ESA CONTRACT REPORT

Contract Report to the European Space Agency

Support-to-Science-Element (STSE) Study
EarthCARE Assimilation

WP-3200 report: Assimilation experiments for radar and lidar

January 2014

Authors: M. Janisková

ESA ESTEC contract 4000102816/11/NL/CT

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

ECMWF

Series: ECMWF ESA Project Report Series

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

Contact: library@ecmwf.int

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

Contract Report to the European Space Agency

Support-to-Science-Element (STSE) Study EarthCARE Assimilation

WP-3200 report: Assimilation experiments for radar and lidar

Authors: M. Janisková

ESA ESTEC contract 4000102816/11/NL/CT

January 2014

ABSTRACT

1D-Var assimilation experiments have been performed using observations of cloud radar reflectivity and lidar backscatter, either separately or in combination. Information on temperature and specific humidity retrieved from 1D-Var was used as pseudo-observations in the 4D-Var system. Obtained results indicate that 1D-Var analyses get closer to assimilated and also independent observations. However, impact of the cloud radar reflectivity is larger than that of the lidar backscatter. 1D+4D-Var analysis reduces analysis departures for q pseudo-observations and provides analysis departures closer to T pseudo-observations than would be obtained if temperature pseudo-observations were not assimilated. Impact on the first-guess and analysis departure statistics when verified against other observation types assimilated in 4D-Var is small. The performed 1D+4D-Var assimilation experiments also suggest a positive impact of the new observations on the subsequent forecast.

Contents

1 Introduction

During the last decade the representation of precipitation and clouds in the global NWP models has greatly improved and has achieved a reasonable degree of realism. This opens new possibilities for improvement of the atmospheric initial state and for model improvement to be explored through assimilation of data related to clouds from active and passive instruments. Observations providing three dimensional information on clouds from space-borne active instruments on board of CloudSat and CALIPSO are already available and new ones, such as EarthCARE, should appear in the near future. The challenge is to identify information that can be extracted from such data sources and transformed through the model into knowledge about the atmospheric state.

In order to study the impact of the new observations on analyses and subsequent forecast, a technique combining one-dimensional variational (1D-Var) assimilation with four-dimensional variational (4D-Var) data assimilation has been selected since it would be difficult to start our study in the framework of the full 4D-Var system which is very complex and thus quite difficult to interpret. In the past, it was proven that the 1D-Var approach can provide very useful experience on how to assimilate new types of observations (e.g., ECMWF studies towards the assimilation of cloud/rain-affected microwave/infrared radiances - Marécal and Mahfouf, 2000; Moreau et al., 2004; Bauer et al., 2006a). Experience from assimilation studies using ARM (Atmospheric Radiation Measurement) radar data in 1D-Var (Benedetti and Janisková, 2004; Janisková, 2004) or even 2D-Var (Lopez et al., 2006) can also be beneficial. Therefore 1D-Var systems for the assimilation of different cloud-related observations from CloudSat and CALIPSO have been developed and are presented in this report.

The two-step 1D+4D-Var approach proposed for this study was used operationally for assimilation of precipitation related observations (Bauer et al., 2006a,b) at ECMWF from June 2005 to June 2008. Before that, benefits of such method had been demonstrated for rain rate observations by Marécal and Mahfouf (2000, 2002). This technique has also proven to be successful in early implementation of clear-sky infrared radiance assimilation in the past (e.g. Eyre et al., 1993; Phalippou, 2005) and for SSM/I brightness temperatures in clear sky areas (Gérard and Saunders, 1999).

In this study, the 1D-Var retrieval first runs on the set of CloudSat radar reflectivity and CALIPSO lidar backscatter observations either separately or in combination to produce pseudo-observations of temperature and specific humidity, which are then assimilated in ECMWF 4D-Var system. Section 2 describes the general methodology of the 1D-Var system and provides information about the observation operators (moist parametrization schemes, reflectivity model and parametrization of lidar backscatter due to clouds), observation errors and handling of observations such as quality control and bias correction. Results from 1D-Var experiments for CloudSat cloud radar reflectivity are presented in Section 3. Section 4 presents results from 1D-Var experiments for CALIPSO cloud lidar backscatter, while the results from 1D-Var experiments for combined radar and lidar observations are summarized in Section 5. A summary of 1D-Var experiments is provided in Section 6. Section 7 introduces the 1D+4D-Var technique applied to the assimilation of cloud observations from CloudSat and CALIPSO, as well as briefly summarizes the 4D-Var system together with pseudo-observations and their errors used in 4D-Var. Results from the 1D+4D-Var experiments are introduced and discussed in Section 8. Finally, Section 9 provides a summary and draws the conclusions of this study.

2 1D-Var assimilation system

2.1 Description of the 1D-Var system

When using 1D-Var assimilation technique, for a given observation **y**, 1D-Var searches for the model state **x** (defined by temperature and humidity profiles) that minimizes the following cost function

$$\mathscr{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b) + \frac{1}{2}(H(\mathbf{x})-\mathbf{y}^o)^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}^o) \tag{2.1}$$

where **B** is the covariance matrix of the background error taken from the ECMWF 4D-Var system. **R** represents observation and representativeness error covariance matrix. H is the observation operator providing the equivalent of the data from the model variable **x**. In the case of radar/lidar assimilation, it employs physical parametrization schemes for moist processes (convection and large-scale cloud schemes), parametrization of cloud radar reflectivity and lidar backscatter due to clouds.

2.2 Model background part of the 1D-Var system

2.2.1 Background values

The background values have been taken from the 12-hour forecast of the ECMWF model with T799 spectral truncation (corresponding to approximately 25 km) and 91 vertical levels. The forecast results have been stored every half an hour in order to use observations in 1D-Var in similar way as in the operational 4D-Var system where all observations are split into half-hour time slots.

The profiles of temperature T and specific humidity q, along with surface pressure, p_s, tendencies, and surface quantities are first used in the moist physics routines (simplified convection and cloud schemes described below) to compute cloud properties (cloud cover, ice and liquid-water contents) and precipitation fluxes. A radar or lidar observation operator is then applied to the model fields to obtain the equivalent model reflectivity or backscatter due to clouds, respectively.

2.2.2 Background error statistics

The background error covariance matrix **B** provides to the variational analysis the appropriate information about the 1D (vertical) statistical structure of the forecast errors. The covariance matrix of the background error **B** is taken from the operational ECMWF 4D-Var system. No cross correlations between the background errors of specific humidity and temperature are considered. The standard deviation of temperature over the vertical is about 0.5 K to around 800 hPa, then slowly decreasing to about 0.3 K around 50 hPa and finally grows to 0.7 K in the top model levels. The standard deviation for specific humidity has been empirically specified by Rabier et al. (1998) as a function of temperature and specific humidity profiles. The vertical profile has a maximum around 850 hPa, an exponential decrease above and lower values in the boundary layer. An example of vertical profile of standard deviations for temperature and humidity is illustrated on Fig. 2.1.

