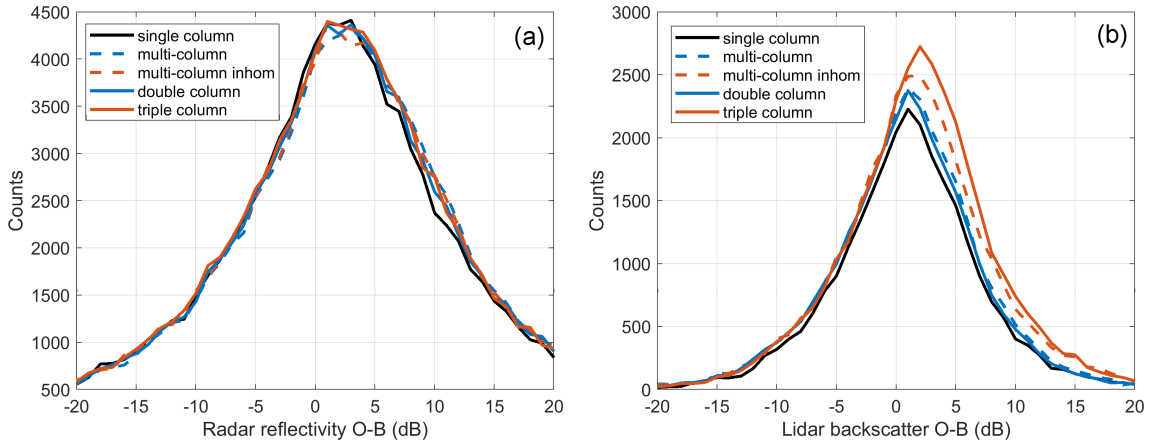

Figure 4.3: Histograms of observation minus background (O-B) departures for the 12 hour data assimilation window of the 20070801 00Z analysis for (a) CloudSat radar reflectivity and (b) CALIPSO lidar backscatter using the single column (black line), multi-column with and without representing subgrid condensate variability (red and blue dashed lines respectively), double column (blue line) and triple column (red line) methods.

tion increases attenuation. This can be seen in the ice cloud and cloud tops in Fig. 4.4d. However, even in contiguous cloudy layers, if the cloud fraction is increasing then accounting for cloud overlap can increase the backscatter because a portion of the cloud layer is exposed to previously unattenuated lidar signal. Elsewhere, multiple layer clouds are assumed to be randomly overlapped so accounting for cloud overlap should decrease the grid-box mean attenuation. An example of this is that cumulus clouds underneath stratocumulus should be visible when using the double-column approach.

In general, the effect of the triple-column approach (Fig. 4.4e) is similar to the double-column approach.

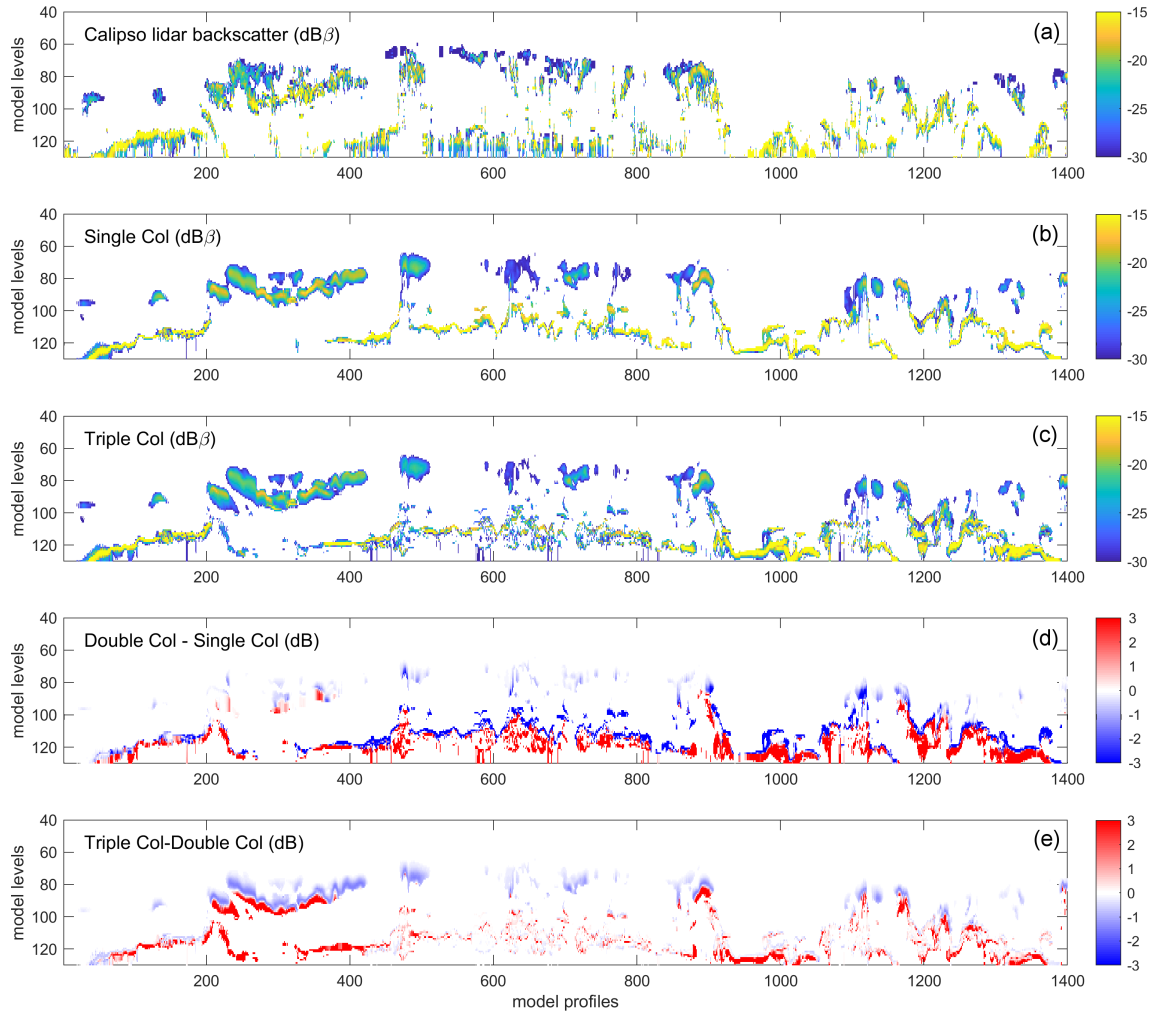

Figure 4.4: Same as 4.1, but for forward modelled attenuated lidar backscatter and observed CALIPSO attenuated lidar backscatter.

At the top of ice clouds, lidar backscatter tends to be reduced as more of the lidar signal passes through the cloud in the thin column whilst getting strongly attenuated in the thick column. Once the lidar signal is fully attenuated in the thick column, the triple-column approach tends to increase the observed backscatter as the lidar signal penetrates deeper into the cloud than it would for either the single or double column approaches. The effect of having less attenuation in the thin column leads to greater backscatter in non-contiguous cloud layers underneath an upper cloud in multi-layer situations.

Quantitatively, the overall bias in forward-modelled lidar backscatter when compared with CALIPSO is increased for the double-column method, and decreased with the triple-column method (Fig. 4.5). However, the greatest effect is on the standard deviation of lidar backscatter; the triple column method has the smallest standard deviation of all the methods, including the multi-column approach. The triple-column method should give very similar results to the multi-column method with inhomogeneity, but the finite number of sub-columns in the multi-column approach introduces additional noise, which could increase the standard deviation of observation minus model statistics.

Finally, the number of additional lidar observations assimilated can be deduced from Fig. 4.3b. The triple-column method allows 20% more observations than the single-column method to pass the FG

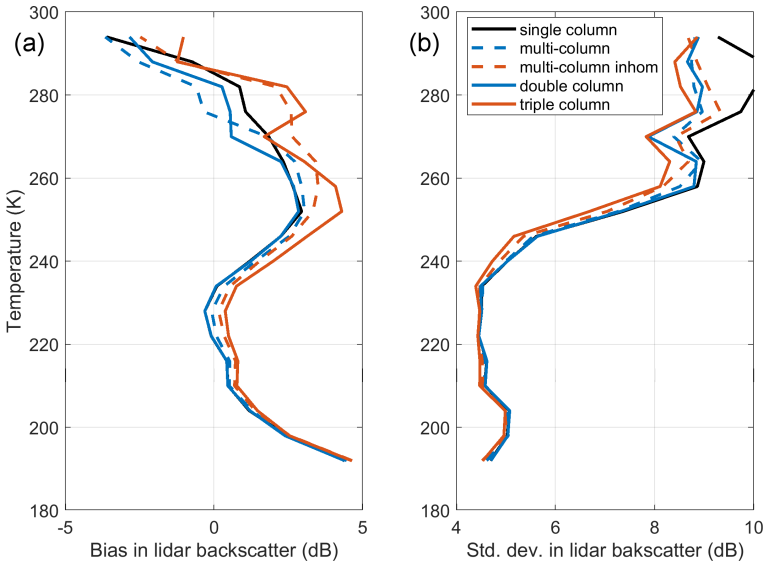

Figure 4.5: Same as 4.2, but for forward modelled attenuated lidar backscatter and observed CALIPSO attenuated lidar backscatter.

departure check of ± 10. Overall, the bias is slightly increased (the overall model backscatter matched to observations is reduced), but, as with the radar, this bias could be mitigated by a revision of the microphysical assumptions.

4.2 Accounting for observation error correlations

4.2.1 Theory and justification

Observation errors are a crucial part of data assimilation as they (in conjunction with background errors) control the weight that is placed on the observations. In variational assimilation, the observation errors are specified in the observation error covariance matrix, $\mathbf{R}$, which forms a cost function, such as the incremental 4D-Var cost function given in Eq. 2.1. For data assimilation of radar and lidar observations we use the flow-dependent observation error model developed in Fielding and Janisková (2017a). Previously, we have assumed observation errors are uncorrelated in the horizontal and vertical. In this case $\mathbf{R}$ reduces to a diagonal matrix, with each element in the diagonal corresponding to the variance of the observation error for each corresponding observation.

To prevent over-fitting to observations that are correlated, observation errors must be inflated. This was demonstrated in Sec. 3, where doubling the observation errors gave the best results. Most observations assimilated operationally at ECMWF have their errors inflated to account for correlations, with the main exception being infrared sounding measurements whose error correlations are specified (Bormann et al., 2016).