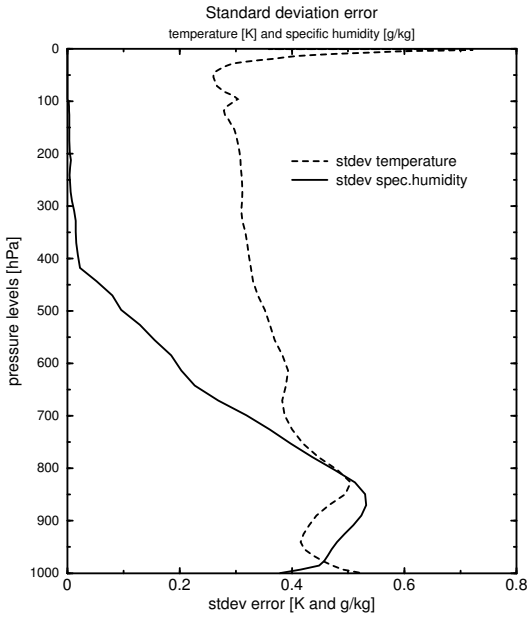

Figure 2.1: Vertical profiles of typical values of the standard deviation of the ECMWF model background errors for temperature (dashed line) and specific humidity (solid line). Units are in K and g kg^{-1} respectively.

2.3 Observational part of the 1D-Var system

2.3.1 Observations

In our study, measurements of cloud radar reflectivity, converted to mm^6 m^{-3} (level-1 product), from the CloudSat 94 GHz radar and/or lidar backscatter (km^{-1} sr^{-1}) due to clouds at 532 nm from CALIPSO are assimilated by 1D-Var system.

2.3.2 Observation errors

The impact of any type of observation in data assimilation is partly determined by the errors that are assigned to them. The magnitude of the observation error determines the weight to be given to the observation during the assimilation process. If the assumed error is large, this weight is small and so is the potential impact of the corresponding observation. However, under-weighting of observations will not lead to larger analysis errors. On the other hand, an underestimation of the error may have a detrimental impact on the assimilation process as the analysis is drawn closer to the observation than the actual quality (reliability) of the measured values would justify. These assumed observation errors take into account the instrument errors, together with forward modelling and representativity errors due to the narrow field of view. For CloudSat, instrument random error is used as described by Di Michele et al. (2014b). For CALIPSO, instrument errors are evaluated from the level-1 data according to (Liu et al., 2006). Forward modelling errors for both CloudSat and CALIPSO are based on evaluation of uncertainty in the microphysical assumptions. These errors are defined through differences between the perturbed state (obtained by small modification in microphysical assumptions) and the reference configuration (Di Michele et al., 2014b,a). This is done for different ranges of temperature. The used representativity error provides the flow dependent error which is estimated based on a statistical approach using the structure function maximum according to Stiller (2010) . This error is defined for the different altitudes and geographical regions.

2.3.3 *Observation operators*

In the case of radar/lidar assimilation, the observation operator H employs physical parametrization schemes for moist processes, i.e. the convection scheme and cloud scheme simulating large-scale condensation and precipitation processes (Lopez and Moreau, 2005; Tompkins and Janisková, 2004; Janisková and Lopez, 2013). Further parametrizations of cloud radar reflectivity and lidar backscatter due to clouds (Di Michele et al., 2012, 2014b,a) are required to convert model fields to reflectivity and backscatter, respectively. Input cloud and precipitation fields to radar and lidar models are computed by the moist physics parametrizations, the main input variables of which are temperature and humidity fields (i.e., control variables of the 1D-Var assimilation system).

(a) Convection scheme

All types of convection (shallow, mid-level and deep) are considered by the simplified convection scheme as described by Lopez and Moreau (2005). For deep and mid-level convection the link between the model control variables and the cloud base mass flux (i.e. closure assumption) is based on the release of convective available potential energy in time. For shallow convection, the closure assumption links the moist energy excess at cloud base to the moisture energy convergence inside the sub-cloud layer. The formation of precipitation from the cloud water contained in the updraught is parametrized according to Sundqvist et al. (1989) and a simple representation of precipitation evaporation is included. Precipitation formed from cloud liquid water at temperatures below the freezing point is assumed to freeze instantly.

(b) Cloud scheme

The original version of the simplified diagnostic scheme for the large-scale clouds and precipitation used in the variational assimilation at ECMWF is described in Tompkins and Janisková (2004). The scheme diagnoses cloud cover and cloud water (liquid+ice) from the input profiles of temperature and humidity. It describes the formation of stratiform cloud through large-scale motions, convective generation of cloud, precipitation generation and evaporation. The scheme assumes the subgrid scale variability of humidity represented by a uniform distribution allowing the in-cloud liquid water to be determined by integrating the saturated portion of the grid box. The impact of convective activity on large-scale clouds (important in the tropics and mid-latitudes summers) is accounted for through the detrainment rate produced by the convection scheme. This detrainment term is used to compute the additional cloud cover and cloud condensate resulting from convection. The formation of precipitation from cloud condensate is parametrized according to Sundqvist et al. (1989). The precipitation evaporation is estimated from the overlap of precipitation with the uniformly distributed subgrid fluctuations of humidity inside the clear fraction of the grid box.

(c) Reflectivity model

The reflectivity model for variational assimilation (ZmVar) is designed to meet the requirements of data assimilation system, i.e. to be computationally efficient and to allow the coding of its adjoint counterpart. A detail description can be found in the WP-1100 report (Di Michele et al., 2014b). For computational efficiency, ZmVAR works using a pre-calculated table of hydrometeor optical properties, where particle single scattering properties are obtained through Mie calculations based on the spherical assumption. Ice particle density is a function of the diameter. For rain, the normalized gamma particle size distribution (PSD) proposed by Illingworth and Blackman (2002) is used enabling to describe the shape of the size distribution for a specific precipitation regime (from light stratiform to convective precipitation). For the snow PSD, the parametrization developed by Field et al. (2007) is used. PSD of cloud ice is based on a modified version of the Field et al. (2007) distribution which does not consider aggregation. Based on radar frequency, a table of hydrometeor optical properties containing volumetric extinction and equivalent reflectivity (specified for predefined hydrometeor types on a range of temperatures and water contents) is pre-calculated. Accurate simulation of reflectivities at the CloudSat frequency also requires a proper modelling of the optical properties of frozen cloud particles.

(d) Backscatter model

The lidar simulator in clouds has been developed as an extension of the reflectivity model for variational assimilation (ZmVar), described in Di Michele et al. (2014b). A new look-up table of hydrometeor optical properties has been generated to allow ZmVar to work at the lidar wavelengths (Di Michele et al., 2014a). The PSDs assumed for each hydrometeor for the radar are also used for the generation of the lidar table. At visible wavelengths used by the lidar, the signal is subject to scattering (with negligible absorption) from the molecular components of the atmosphere. In order to simulate the apparent backscatter as measured by the lidar, the propagation of the backscatter signal along the vertical is modelled involving the evaluation of the signal attenuation and the multiple scattering effects, which are not negligible in a lidar system. A very simple multiple scattering approach is used for the linearized model (used in the assimilation) based on a small angle correction function as described in WP-1200 (Di Michele et al., 2014a).

2.4 Data handling

For a proper handling of observations in the context of an assimilation system, an appropriate quality control strategy and a scheme for bias correction (Di Michele et al., 2014b,a) are required.

2.4.1 Data quality control

A quality control of observations is required in order to discard measurements flagged as of poor quality or leading to excessive first guess (FG) departures (i.e. very large differences between observed and corresponding simulated parameters). The aim of quality control is to avoid the assimilation system to perform sub-optimally in cases of large model uncertainties or excessive non-linear effects in the observation operator. Using an approach described in WP-1100 (Di Michele et al., 2014b) for radar and in WP-1200 (Di Michele et al., 2014a) for lidar, the quality control of CloudSat radar and CALIPSO lidar observations is based on statistics of the FG departures, i.e. the differences between observed reflectivities/backscatters and corresponding simulated observations generated by the ZmVar forward operator. Observations with extreme FG departures are blacklisted, i.e. not taken into account. In addition, cases with intense convection as identified by the model are also discarded because of the model difficulty to accurately represent these events.

2.4.2 Bias correction

Since assimilation relies on the assumption that both observations and model background are unbiased quantities, it is necessary to quantify these systematic errors and to remove them. This is achieved by applying a bias correction scheme which is able to guarantee zero-mean statistics. Based on the statistics of FG departures, the applied bias correction is designed using temperature and altitude as predictors, separately over seasons and geographical regions to take into account the variability linked to the different weather regimes and phase of hydrometeors of clouds.

2.5 Experimental framework

Using 1D-Var system, different experiments have been performed for the selected situations:

- single track displayed in Fig. 3.1 (also red track in Fig. 8.1) between 23:50 UTC on 23 January 2007 and 00:26 UTC on 24 January 2007 over the whole Pacific Ocean from approximately 62°N to 62°S covering a variety of meteorological situations (e.g. tropical convection and an extra-tropical cyclone in the north),

- multiple tracks (Fig. 8.1) covering the 12-hour period from 21:00 UTC on 23 January 2007 to 09:00 UTC on 24 January 2007 which corresponds to the length of the 4D-Var assimilation window used at ECMWF.

The experiments using the following observations have been done:

- **R** (or **radar**) - cloud radar reflectivity from CloudSat 94 GHz radar,
- **L** (or **lidar**) - lidar backscatter due to clouds from CALIPSO lidar at 532 nm,
- **C** (or **combi**) - combination of cloud radar reflectivity and lidar backscatter.

Observations have been averaged in the grid-box at T799 resolution (corresponding to approximately 25 km) and have been used with:

- **nreg** - simple observation error estimate and rough quality control (as defined in Janisková et al., 2010; Janisková et al., 2012) without any bias correction,
- **bc** - applying bias correction on top of **nreg**,
- **qcbc** - applying bias correction and quality control with simple observation error definition (as **nreg**),
- **qcbcer** - applying complex observation error definition, bias correction and quality control.

The performance of 1D-Var assimilation has been verified against independent observations (i.e. observations which were not assimilated), such as cloud optical depth from MODIS (at the standard reference wavelength of 0.55 μm) or radar reflectivity and lidar backscatter when not assimilated. Analysis increments of temperature and specific humidity have also been evaluated since they can provide information about the impact of assimilated observations on the control variables of 1D-Var system.

3 1D-Var experiments for CloudSat cloud radar reflectivity

3.1 Experimental setup

Several 1D-Var experiments have been run mostly for the situation over the Pacific Ocean on 23-24 January 2007 using cloud radar reflectivity observations from 94 GHz CloudSat radar with the different setups described in Section 2.5 in oder to study the impact of the complex observation error definition, bias correction and quality control on the performance of 1D-Var assimilation. The satellite track for this situation is displayed in Fig. 3.1 and for CloudSat data the track covers 705 model grid points at T799 model resolution. This track crosses a large extra-tropical cyclone in the north; tropical convection between 10°N and 20°S; a large, vertically well developed system in the southern hemisphere between 22°S and 35°S; and another system in the south between 45°S and 60°S.

The model background values have been taken from the 12-hour forecast of the ECMWF model with T799 spectral truncation and 91 vertical levels.

3.2 Results

Figure 3.2 shows a comparison of the first-guess and 1D-Var retrieved radar reflectivities versus CloudSat 94 GHz radar over the Pacific Ocean. Though all assimilation experiments with the different observation handling (i.e. **nreg**, **bc**, **qcbc**, **qcbcer**) in general provide the 1D-Var analysis closer to the observations for most of the profiles, there are apparent differences between them. At first glance, using a basic treatment for observations and their errors (**nreg**) allows analysis to get closer to the observations than any other experiments. Only a close look reveals that the 1D-Var without stricter control of observation usage can be over-active and

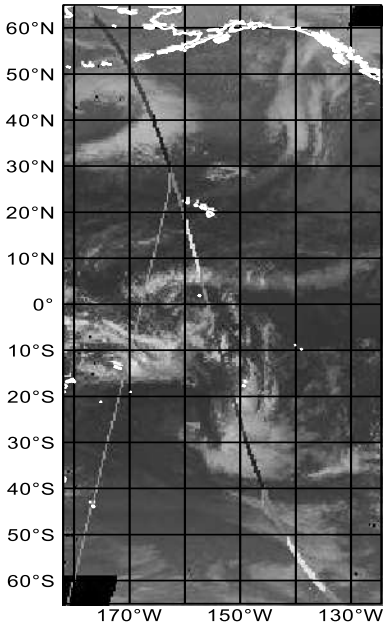

Figure 3.1: Situation over the Pacific Ocean between 23:50 UTC on 23 January 2007 and 0:26 UTC on 24 January 2007 used in 1D-Var experimentation

therefore removing or reducing clouds further than indicated by the observations, which is obvious for instance around 167 °-172°W or 137°-141°W. Applying quality control (**qcbc**) removes situations with large differences between the observed and the simulated reflectivities from 1D-Var computations, such as as between 156° -158°W or around 150°W (mostly up to 5km). The role of complex error definition, especially flow dependent representativity error, is mostly obvious in the tropical area approximately between 20°N and 5°S. In this area with the tropical convection creating localized clouds, the representativity error is large which leads to only small modifications in 1D-Var analysis with respect to FG to account for the fact that "patchy" clouds in observations may not be representative of the model grid-box situation. This behaviour is important for an optimal performance of the 1D-Var system which should only draw the model state close to the observations within their error bars and which the model is able to represent properly.

The retrieved values of temperature and specific humidity bring the 1D-Var analysis closer not only to the assimilated cloud radar reflectivity observations, but also to independent observations of lidar backscatter due to clouds as illustrated in Fig. 3.3. This is also demonstrated in Fig. 3.5 displaying the vertical profiles of root mean square (rms) errors for differences between the first-guess departures (differences between observations and the model first guess) and analysis (AN) departures (differences between observations and analysis) for both CloudSat radar reflectivity (a) and CALIPSO lidar backscatter (b). Only when applying bias correction, quality control and complex error definition (experiment **qcbcer**), the analysis departures are smaller than the first-guess ones through the whole vertical profile for independent lidar backscatter observations confirming the importance of proper observation handling in the assimilation system.