Specifying the observation error correlations has two main potential benefits. The first is that it give a physically justified way to reduce the weight placed on the observations. The greater the correlation in the observation errors, the less weight is given to the observations. Thus, in theory, if the observation error correlation was known perfectly, no arbitrary error inflation would be required. Secondly, specifying error correlations should improve the model fit to smaller scales, particularly if the error correlations

are different to the background error correlations (Fowler et al., 2018). Ultimately, it is hoped that by including observation error correlations more information from the radar and lidar observations can be extracted by the assimilation system, therefore improving their impact.

4.2.2 *Characterising the observation error correlation*

With analogy to characterising the observation errors, there are two main ways to characterise the observation error correlations for a particular observation type. The first is to take a statistical approach using first-guess departures; in a linear system the covariance of the background departures $E(\mathbf{dd}^{-T})$ should be equal to $\mathbf{H}BH^T + \mathbf{R}$, where E is the expectation operator and it has been assumed there are no correlations between the background and the observation errors. For cloudy observations it is often assumed (e.g., Geer and Bauer, 2011) that the observation errors dominate the model errors in observation space ($\mathbf{H}BH^T$). In this case the covariance of the background departures can be used. While using this approach is relatively simple to implement, it produces a static error covariance that does not depend on the regime or situation. For radar and lidar observations, it was shown in Fielding and Stiller (2019) that the observation error were strongly situation dependent, implying that the error correlations will be too.

The second approach is to take an 'error inventory' approach by estimating individual error sources. According to Fielding and Janisková (2017a) there are three main sources of error for radar and lidar observations. The first is the measurement or instrument error, which can be thought of as the instrument noise. It is also the smallest of the errors, apart from when the signal is close to the instrument sensitivity. The measurement error is also mitigated by the superobbing of observations. Given its generally small contribution, we neglect the correlation in measurement error. The second is the representativity error. In the literature, representativity error has many different definitions, but here we define it as the sampling error of the observations that arises from the typically narrow footprint of radar and lidar measurements relative to the model scale.

Fielding and Stiller (2019) showed how the vertical correlation in representativity error turns out to be directly proportional to the correlation in the measurements themselves, i.e.:

$$Cov(RE_i, RE_j) \approx \sigma_{RE,i}\sigma_{RE,j}\rho_v(z_i, z_j)\,. \tag{4.6}$$

where $\sigma_{RE,j}$ is the standard deviation of representativity error at level j and ρ_v is the correlation of the measurements in the vertical. To implement this approach, all that is required is an estimate of $\rho_v(z_i, z_j)$. While σ_{RE} is made flow-dependent by using the local standard deviation of measurements along the track, we choose to use a global estimate of $\rho_v(z_i, z_j)$ from a climatology of observation measurements. For radar reflectivity, we use one month of CloudSat observations to produce the correlation matrix shown in Fig. 4.6a. Its features agree with previous investigations into the correlation of cloud measurements. For example the e-folding length scale is approximately 2 km at all heights; Hogan and Illingworth (2000) found similar results in ground based radar measurements to estimate the decorrelation in cloud overlap. Intuitively it makes sense that the decorrelation length in cloud overlap would be similar to that of the representativity error as you would expect overlapped clouds to have the same representativity error (i.e. the mismatch between model and observation footprint in one layer is very similar to the layer directly underneath).

The third source of error is the observation operator error. The observation operator error itself encompasses several distinct sources including the error in radiative transfer (e.g., inadequate representation of multiple scattering), macrophysical assumptions (e.g., cloud overlap or condensate variability) and

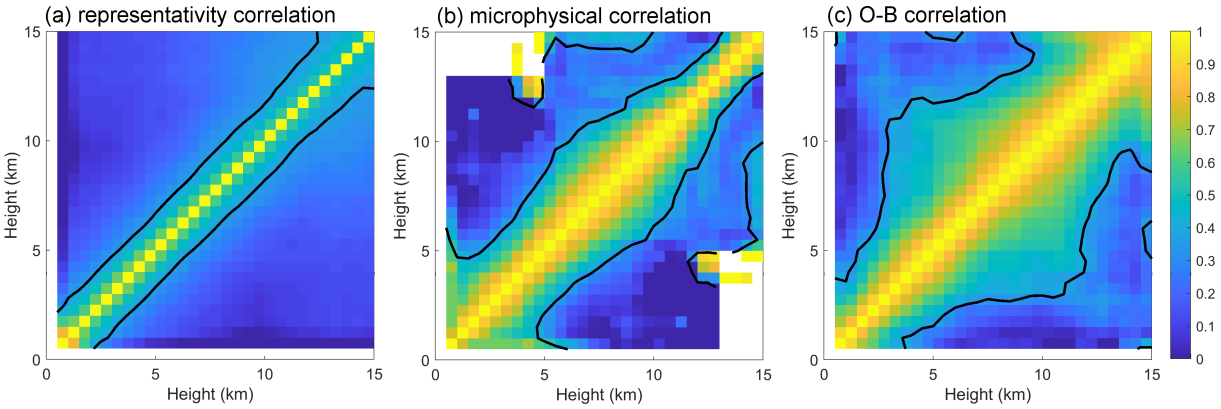

Figure 4.6: Vertical error correlation matrices for cloud radar reflectivity. Panels show (a) correlation in representativity error derived from CloudSat observations, (b) correlation in microphysical errors derived from synergistic retrievals and (c) correlation in observation minus background (O-B) departures. The black lines are isolines of the e-folding distance (e^{-1}).

microphysical assumptions (e.g., particle habit, particle density or particle size). Following the conclusion of Fielding and Janisková (2020) that errors in the microphysical assumptions dominate, we seek a method to estimate their correlations.

To estimate the errors in microphysical assumptions we exploit synergistic retrievals from CAPTIVATE (Mason et al., 2018), which has has been developed to retrieve cloud and precipitation properties from the combination of radar, lidar and radiometer measurements. The combination of instruments in the retrieval allows multiple moments of the size distribution to be retrieved, such as both ice water content and ice particle number concentration. First we define the retrievals of ice water content, liquid water content and rain water content as a quasi-truth, $\mathbf{x}_t$ and, assuming no errors in the CAPTIVATE forward model, define the CAPTIVATE forward modelled radar reflectivity as the truth, $\mathbf{Z}_t = h_{cap}(\mathbf{x}_t)$, where h_{cap} is the CAPTIVATE forward model for radar reflectivity. If we then use our 4D-Var observation operator to forward-model radar reflectivity (h_{EC}), so that $\mathbf{Z}_{EC} = h_{EC}(\mathbf{x}_t)$, we can approximate the correlation in observation operator error using:

$$\mathbf{R} = E(\mathbf{Z}_{EC} - \mathbf{Z}_t)(\mathbf{Z}_{EC} - \mathbf{Z}_t)^T. \tag{4.7}$$

Because the CAPTIVATE algorithm has a more sophisticated forward model for multiple scattering, the correlations also partly include correlations due to errors in the radiative transfer.

To compute a covariance matrix from eq. 4.7, values for $\mathbf{x}_t$ were obtained from CAPTIVATE retrievals of A-train measurements for a 12 hour period between 00 - 12 UTC 01 February 2008. Using a 12-hour period ensures a variety of cloud regimes and situations are captured in the statistics. The retrieved $\mathbf{x}_t$ were then forward modelled using the 4D-Var observation operator. The correlation matrix in Fig. 4.6b was then obtained by converting the covariance matrix using standard formula.

Comparing the first two panels in Fig. 4.6, the microphysical error correlation turns out to be much broader than the representativity error, with an average e-folding distance of around 5 km to 2 km respectively. This suggests that the microphysical error is dominated by the regime or synoptic situation rather than finer-scale turbulent motions, especially at the horizontal scales considered here (observations and model are averaged to around 72 km, see 3.1). Interestingly, there are two peaks in the microphysical correlation, one for radar reflectivity at heights of 2 km, and another for radar reflectivity at 8 km. These

broadly correspond to the average heights for rain and stratiform ice cloud respectively, suggesting that microphysical phase could be an important indicator of microphysical error correlation.

Both the representativity error correlation and the microphysical error correlation are narrower than the error correlation computed from first guess departures (Fig. 4.6c). This is a useful sanity check on our characterisation of the error correlation as, providing the observation errors are a sufficiently large compared to the background errors, the observation error correlations should show some similarity; here the microphysical error correlations (derived independently) are strikingly similar to the O-B error correlations.

4.2.3 Implementation, validation and further work

Ideally for use in assimilation experiments, the observation error correlation matrices should be multiplied by their respective observation error components to form the final observation error covariance matrix used within 4D-Var. However, due to the complexity of the IFS and the 'plumbing' required to move different information to different parts of the code, a pragmatic approach was taken. Figure 4.7 shows the mean correlations for both the representativity error (red line) and microphysical error (blue line) correlations as a function of vertical displacement.

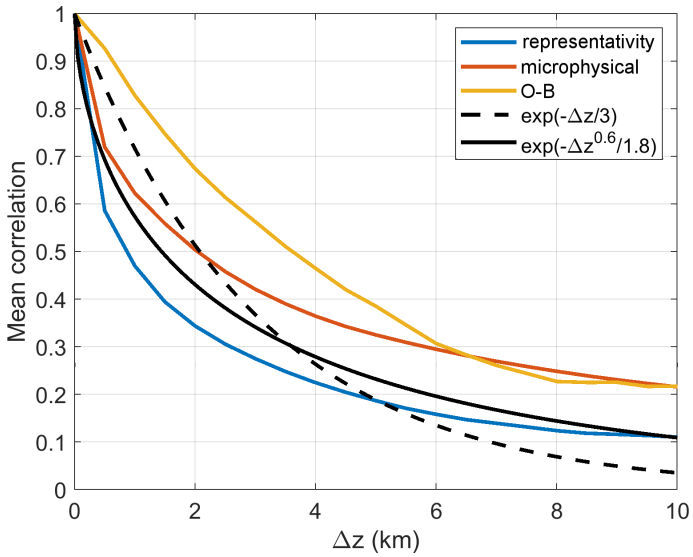

Figure 4.7: Mean correlation as a function of vertical displacement (Δz) derived from the three correlation matrices in Fig. 4.6. Also shown (black lines) are two different fits to the data (see legend).