Figure 3.4 shows the probability distribution function (PDF) of the FG and AN departures for the assimilated cloud radar reflectivity and independent lidar backscatter obtained from the different assimilation experiments (**nreg**, **qcbc**, **qcbcer**). The PDFs of analysis departures for both cloud radar reflectivity and lidar backscatter become more narrow compared to the FG departures, indicating that the analyses are getting closer to the observations. The most symmetric PDF shapes for assimilated and independent observations are achieved by **qcbcer** experiments.

Comparisons of the first guess and analysis against independent MODIS observations of cloud optical depth

are summarized in Fig. 3.6-3.7. The statistics have been run for the cases when the model values (first-guess and analyzed) were smaller or equal to 50, which is the maximum range of this observational product. The analyses get closer to the MODIS cloud optical depth for all assimilation experiments with the best results being obtained, as expected, by the **qcbcer** experiment.

Results from the assimilation experiments are also summarized in Table 3.1 in terms of bias, standard deviation (stdv), as well as mean absolute errors (mae) and root-mean square errors (rms) and in Table 3.2 in terms of differences between the absolute value of analysis minus observations and absolute value of the first guess minus observations as well as differences of rms values for the FG and AN departures from the CloudSat cloud radar reflectivity and the CALIPSO cloud backscatter. The smallest errors (bias, stdv, mae, rms) and the largest reduction of rms errors for cloud radar reflectivity are achieved by the **qcbcer** experiment.

Analysis increments of specific humidity (Fig. 3.8) and temperature (Fig. 3.9) have also been evaluated. This evaluation revealed that both increments are modified by the assimilation of cloud radar reflectivity. Therefore the pseudo-observations of both temperature and specific humidity profiles from 1D-Var retrievals should be included in the 4D-Var system. When applying bias correction, quality control and complex observation error specification, increments are smaller since (i) some situations with extreme departures are excluded from 1D-Var assimilation by the quality control, (ii) 1D-Var is not correcting for the model bias and (iii) analysis is not driven as close to the observations due to observation error definition, especially in the cases with representativity issues, like in the tropics.

	reflectivity				backscatter			
	bias	stdv	mae	rms	bias	stdv	mae	rms
FG	0.55	4.26	1.86	4.30	7.72E-03	3.30E-02	1.46E-02	3.39E-02
AN-R_nreg	0.73	3.73	1.34	3.80	6.90E-03	3.19E-02	1.37E-02	3.27E-02
AN-R_qcbc	0.67	3.85	1.47	3.91	7.77E-03	3.25E-02	1.40E-02	3.34E-02
AN-R_qcbcer	0.48	3.70	1.38	3.74	7.33E-03	3.22E-02	1.39E-02	3.30E-02

Table 3.1: Bias, standard deviation (stdv), mean absolute error (mae) and root mean square error (rms) of the first guess (FG) and analysis (AN) departures for the different assimilation experiments (see text for experiment description) from the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) and the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) observations. 705 profiles are included in the statistics for the case of 24 January 2007.

	reflectivity		backscatter	
	DIFF	RMSD	DIFF	RMSD
AN-R_nreg	-0.635	-0.682	-7.879E-04	-9.592E-04
AN-R_qcbc	-0.677	-0.706	-7.611E-04	-8.096E-04
AN-R_qcbcer	-0.651	-0.733	-7.031E-04	-9.871E-04

Table 3.2: Difference between the absolute value of analysis (AN) minus observation (OBS) and absolute value of the first guess (FG) minus OBS (DIFF), as well as difference of rms values for the FG and AN departures (RMSD) from the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) and the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) observations. Results are shown for the different assimilation experiments (see text for experiment description). 705 profiles are included in the statistics for the case of 24 January 2007.

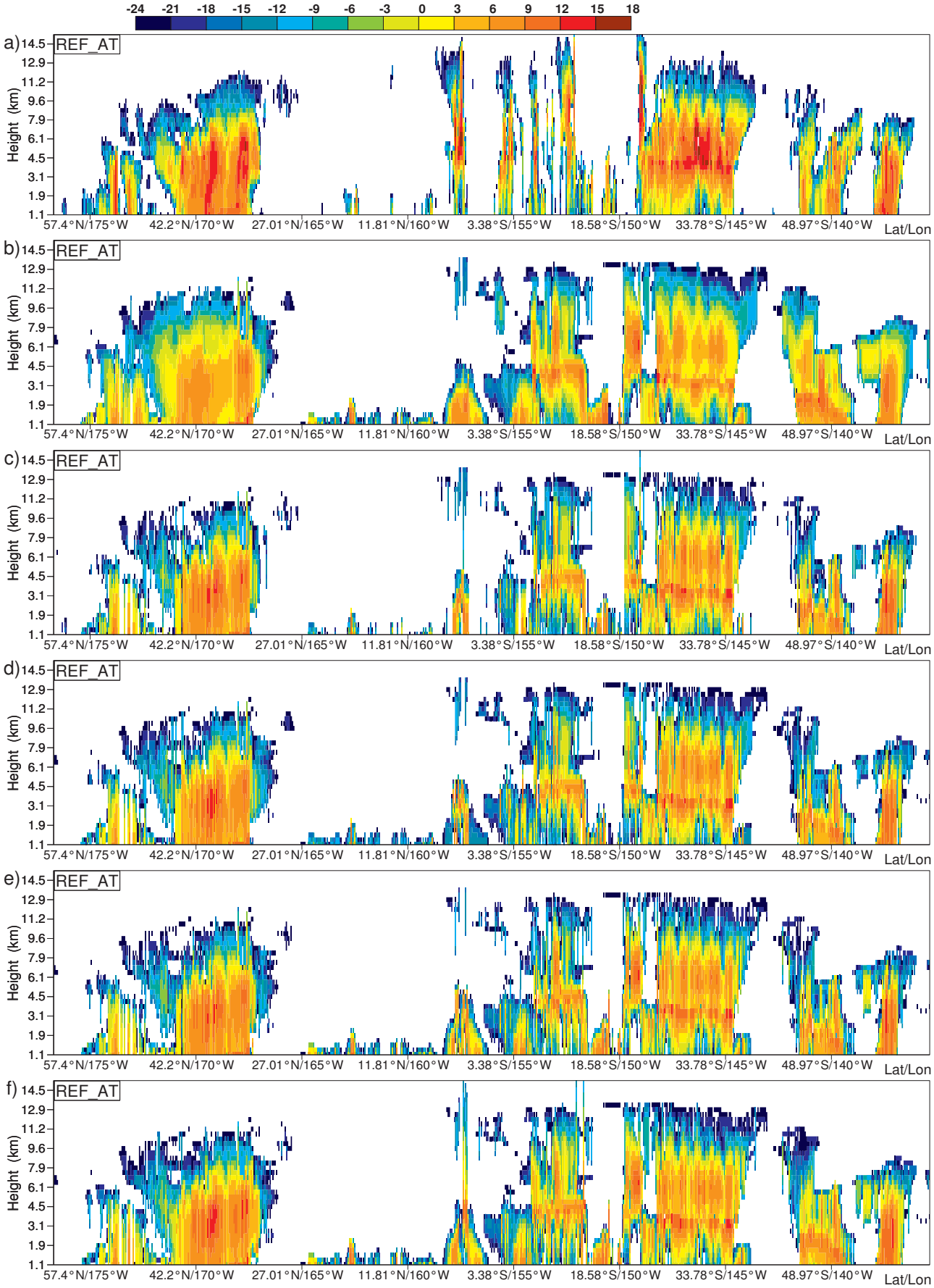

*Figure 3.2: Cloud radar reflectivity (dBZ) for the situation on 24 January 2007 over the Pacific Ocean. (a) CloudSat observations from 94 GHz radar, (b) model first guess (FG) and (c-f) 1D-Var retrievals using cloud reflectivity observations with: (c) rough quality control, no bias correction and simple observation error definition, (d) bias correction applied (**bc**), (e) bias correction and quality control (**qcbc**) and (f) complex observation error definition applied to **qcbc**.*