Using the assumption that both error sources have roughly equal weight we seek to find a function that is a suitable fit to both error sources. Given that correlation functions often exhibit exponential decay we choose to fit the correlation using:

$$\rho(\Delta z) = \exp(-\frac{\Delta z^{\alpha}}{\beta}), \tag{4.8}$$

where Δz is the vertical displacement and α and β are fitting parameters.

Two fits to the data have been chosen. The first uses $\alpha = 1$ and $\beta = 3$, chosen to match the mean e-folding distance of the two error correlation functions. The second uses $\alpha = 0.6$ and $\beta = 1.8$, chosen

to give the best overall to both error correlation functions. The first fit gives greater correlation for the first few kms of displacement before decaying more rapidly than the second fit. Both fitted correlation functions are always less than the O-B correlation.

To validate the observation error correlation implementation, a set of 4D-Var experiments assimilating with and without observation error correlations for radar and lidar have been performed on a single 12-hour data assimilation window with the analysis at 00 UTC 20070801. Both radar and lidar share the same error correlation model.

Fig. 4.8a shows that the greatest reduction in analysis departures of radar reflectivity is seen when the radar and lidar are assimilated, but no observation error correlations are assumed. The experiments using correlations reduce the standard deviation of the departures by about half that of the no correlation experiment; we expect the correlation to reduce the impact of the observations by overall reducing their weight in the cost function. The impact on other metrics such as the first guess departures of other observations is explored in Sec. 5.

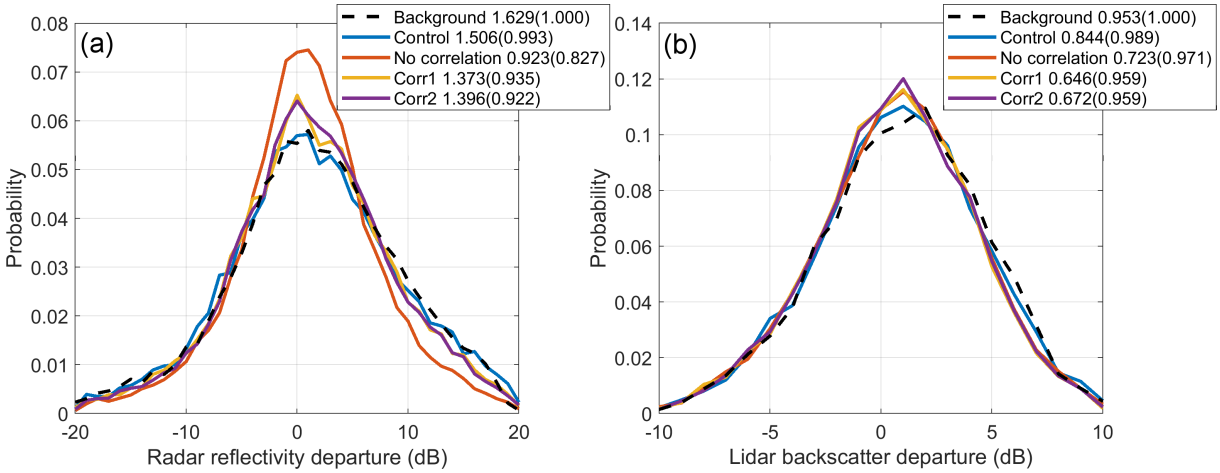

Figure 4.8: Probability functions of (a) radar reflectivity and (b) lidar backscatter showing first guess (dashed) and analysis (solid) departures (dB), for 4D-Var assimilation experiments for one 12 hour assimilation window where: only routine observations are assimilated (control; blue line), radar and lidar assimilated on top of routine observations (no correlation; red line), same as previous but radar and lidar observation errors are assumed correlated using eq. 4.8 with $\alpha = 1$ and $\beta = 3$ (yellow line) and same as previous, but with $\alpha = 0.6$ and $\beta = 1.8$ (purple line). The bias (and standard deviation normalised by the background) of the departures is shown in the legend.

The impact of the correlated observation errors on lidar backscatter is less, due to the reduced impact of the observations relative to the radar. However, including the observation error correlation actually increases the analysis fit to the observations. This could be because the analysis is over-fitted to the radar observations in the 'no correlation' case, so reducing the weight of the radar observations gives a better overall analysis and thus a better fit to the lidar observations.

Finally, although not implemented here, it is worth noting that the hydrometeor phase could have a dramatic impact on the microphysical error correlation. Figure 4.9 shows that when considering the correlation between two observations at different levels with the same hydrometeor phase (neglecting super-cooled liquid) the correlations broadly agree with Fig. 4.6b. However, when considering two layers with differing phase, there is almost no correlation. In these situations, only the representativity error would provide any correlation, thus reducing the overall correlation significantly. Accounting for the particle phase in the observation error correlation model could therefore improve its performance and

is suggested as future work.

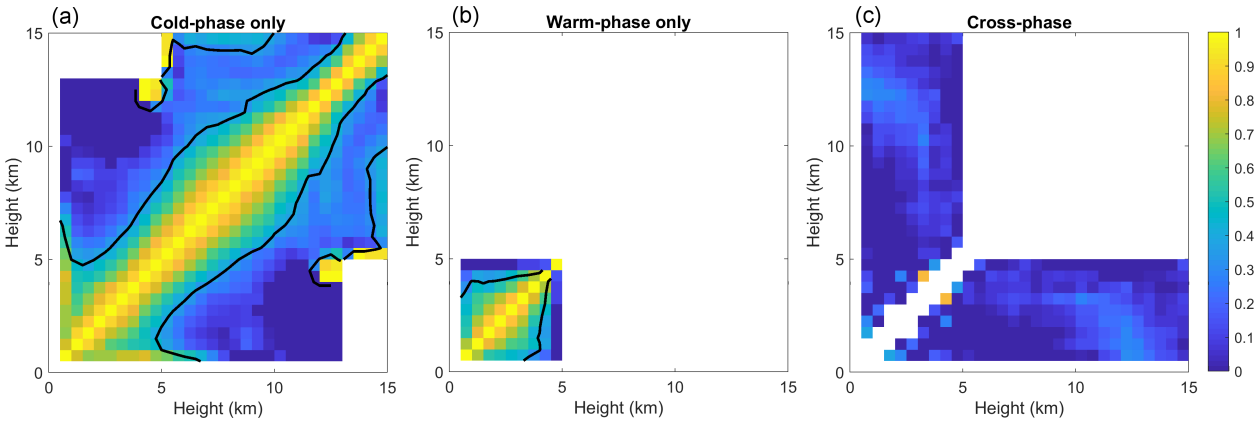

Figure 4.9: Same as Fig. 4.6b, but considering pairs of data points on levels that have temperature (a) both less than 271 K (b) both more than 275 K and (c) one level less than 271 K and one level more than 275 K.

4.3 Bias correction scheme

Since 4D-Var assimilation relies on the assumption that both observations and model background are unbiased quantities, it is essential that any systematic error in the first guess departures are removed. This is achieved by applying a bias correction scheme. Though the 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 the result of forecast model inaccuracies as well as deficiencies in the observation operators which are used to provide the model equivalent to the observations. Now that both the observation operator and the model have been updated, it is necessary to update the bias correction scheme as the biases have changed and the bias correction scheme does not give unbiased results (e.g. Fig. 4.2a shows the triple-column method introduces different biases).

Operationally at ECMWF, most observations use variational bias correction (VarBC) to automatically correct for biases. VarBC works by including an additional term in the 4D-Var cost function that penalises biases according to a set of indicators. As is common practice for assimilation experiments of new observations types, we employ a fixed bias correction based on a climatology of first guess departures rather than using VarBC. The implementation of VarBC for new observation types requires development and careful tuning and it is planned to explore its implementation for radar and lidar observations in the future.

A fully functioning climatological bias correction scheme for radar and lidar observations was developed in Fielding and Janisková (2017b). In this scheme, height and temperature were used as indicators to provide an implicit regime dependence. The scheme used look-up tables for the climatological bias based on first guess departures. However, when implementing this approach for the new triple-column approach in the observation operator, some regional biases remained; in particular, there were some strong biases for arctic clouds (not shown). Also, with the plans to investigate the use of VarBC in mind, a parametric approach that is more suitable to parameter estimation has been devised.

The new bias correction scheme has been chosen to be as simple as possible, whilst capturing the majority of the biases inherent in the model and observation operator. The bias correction scheme uses a fit to a 5th order polynomial as a function of model temperature, T:

$$\eta(T) = a_0 + a_1 T + a_2 T^2 + a_3 T^3 + a_4 T^4 + a_5 T^5, \tag{4.9}$$

where a_i is the set of fitted coefficients and η is the bias. Temperature is chosen as the indicator as the microphysical assumptions are sensitive to temperature, and the microphysical assumptions are one of the dominant sources of error. To account for regime dependent biases, five different regions are used (Antarctic, Southern Hemisphere, Tropics, North Hemisphere and Arctic) based on thresholds in Latitude. Also, because the biases are found to be sensitive to hydrometeor phase, two different parameter sets are used for each region (one for model temperatures greater than 273 K and one for model temperatures less than 273 K).

To fit the coefficients in eq. 4.9, a month-long monitor-only (radar and lidar observations are not assimilated) is first run using the new observation operator and using model cycle 46R1. The biases in first guess departures with respect to CloudSat and CALIPSO observations (Fig. 4.10a and 4.11a respectively) from the monitor-only experiment are then fitted to eq. 4.9 using non-linear regression to obtain a smooth bias correction (Figs. 4.10b and 4.11b).