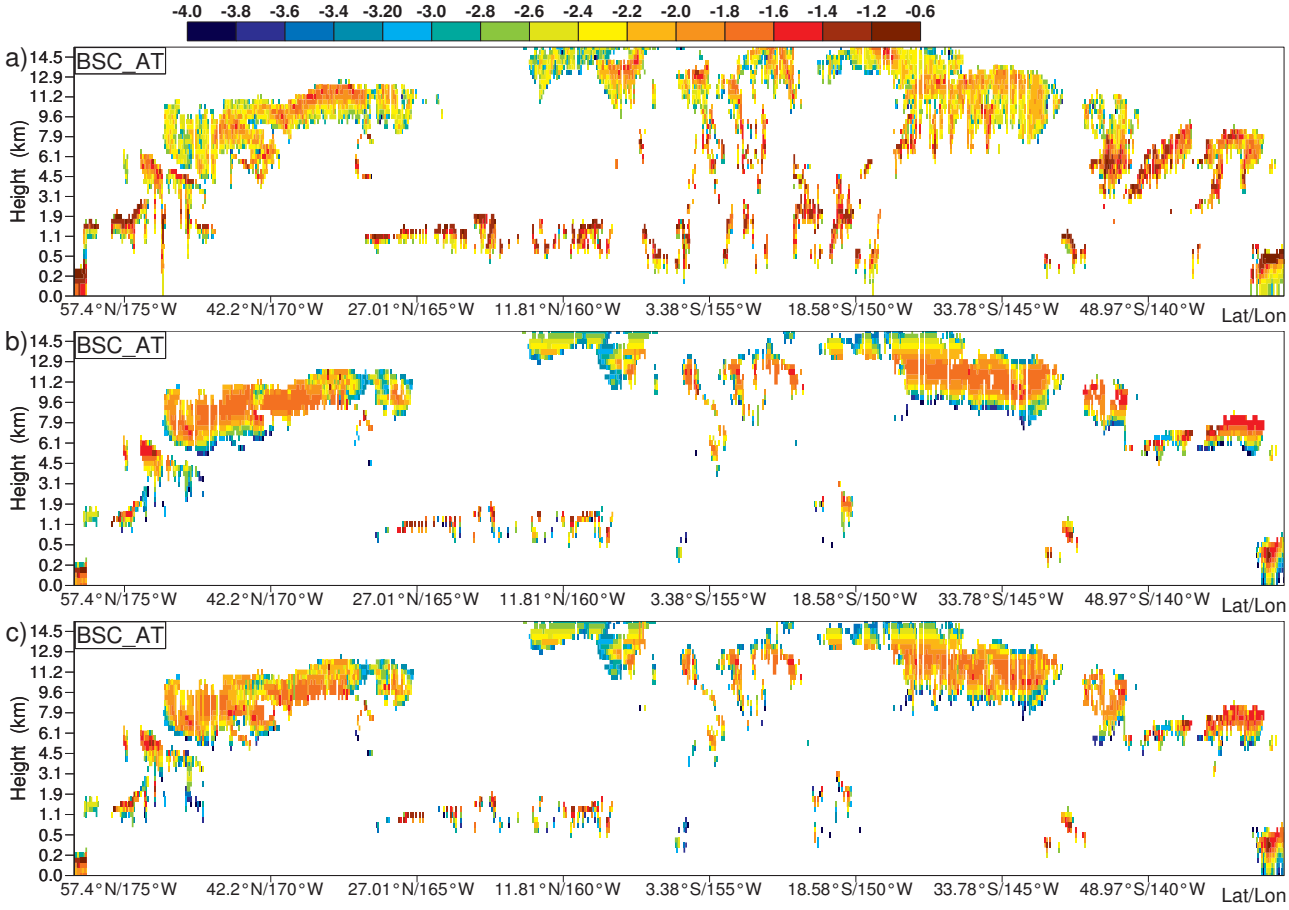

*Figure 3.3: Lidar backscatter due to clouds (in km^{-1} sr^{-1}) for the same situation as on Fig. 3.2: (a) CALIPSO observations, (b) model first guess and (c) analyzed values obtained from the modified profiles of temperature and specific humidity coming from the 1D-Var retrievals using cloud reflectivity observations when applying bias correction, quality control and complex observation error definition (***qcbcer***).*

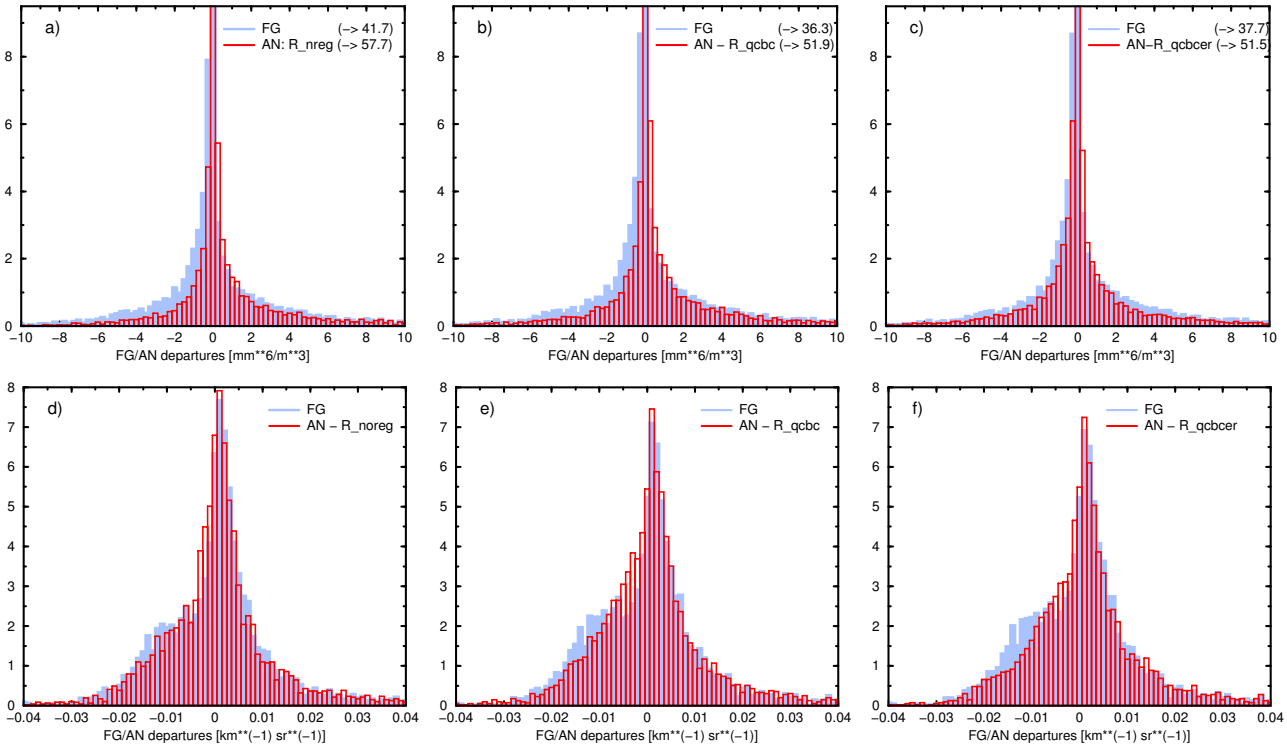

*Figure 3.4: Probability distribution functions of the first-guess (light blue shading) and analysis (red line) departures for (a-c) cloud radar reflectivity (in $mm^6\ m^{-3}$) and (d-f) lidar backscatter due to clouds (in $km^{-1}\ sr^{-1}$) coming from 1D-Var retrievals using CloudSat radar reflectivity observations (a, d) with rough quality control, no bias correction and simple observation error definition (***nreg***), (b, e) applying bias correction and quality control (***qcbc***) on top of* ***nreg*** *and (c, f) with complex observation error definition (***qcbcer***) applied on top of* ***qcbc****. Situation over the Pacific Ocean on 24 January 2007.*

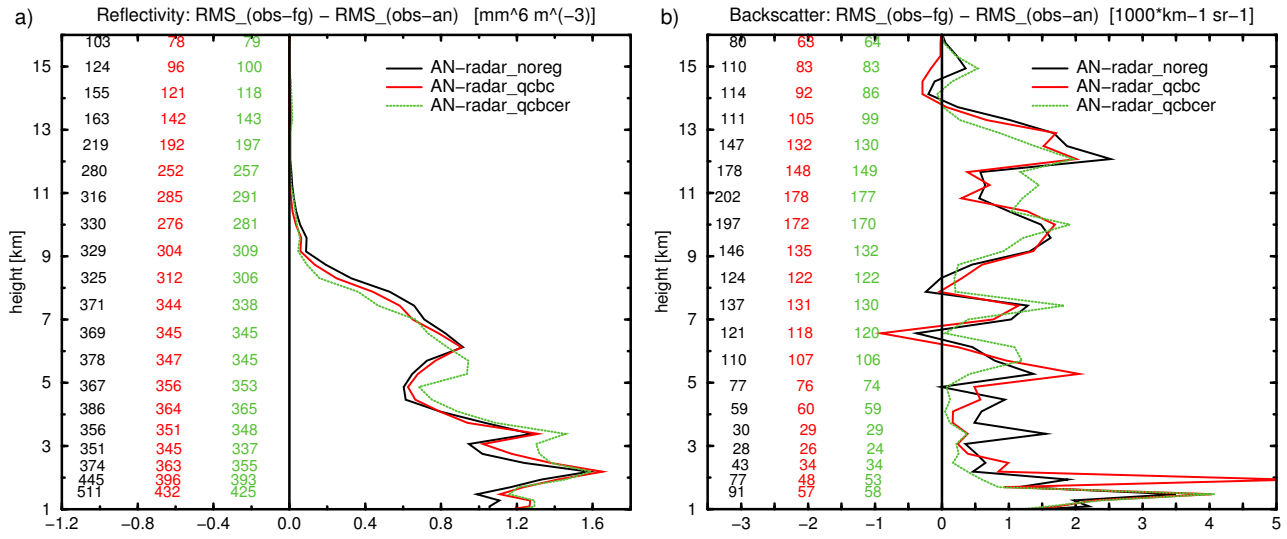

Figure 3.5: Difference of (a) CloudSat radar reflectivity rms errors (in $mm^6\ m^{-3}$) and (b) CALIPSO lidar backscatter rms errors (in 1000 $km^{-1}\ sr^{-1}$) for differences between the first guess (FG) departures and analysis (AN) departures when assimilating observations of cloud radar reflectivity. Results are shown for the same experiments (**nreg**, **qcbc** *and* **qcbcer***) as in Fig. 3.4. Colour number on the left side of (a) and (b) indicates number of observations considered for statistics by the different experiments.*