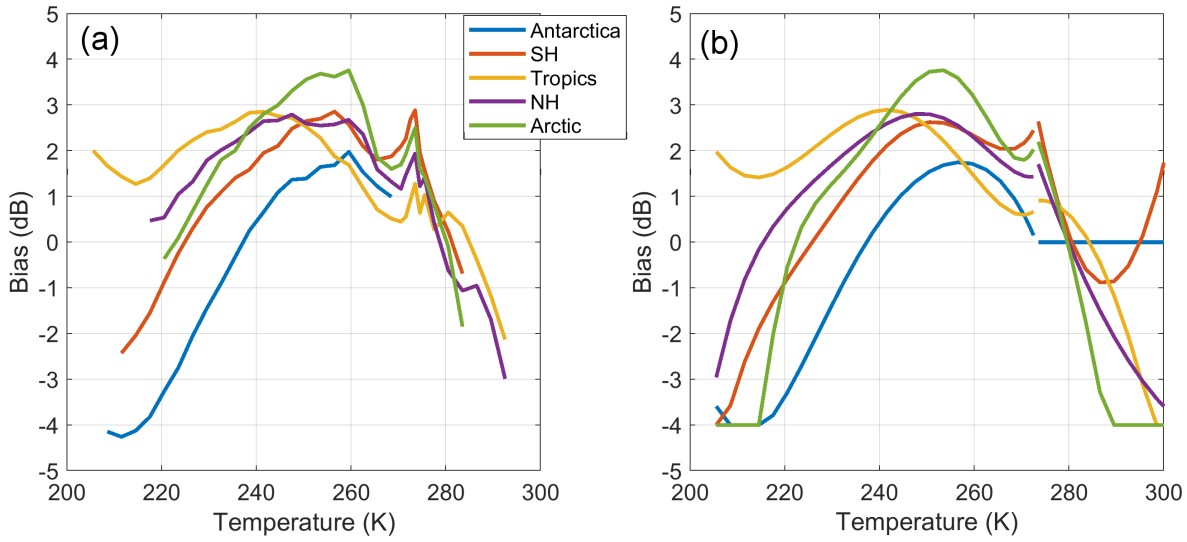

Figure 4.10: Panels show (a) Bias in first guess departures (dB) for CloudSat radar reflectivity during August 2007 as a function of model temperature and region, and (b) corresponding regime and temperature-dependent bias correction.

Firstly, for the biases in first guess departures of radar reflectivity (Fig. 4.10), there is a reasonably large range in bias for different regions, justifying the need for a regime dependent approach. Interestingly, observation minus model biases in the Antarctic are around 3 dB different in the Arctic. This difference warrants further investigation, but for now it is left for the bias correction scheme to correct. The temperature dependency of the bias is also justified with biases a strong function of temperature for all regions. A sharp local maxima in the biases is observed around 273 K corresponding to the melting layer. The cause of this bias is unlikely to be due to melting snowflakes themselves as melting particles do not produce a stronger signal at W-band Sassen et al. (2007), but more likely an underestimation of particle size by the observation operator or the too-rapid melting of snowflakes by the model.

Next, for the biases in first guess departures of lidar backscatter (4.11), a similar range in temperature-dependent bias is seen, but the regime dependence is less. In particular, the Arctic and Antarctic regions have very similar biases, suggesting that the different biases in radar reflectivity relates to precipitation,

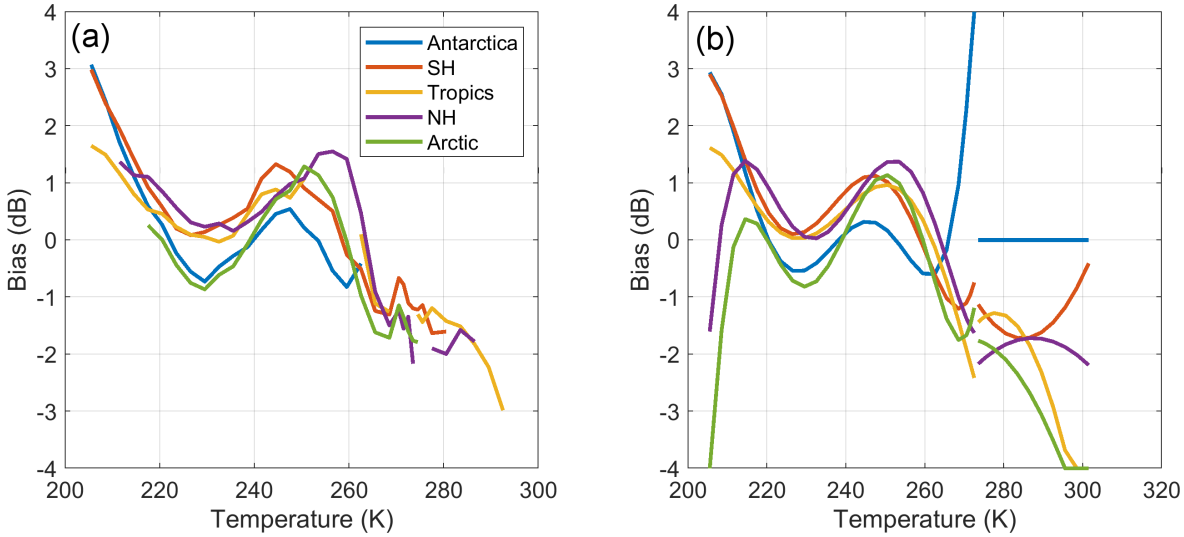

Figure 4.11: Same as Fig. 4.10, but for CALIPSO lidar backscatter.

of which the radar is more sensitive to. Between 220 K and 260 K, the bias is actually small (within 1 dB), so the bias correction is almost not required in these regions. For warmer temperatures, the model tends to have greater backscatter than the observations, while the converse is true for very cold clouds (less than 210K).

Once the bias correction schemes have been calibrated, their performance can be assessed by applying them to the monitor-only first guess departures and analysing the mean bias-corrected departures as a function of variables that are not directly dependent on temperature. For example, Fig. 4.12 shows the bias in radar reflectivity before and after the new bias correction as a function of latitude and pressure. Generally, the bias is completely removed, particularly for regions with the greatest amounts of clouds (mid-latitudes and tropical high-cloud). A small negative bias remains for the mid-troposphere relating to an over-correction of the bias for convective regions. The impact of this remaining bias is mitigated by the fact the observation errors are generally large in these regions due to the greater representativity error. Nevertheless, an improvement in the bias correction could likely be made if an indicator for convective situations was included.

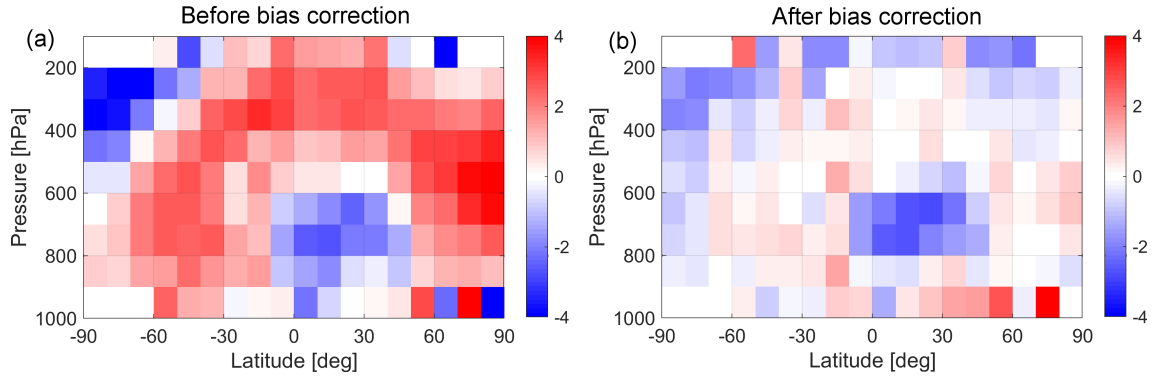

Figure 4.12: Bias in first guess departures (dB) for CloudSat radar reflectivity during August 2007 as a function of latitude and pressure. Panels show (a) before bias correction and (b) after bias correction.

Finally, the mean of the bias-corrected first guess departures for lidar backscatter (Fig. 4.13) performs extremely well, with no significant biases remaining when stratifying the biases by latitude and pressure. The lack of large regime-dependent biases probably helps the bias correction scheme, and also suggests that the model ice cloud is consistently represented across different regimes. The positive bias that remains for high ice-cloud in the Antarctic is not significant as very few clouds are found above 200 hPa at these latitudes (Stubenrauch et al., 2010; Hagihara et al., 2014).

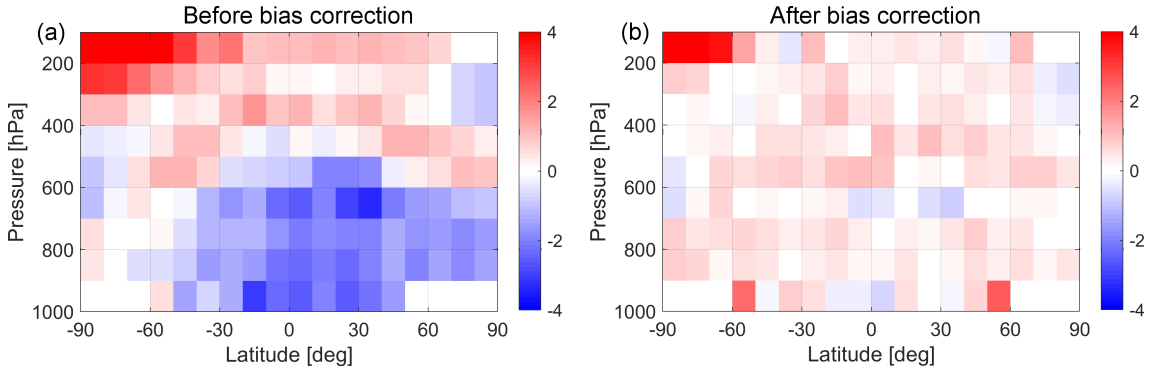

Figure 4.13: Same as Fig. 4.12, but for CALIPSO lidar backscatter.

5 Impact of updates

5.1 Experimental setup

Several assimilation experiments have been run using the full system of regularly assimilated observations at ECMWF to assess the impact of the updates described in Sec. 4. In addition to the updates, the model has also been updated to CY46R1 (see WP-1000). All experiments shown have been performed over the period of one months between 1 August 2007 and 31 August 2007 using cloud radar reflectivity from CloudSat and cloud lidar backscatter from CALIPSO. The results from the following experiments will be presented:

- **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;

- **baseline**: experiment assimilating cloud radar and lidar observations on top of all other normally assimilated observations using the observation operator in CY43R1. All screening, observation pre-processing and observation errors have been kept as close as possible to **RADLID_2err_Ha**, with cloud radar and lidar observations averaged to approximately 72 km.

- **updated**: same as **baseline**, but using the triple-column version of the observation operator and updated bias correction scheme.

- **updated+corr**: same as **updated**, but observation errors are assumed to be correlated using eq. 4.8 with $\alpha = 0.6$ and $\beta = 1.8$.