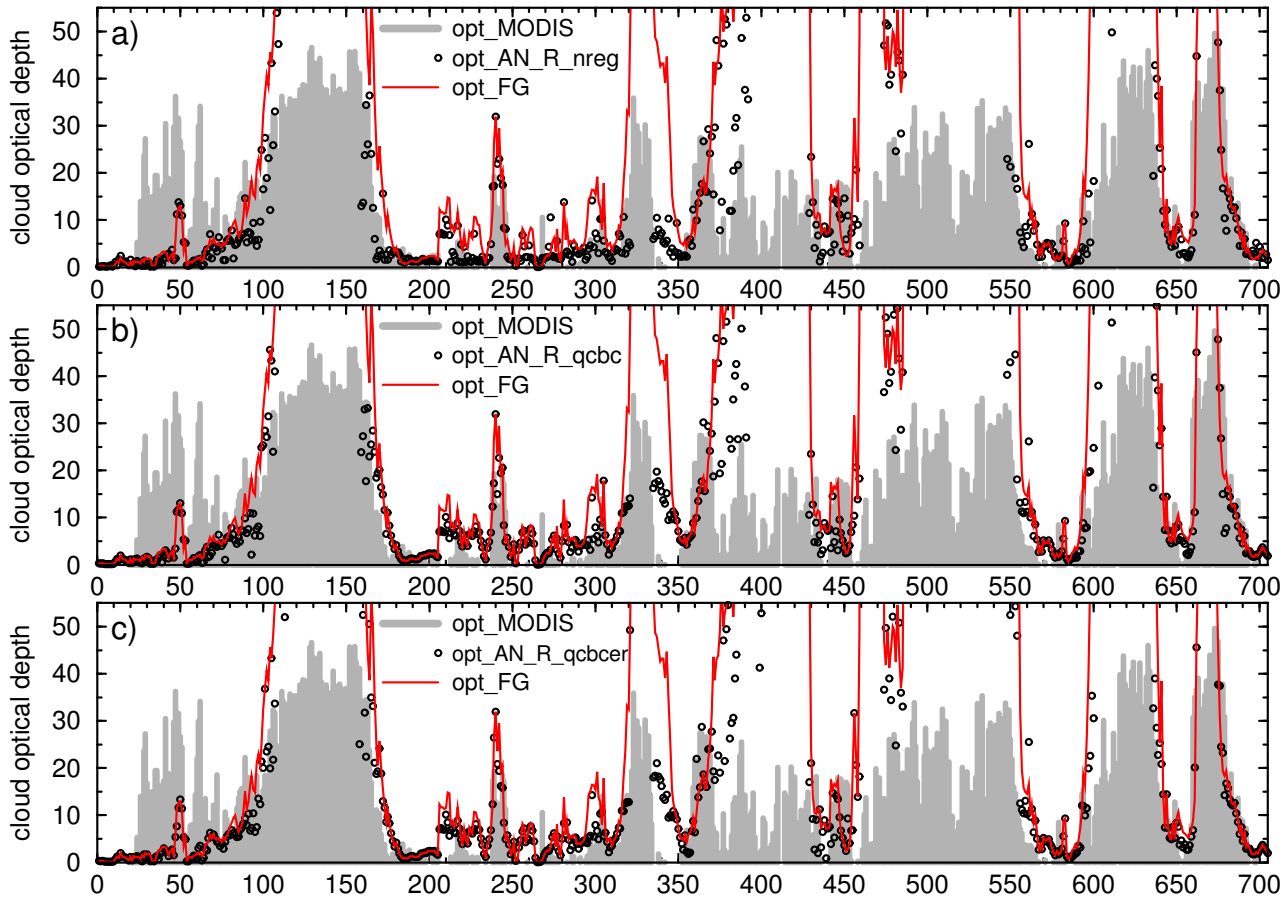

Figure 3.6: Comparison of the model first guess (red line) and the analyzed values of cloud optical depth from 1D-Var retrievals using observations of cloud radar reflectivity (black circles) against MODIS cloud optical depth observations (grey shading). Results are shown for the same experiments as in Fig. 3.4: (a) **nreg***, (b)* **qcbc** *and (c)* **qcbcer** *for the situation over the Pacific Ocean on 24 January 2007. The horizontal axis is in grid points.*

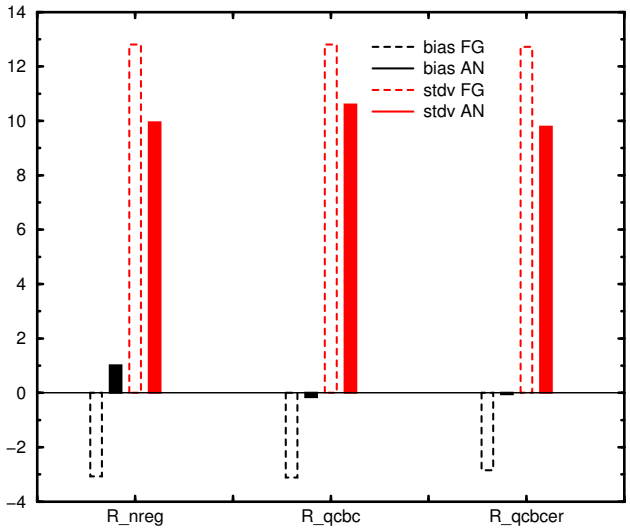

Figure 3.7: Bias (in black) and standard deviation (in red) of the FG (dashed bar) and AN (solid filled bar) departures from MODIS cloud optical depth for the same 1D-Var experiments as in Fig. 3.4. Situation of 24 January 2013.

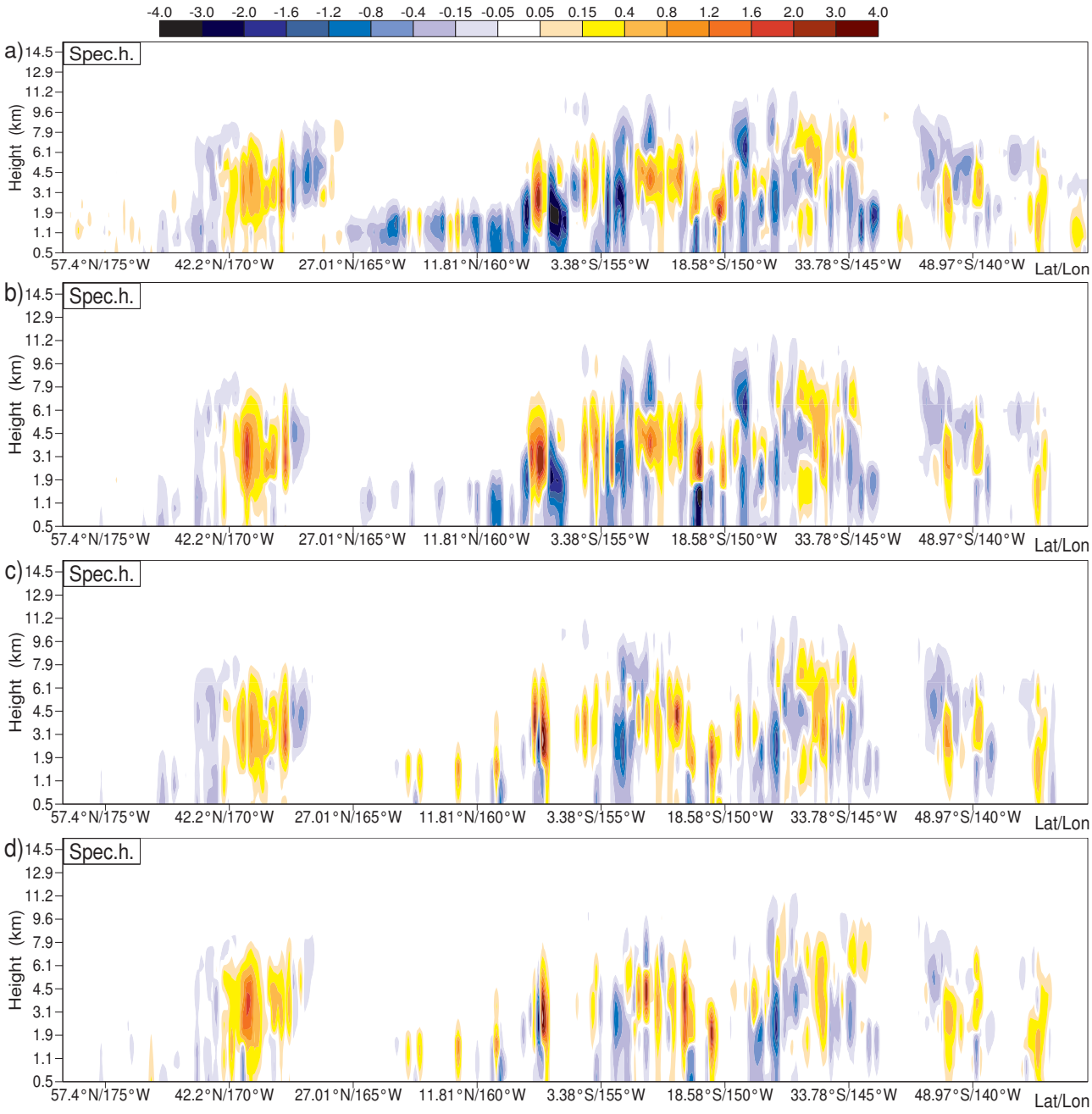

*Figure 3.8: Analysis increments for specific humidity in g kg^{-1} from 1D-Var assimilation of CloudSat cloud radar reflectivity (shown in Fig. 3.2): (a) using rough quality control, no bias correction and simple observation error definition (***nreg***), (b) applying bias correction (***bc***) only on top of* ***nreg*** *or (c) together with the quality control (***qcbc***) and (d) applying complex observation error definition on top of* ***qcbc****. Situation over the Pacific Ocean on 24 January 2007.*