5.2 Impact on first guess and analysis departures of own observations

5.2.1 Radar reflectivity

The first step in evaluating the impact of assimilating the cloud and radar observations is to assess the impact on the observations themselves. Figure 5.1 shows the distribution of both first guess and analysis departures of radar reflectivity for the different experiments. Looking at the results globally, we can see that for all experiments, the standard deviation of analysis departures is smaller than the first guess departures, showing that the analysis brings the model closer to the observations. We can also see that experiments that assimilate the radar and lidar observations have better fits to the observations than the control. The closest fit to observations in the analysis is given by the **updated** experiment.

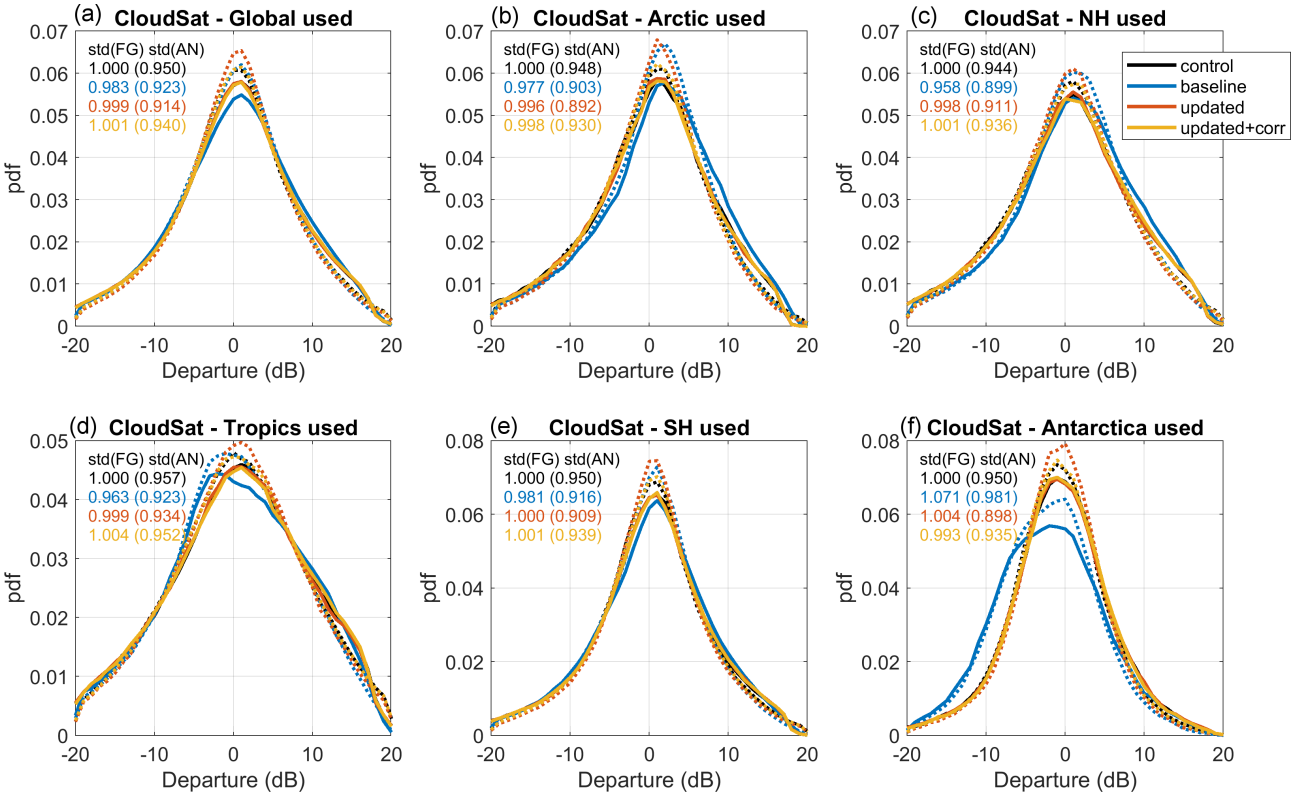

Figure 5.1: Probability density functions for first guess (solid lines) and analysis (dotted lines) departures between the control experiment (black), **baseline** *experiment (blue line),* **updated** *experiment (red line) and* **updated+corr** *experiment (yellow line) and cloud radar reflectivity using CloudSat observations. Panels show results subset by region: (a) Global, (b) Arctic, (c) Northern Hemisphere, (d) Tropics, (e) Southern Hemisphere and (f) Antarctica. Also shown in the top left of each panel are the standard deviation of first guess and analysis departures, both normalised by the standard deviation of first guess departures in the control.*

Taking a regional perspective, the new triple-column approach (used in all experiment except the **baseline** experiment) provides smooth unbiased Gaussian-like distributions of first guess and analysis departures. The improvements in the updates are particularly apparent in the Tropics and Antarctica, where the **baseline** experiment does not produce symmetric curves. In the tropics, the more-physical representation of attenuation given by the triple-column approach has improved the fit, while in the Antarctic, it is a combination of the bias correction and the operator updates. In general, the analysis departures of the **updated+corr** experiment are larger than the **updated** experiment but smaller than the control

experiment showing that the observation error correlations have reduced the impact of the observations. This is not surprising given that the observation error variances have the same scaling in each experiment. This suggests the observation errors could be reduced when applying the vertical observation error correlations in future experiments.

The performance of the bias correction scheme developed in Sec. 4.3 is evaluated in Fig. 5.2. In the global average above 800 hPa, the bias corrected first guess departures are almost entirely unbiased. The increase in bias for observations below 800 hPa is not worrying as, due to screening of observations affected by ground clutter below 1 km, there are few observations below 800 hPa that pass screening. However, in all experiments, the model analysis departures below 800 hPa tend to be larger, suggesting that the model has a tendency to try to remove precipitation from the background. Interestingly, this effect is almost as strong in the control (REF) as the other experiments, suggesting it is caused by other observations and their interaction with the model and 4D-Var. This is corroborated by the fact the mean analysis increments of cloud fraction and relative humidity in the REF experiment are both negative in the Southern Hemisphere and the tropics (Fig. 5.3).

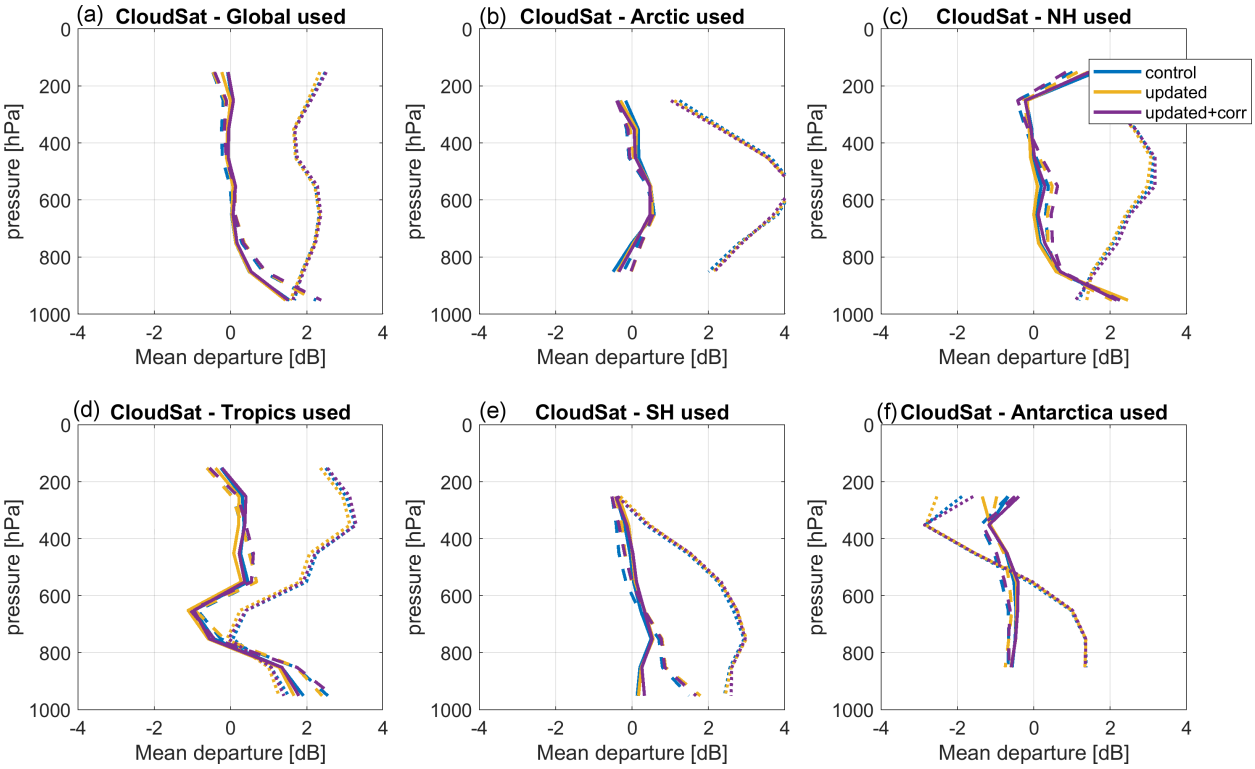

Figure 5.2: Same as Fig. 5.1, but for mean first guess (solid), first guess before bias correction (dotted) and analysis (dashed) departures as a function of pressure (hPa).

Finally, the height dependence of the first guess and analysis departures for radar reflectivity is evaluated in Fig. 5.4. For all regions and heights, the **updated** experiment has the smallest analysis departures. Again, the analysis fit to own observations for the experiment with correlated observation errors is not as strong as the experiment without vertical correlations, but they are improved compared to the control (REF) experiment. Reducing the inflation of the observation errors should allow a better analysis fit to the observations.

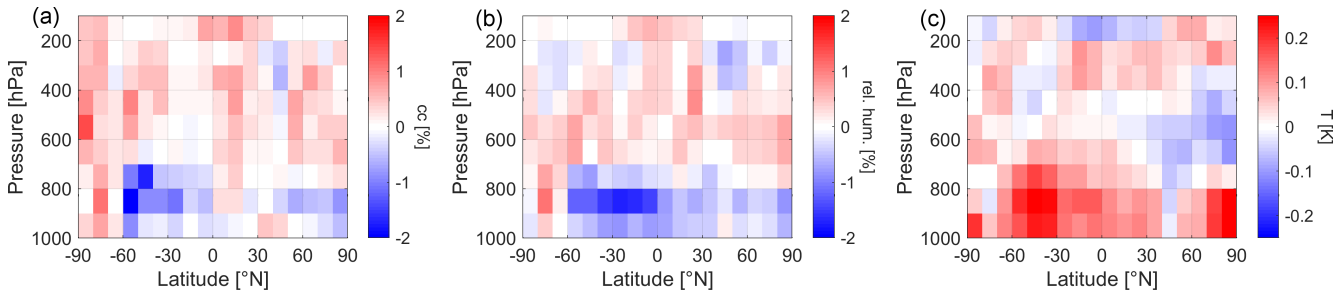

Figure 5.3: Mean analysis increments co-located with CloudSat observations for the REF experiment of (a) cloud fraction (%), (b) relative humidity (%) and (c) temperature (K).

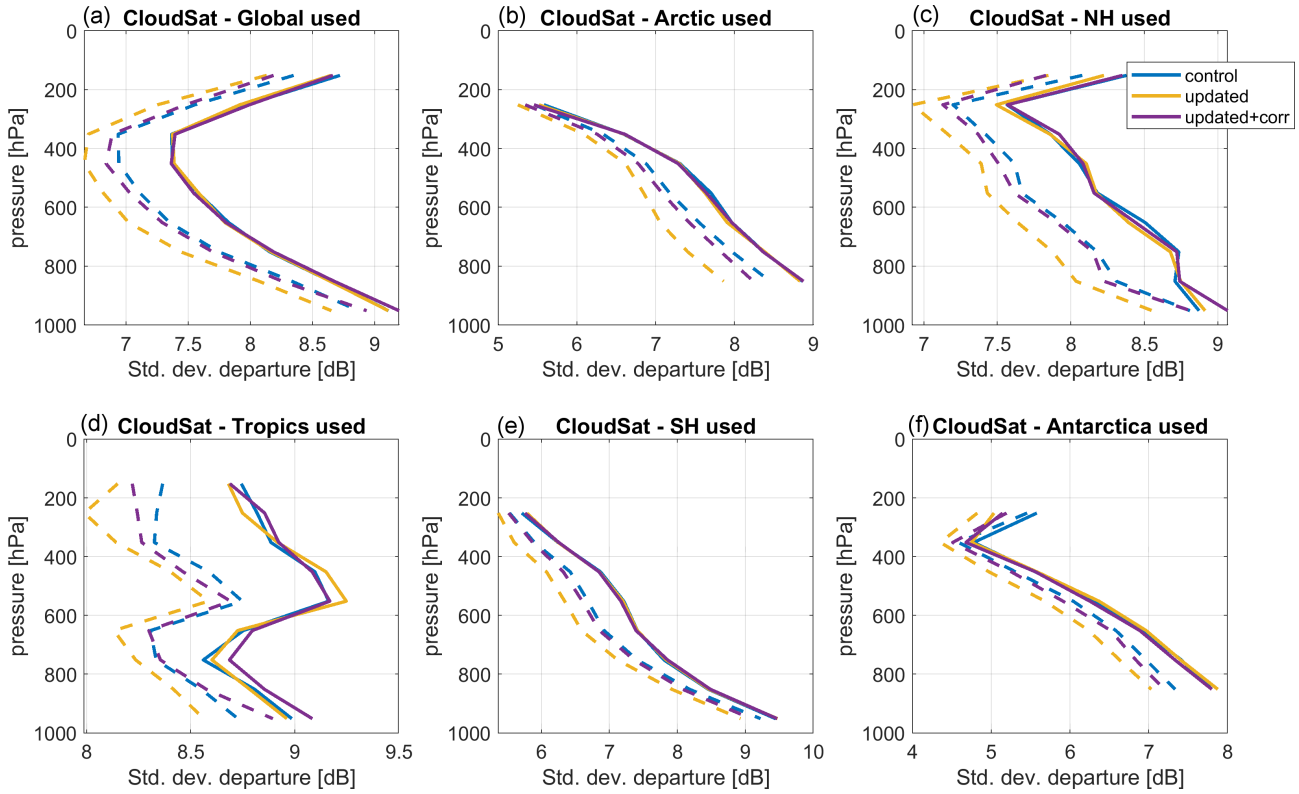

Figure 5.4: Same as Fig. 5.1, but for the standard deviation of first guess (solid) and analysis (dashed) departures.

5.2.2 *Lidar backscatter*

As was done for the radar, it is important to first evaluate the impact of observations on the analysis fit to themselves. Figure 5.5 shows the distribution of both first guess and analysis departures of lidar backscatter for the different experiments. In the global sense, whereas the single column operator used in the **baseline** experiment shows only a 1% reduction in analysis departure compared to the first guess, the **updated** experiment shows a 5% decrease in standard deviation of the model fit to observations. This result is very promising and shows the increased potential for gaining information from the lidar observations. A pair of experiments assimilating radar-only and lidar-only would help to elucidate how much benefit in increasing fit to the observations was coming from each instrument. However, it can be said that, as was seen in for the radar departures, all experiments that assimilate the radar and lidar

observations using the triple-column observation operator have better fits to the observations than the control.

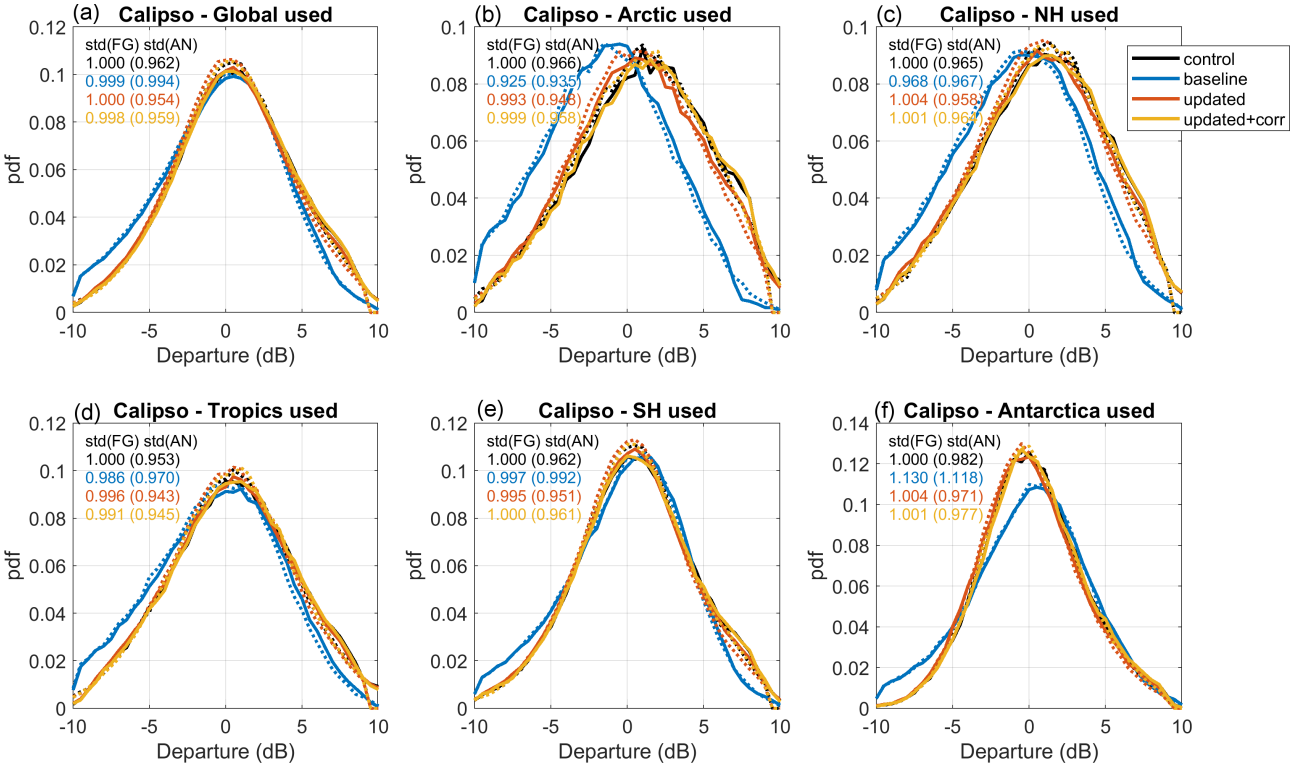

Figure 5.5: Same as Fig. 5.1, but for CALIPSO lidar backscatter.

The effect of the bias correction scheme and the presence of any biases in the analysis increments of lidar backscatter are evaluated in Fig. 5.6. While considering the biases it is important to remember that the majority of useful lidar observations are of the upper regions of ice cloud and away from polar regions this is likely to be at 500 hPa and above. For regions above 500 hPa, the bias correction scheme works well with mean departures less than 1 dB for all regions. In general the bias tends to increase towards the surface suggesting too much attenuation or insufficient representation of multiple scattering. One way to adjust this without using the bias correction scheme would be to increase the fractional standard deviation setting in the triple-column approach. Unlike the radar, there are no obvious biases in the mean analysis increments in lidar backscatter probably because the model and 4D-Var appear better balanced in the upper troposphere (Fig. 5.3). However, for the Arctic and other areas, there is a slight positive analysis increment (reduction in mean departure in Fig. 5.6f, suggesting the data assimilation is adding cloud).

5.3 Impact on first guess departures of other assimilated observations

As was done for the experiments using CY43R1, the next step in evaluating the impact of the assimilated cloud radar reflectivity and lidar backscatter is to compare results against other assimilated observations. This has been done by comparing the first-guess fits to other routinely assimilated observations for the one-month period of the experiments. For all figures, the results are based on statistics using averaged values over the whole globe.

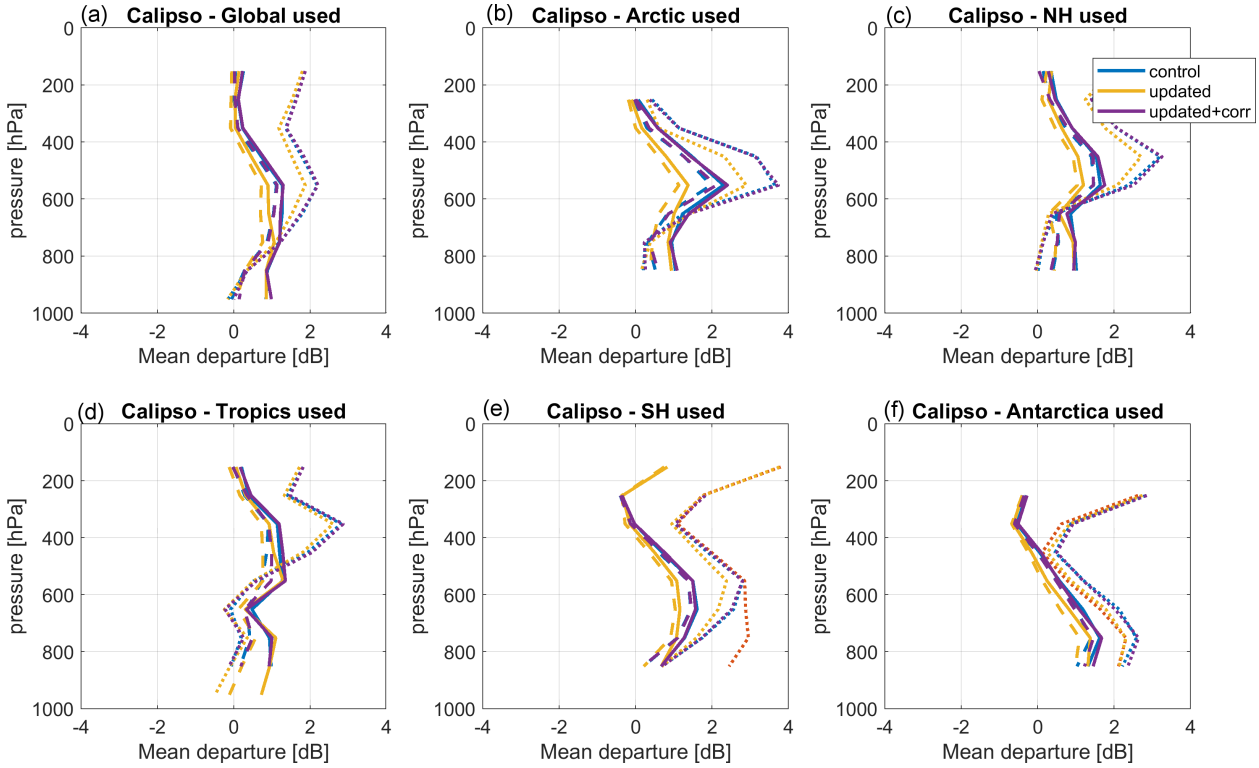

Figure 5.6: Same as Fig. 5.2, but for CALIPSO lidar backscatter.

Firstly, comparing the **baseline** experiment with the **RADLID_2err_Ha** experiment will assess the impact of moving cycles from CY43R1 to CY46R1. Overall, the first guess departures for selected satellite observations in the **baseline** experiment (Fig. 5.7) are very similar to the corresponding departures from the **RADLID_2err_Ha** experiment (Fig. 3.9). For AMSU-A, both experiments show improvements in channels 9-11 while the results are neutral for other channels. For HIRS, no channels show a significant signal in the **baseline** experiment relative to the reference experiment, as was the case for the **RADLID_2err_Ha** experiment. Finally, the departures for the motion vector winds provided by the SATOB measurements appear similar for both experiments, with a slight improvement at 250 hPa. Similar conclusions are drawn from comparing the first guess fits to conventional observations in both experiments (Fig. 3.10 vs. Fig. 5.8). From these results, it is clear that much longer experiments would be needed to differentiate the impact of cloud radar and lidar observations on CY43R1 and CY46R1, so we can conclude that the impact of moving cycles on the impact of assimilating cloud radar and lidar observations is small.

Next we can consider the relative impact of the updated observation operator on other assimilated observations by comparing the **updated** experiment with the **baseline** experiment. For the selected satellite observations in Fig. 5.7, the main impact appears to be in the upper tropospheric and lower stratospheric AMSUA-A channels 8-12 that are sensitive to temperature. AMSU-A channel 8's weighting function peaks at around 200 hPa, while channel 9 peaks around the stratopause at 100 hPa Kazumori et al. (2016). The weighting functions for channels 10-12 peak increasingly higher the atmosphere. Further investigation reveals (not shown) reveals that this impact is mainly seen in the tropics. Conversely, in the tropics these channels can be degraded by all-sky measurements (e.g., Kazumori et al., 2016), which has been attributed to gravity wave generation and propagation into the stratosphere. It is possible that the assimilation of the cloud structure provided by cloud radar and lidar observations helps to mitigate

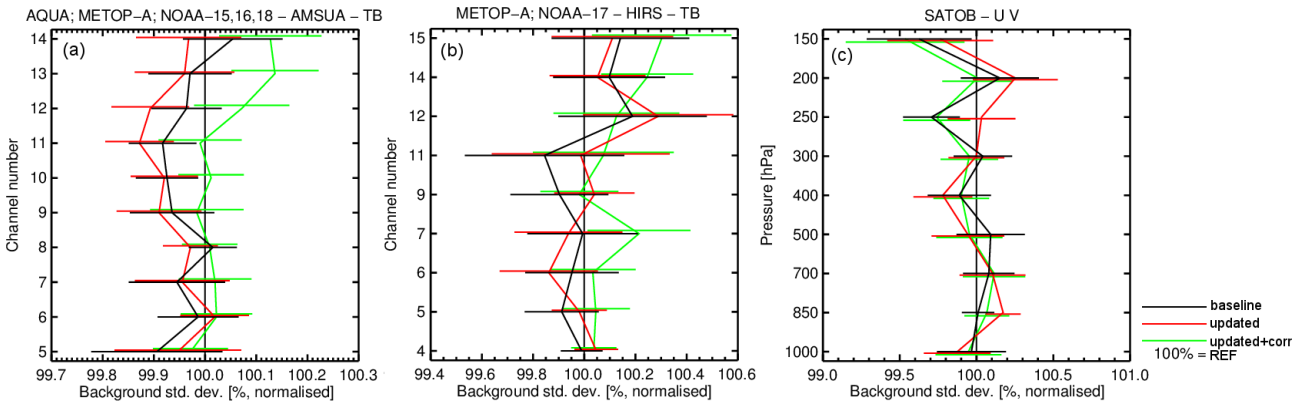

*Figure 5.7: Difference between experimental runs and the reference run (REF) in standard deviation normalised by the standard deviation of REF for the first guess fits to different observations: (a) AMSU-A tropospheric (channels 5-8) and stratospheric (channels 9-14) temperature, (b) HIRS temperature (channels 5-8, 13-15), water vapour (channels 10-12) and (c) SATOB wind. The reference run is represented by the 100% line and negative values correspond to improvements in the experimental runs. Horizontal bars indicate statistical significance at the 95% level. The results for experiments assimilating, in addition to other observations, cloud radar reflectivity and lidar backscatter, using the old operator (***baseline***, in black), updated operator (***updated***, in red) and updated operator with observation error correlations (***updated+corr***, in green), are shown for the whole globe over the period from 1 August to 31 August 2007.*

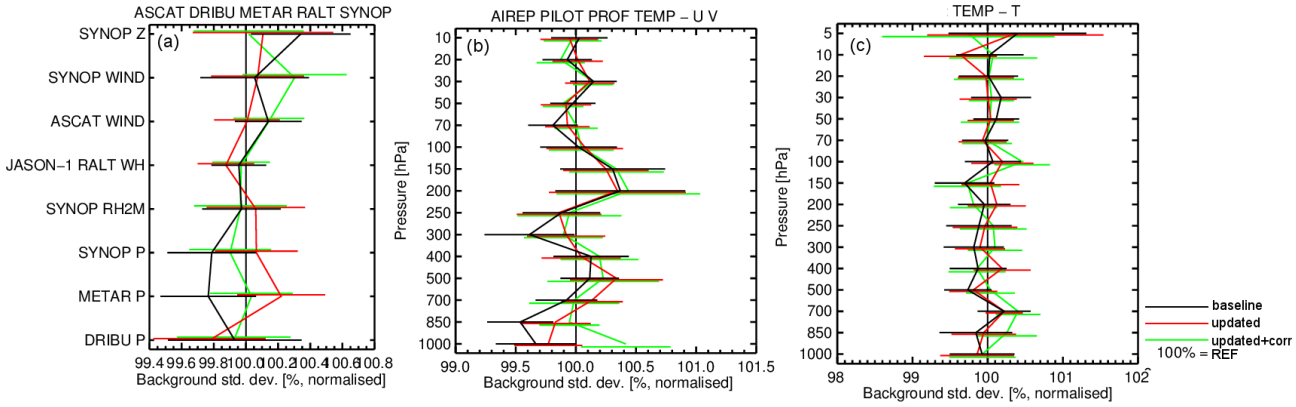

Figure 5.8: Same as Fig. 5.7, but for the first guess fits to different conventional observations: (a) surface pressure, relative humidity and 10-metre wind, (b) wind and (c) temperature profiles.

this gravity wave generation. More investigations are required to understand the source of these lower stratospheric benefits. For all other observation types, including conventional observations (Fig. 5.8), the impact of the updated observation operator is neutral and there are no significant differences between the experiments. Much longer experiments would likely be required to discern significant differences between the **baseline** experiment and the **updated** experiment, but the comparison does validate the new operator does not provide any significant degradation after one month of experimentation.

The effect of adding vertical error correlations to the observation error matrix appears to slightly degrade the first guess fits to satellite observations. For example, for the fit to AMSU-A observations (Fig. 5.7a), the **updated+corr** experiment is always worse than both the **baseline** and the **updated** experiments. Similarly, the fit to HIRS observations (Fig. 5.7b), is generally worse in the **updated+corr** experiment. However, the fit to conventional observations is generally similar to the other experiments. There are two reasons that accounting for the observation error correlations is not having a significantly beneficial im-

provement to the analysis. The first is that adding observation error correlations decreases the weight of the observations in the data assimilation if the observation error variances are not changed (as is the case in **updated+corr**). Therefore the relative degradations could be due to a lack of improvement compared to the reference experiment rather than an absolute degradation from the inclusion of the observation error correlations. Alternatively, the observation error correlations could be causing a degradation to the analysis. A reason for this is that observation error correlations can amplify inter-level biases if they exist (Geer, 2019). Further investigation is required to understand the performance of using observation error correlations, including scaling of the observation errors, verification of the subsequent forecast and assessment of the effect of inter-level biases.

6 Summary and conclusions

In this report, the impact of including observations of cloud radar reflectivity and cloud lidar backscatter within the ECMWF 4D-Var assimilation system was evaluated, with a focus on the impact on the model analysis. The report is separated in two distinct parts. The first part extends the evaluation made in Janisková and Fielding (2018b) from one month to three months. Using the IFS model cycle CY43R1, a wide range of experiments were performed to understand the behaviour of the new observation types in the data assimilation system. For example, a strong sensitivity of the analysis to observation error scaling and observation averaging size was found, with a doubling of observation errors required to avoid a degradation to short-term forecasts of humidity and temperature. When using double errors and the horizontal averaging to 72 km, the results obtained indicate that 4D-Var analyses are brought closer to the new observations and the impact on other assimilated observations is often positive, but generally not significant using a 95% confidence interval. In general it was also found that the impact of cloud radar reflectivity is greater than of the lidar backscatter.

The second part of this report focuses on improving the impact of cloud radar and lidar observations by developing a new triple-column observation operator to account for both cloud overlap and subgrid condensate variability. The new triple-column approach is shown to improve the physical realism of the operator while being cost efficient; the triple-column is much more efficient than the multi-column approach used for model evaluation and is also easily differentiable. Standard deviation of first-guess departures are reduced the most for lidar backscatter, but the new approach also has a substantial effect on radar reflectivity in heavier precipitation. In addition, a specification for the vertical observation error correlation based on an 'error inventory' approach is constructed and an update to the bias correction scheme is detailed. These updates are then merged with the updates of (Janisková and Fielding, 2022) so that testing of these updates can use the ECMWF model cycle CY46R1, which became operational in June 2019.

As was found in CY43R1, assimilating observations of cloud radar reflectivity and lidar backscatter from CloudSat and CALIPSO respectively in CY46R1 brings the model analysis closer to the observations themselves. The new triple-column approach is found to be particularly beneficial to the impact of the lidar backscatter, as it reduces the analysis fit relative to the background significantly compared to when using the single-column method. The impact of the new triple-column operator is mostly neutral when comparing the model first-guess fit to other assimilated observations such as AMSU-A and HIRS. The greatest impact on other assimilated observations appears to be in the lower stratospheric channels of AMSU-A; further investigation is required to deduce if the radar and lidar observations are improving the model representation of gravity waves which could affect these scores. Finally including the representation of vertically correlated observation errors has not had a significant positive impact on the model analysis. This could be for several reasons including: a need for the observation error variances to be re-tuned, amplifications of height-dependent biases or a poor representation of the true error correlations. A first step to improve the error correlation would be to include a phase-dependence to the microphysical error component.

While the report has provided a thorough assessment of the impact of cloud radar and lidar observations on the model analysis, the impact is intrinsically linked with the impact on the subsequent forecast, which will be explored in WP-3000. It is therefore likely that the findings and results in this work-package will be revisited and updated during the investigations for WP-3000 before a final version will appear as part of the final project report for PEARL Cloud project. The length of the experiments shown in this work package (one month), while suitable for assessing any large impacts on the model analysis, will need to be significantly longer for studying forecast impacts (3 months or longer), which are likely to be harder

to detect due to the chaotic nature of weather forecasts. These longer experiments for WP-3000 will have the additional benefit of adding more confidence to the results found in this report.

Acknowledgments

The authors would like to thank all the ECMWF colleagues who helped to port the cloud radar and lidar developments from CY43R1 to CY46R1, especially Peter Lean and Tomas Kral. Robin Hogan is thanked for discussions on the observation operator developments. Shannon Mason is thanked for providing the CAPTIVATE retrievals.

List of Acronyms

4D-Var	Four-Dimensional Variational Assimilation
AN	Analysis
AIREP	AIRcraft Weather REPort
ATLID	ATmospheric LIDar
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
CAPTIVATE	Cloud and Aerosol Property reTrIeval using a VAriational TEchnique
CloudSat	NASA's cloud radar mission
CPR	Cloud Profiling Radar
EarthCARE	Earth, Clouds, Aerosols and Radiation Explorer
ECMWF	European Centre for Medium Range Weather Forecasts
ESA	European Space Agency
FG	First Guess
HIRS	High-resolution Infrared Radiation Sounder
IFS	Integrated Forecasting System of ECMWF
NWP	Numerical Weather Prediction
OBS	OBServations
T95	Model grid with spectral truncation T95
T159	Model grid with spectral truncation T159
T255	Model grid with spectral truncation T255
TCo639	Model cubic octahedral grid with spectral truncation T639
UTC	Universal Time Coordinated
VarBC	Variational Bias Correction

References

Ahlgrimm, M. and R. M. Forbes, 2016: Regime dependence of cloud condensate variability observed at the atmospheric radiation measurement sites, *Quarterly Journal of the Royal Meteorological Society*, **142**(697), 1605–1617.

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.

Bormann, N., M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally, 2016: Enhancing the impact of iasi observations through an updated observation-error covariance matrix, *Quarterly Journal of the Royal Meteorological Society*, **142**(697), 1767–1780.

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., M. Ahlgrimm, R. Forbes, M. Kulie, R. Bennartz, M. Janisková, and P. Bauer, 2012: Iterpreting and evaluation of the ECMWF global model with CloudSat onbservations: ambiguities due to radar reflectivity forwad operator uncertainties, *Q. J. R. Meteorol. Soc.*, **138**, 2047–2065, doi:10.1002/qj.1936.

Fielding, M. and M. Janisková, 2017a: 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.

Fielding, M. and M. Janisková, 2020: Direct 4D-Var assimilation of space-borne cloud radar cloud radar reflectivity and lidar backscatter. Part I: Observation operator and implementation, *Q. J. R. Meteorol. Soc.*, **146**(733), 3877–3899.

Fielding, M. and O. Stiller, 2019: Characterizing the representativity error of cloud profiling observations for data assimilation, *J. Geophys. Res. - Atmospheres*, **124**, 4086–4103.

Fielding, M. D. and M. Janisková, 2017b: 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, 4000116891/16/NL/LvH, 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.

Fowler, A. M., S. L. Dance, and J. A. Waller, 2018: On the interaction of observation and prior error correlations in data assimilation, *Quarterly Journal of the Royal Meteorological Society*, **144**(710), 48–62.

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.

Geer, A. J., 2019: Correlated observation error models for assimilating all-sky infrared radiances, *Atmospheric Measurement Techniques*, **12**(7), 3629–3657.

Geer, A. J. and P. Bauer, 2011: Observation errors in all-sky data assimilation, *Quart. J. Roy. Meteorol. Soc.*, **137**, 2024–2037.

Hagihara, Y., H. Okamoto, and Z. J. Luo, 2014: Joint analysis of cloud top heights from cloudsat and calipso: New insights into cloud top microphysics, *Journal of Geophysical Research: Atmospheres*, **119**(7), 4087–4106.

Hill, P. G., C. J. Morcrette, and I. A. Boutle, 2015: A regime-dependent parametrization of subgrid-scale cloud water content variability, *Quarterly Journal of the Royal Meteorological Society*, **141**(691), 1975–1986.

Hogan, R. J. and A. Bozzo, 2018: A flexible and efficient radiation scheme for the ecmwf model, *Journal of Advances in Modeling Earth Systems*, **10**(8), 1990–2008.

Hogan, R. J. and A. J. Illingworth, 2000: Deriving cloud overlap statistics from radar, *Quarterly Journal of the Royal Meteorological Society*, **126**(569), 2903–2909.

Janisková, M. and M. Fielding, 2018a: Feasibility demonstration of 4D-Var assimilation system using CloudSat and CALIPSO observations, WP-5000 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, 35 pp.

Janisková, M. and M. Fielding, 2018b: Operational Assimilation of Space-borne radar and Lidar Cloud Profile Observations for Numerical Weather Prediction, Contract report to the European Space Agency 4000116891/16/NL/LvH, 164 pp.

Janisková, M. and M. Fielding, 2022: Assimilation system adaptation and maintenance, WP-1000 report for the project Preparations for EarthCARE Assimilation - Radar and Lidar Cloud Observations (PEARL Cloud), ESA ESTEC contract 4000128669/19/NL/CT, 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.

Kazumori, M., A. J. Geer, and S. J. English, 2016: Effects of all-sky assimilation of gcom-w/amsr2 radiances in the ecmwf numerical weather prediction system, *Quarterly Journal of the Royal Meteorological Society*, **142**(695), 721–737.

Marchand, R., G. G. Mace, T. Ackerman, and G. Stephens, 2008: Hydrometeor detection using cloudsat—an earth-orbiting 94-ghz cloud radar, *Journal of Atmospheric and Oceanic Technology*, **25**(4), 519–533.

Mason, S. L., C. J. Chiu, R. J. Hogan, D. Moisseev, and S. Kneifel, 2018: Retrievals of riming and snow density from vertically pointing doppler radars, *Journal of Geophysical Research: Atmospheres*, **123**(24), 13,807–13,834.

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.

Sassen, K., S. Matrosov, and J. Campbell, 2007: Cloudsat spaceborne 94 ghz radar bright bands in the melting layer: An attenuation-driven upside-down lidar analog, *Geophysical Research Letters*, **34**(16).

Shonk, J. K. P., R. J. Hogan, J. M. Edwards, and G. G. Mace, 2010: Effect of improving representation of horizontal and vertical cloud structure on the earth's global radiation budget. part i: Review and parametrization, *Quarterly Journal of the Royal Meteorological Society*, **136**(650), 1191–1204.

Shonk, J. P. and R. J. Hogan, 2008: Tripleclouds: An efficient method for representing horizontal cloud inhomogeneity in 1d radiation schemes by using three regions at each height, *Journal of Climate*, **21**(11), 2352–2370.

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.

Stubenrauch, C. J., S. Cros, A. Guignard, and N. Lamquin, 2010: A 6-year global cloud climatology from the atmospheric infrared sounder airs and a statistical analysis in synergy with calipso and cloudsat, *Atmospheric Chemistry and Physics*, **10**(15), 7197–7214.

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.

Young, S., M. Vaughan, R. Kuehn, and D. Winker, 2013: The retrieval of profiles of particulate extinction from cloud–aerosol lidar and infrared pathfinder satellite observations (calipso) data: Uncertainty and error sensitivity analyses, *J. Atmos. and Ocean. Tech.*, **30**(3), 395–428.