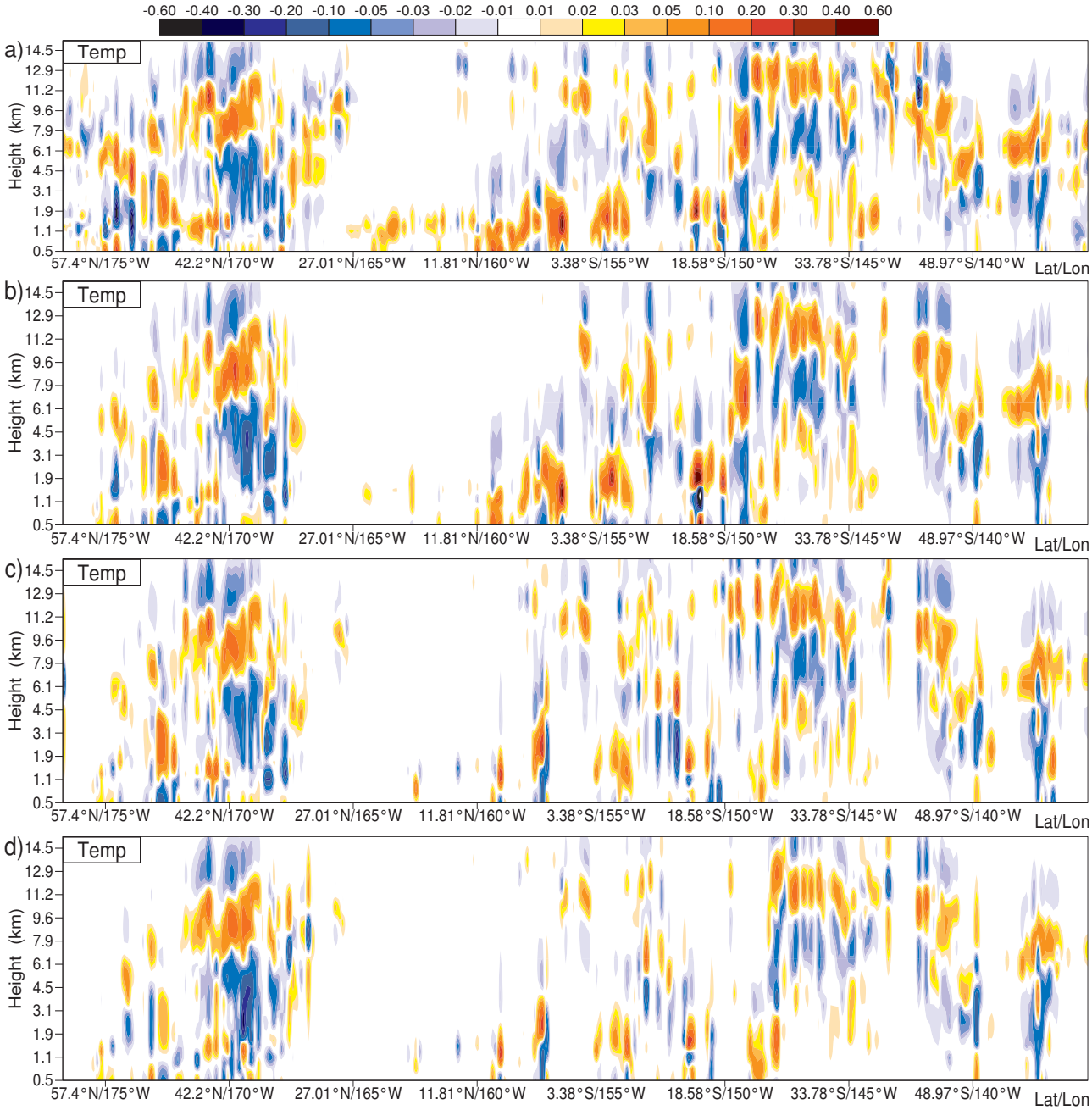

Figure 3.9: Same as Fig. 3.8, but for temperature in K.

4 1D-Var experiments for CALIPSO lidar backscatter due to clouds

4.1 Experimental setup

1D-Var experiments for the situation over the Pacific Ocean on 23-24 January 2007 have also been run using lidar observations due to clouds from CALIPSO lidar at 532 nm. The different observation treatments as described in Section 2.5 have been used. The CALIPSO observations averaged to the model grid at T799 horizontal resolution have been matched with 676 model profiles in this case.

4.2 Results

A comparison of the first-guess and 1D-Var retrieved lidar backscatter against CALIPSO cloud lidar observations at 532 nm over the Pacific Ocean is shown in Fig. 4.1. For easier comparison, the model values (either first-guess or analyzed) are only displayed when the CALIPSO observations are available. Results indicate that contrary to the radar case the analysis fit to observations is only marginally better than for the first guess when applying bias correction and quality control together with the complex observation error definition (i.e. using sum of instrument, forward modelling and representativity errors). This is partly related to the small observation field of view, which leads to large values of representativity error. These larger errors then reduce the weight given to these observations in the analysis. However, only when using such observation errors (experiments **qcbcer**), analysis gets also closer to the independent CloudSat radar reflectivity observations. This can be seen in Table 4.1 and especially in Table 4.2 where rms errors of analysis departures are only reduced with respect to the first-guess ones for the **qcbcer** experiment.

Figure 4.2 shows that the analyzed values of cloud radar reflectivity when assimilating the cloud lidar backscatter get closer to the observations mainly at the top of clouds, for instance in the areas approximately around 172°-174°W, 168°-170°W or 140°-150°W.

The PDFs of analysis departures become narrower compared to the FG departures (Fig. 4.3) for the lidar backscatter with the smallest adjustment in the case of **qcbcer** experiment. Though, as seen in Fig. 4.4, only for this experiment, rms errors of analysis departures are reduced with respect to the first guess ones from independent cloud radar reflectivity for most of the vertical profiles, while the other two experiments (**nreg** and **qcbc**) lead to increased rms errors of the analysis departures. This underlines how important is to apply an appropriate quality control, bias correction and error estimate to ensure the "healthy" 1D-Var performance.

Figures 4.5 and 4.6 show that there is only marginal improvement of the analysis departures compared to the first-guess ones for MODIS cloud optical depth as already indicated by other results presented in Fig. 4.1 - 4.4.

Analysis increments of specific humidity (Fig. 4.7) and temperature (Fig. 4.8) occur at higher altitudes than in the case of assimilation of the cloud radar reflectivity (Fig. 3.8-3.9). By applying bias correction and quality control (**qcbc**), increments are reduced at altitudes and locations where the first-guess departures from observations are large and therefore not used in the 1D-Var analysis. Further reduction appears in the **qcbcer** experiment at lower altitudes because of the large observation error due to the representativity issues for the scenes with the hight variability of cloud cover, such as in the tropics (25°N-15°S) or above 45°N.

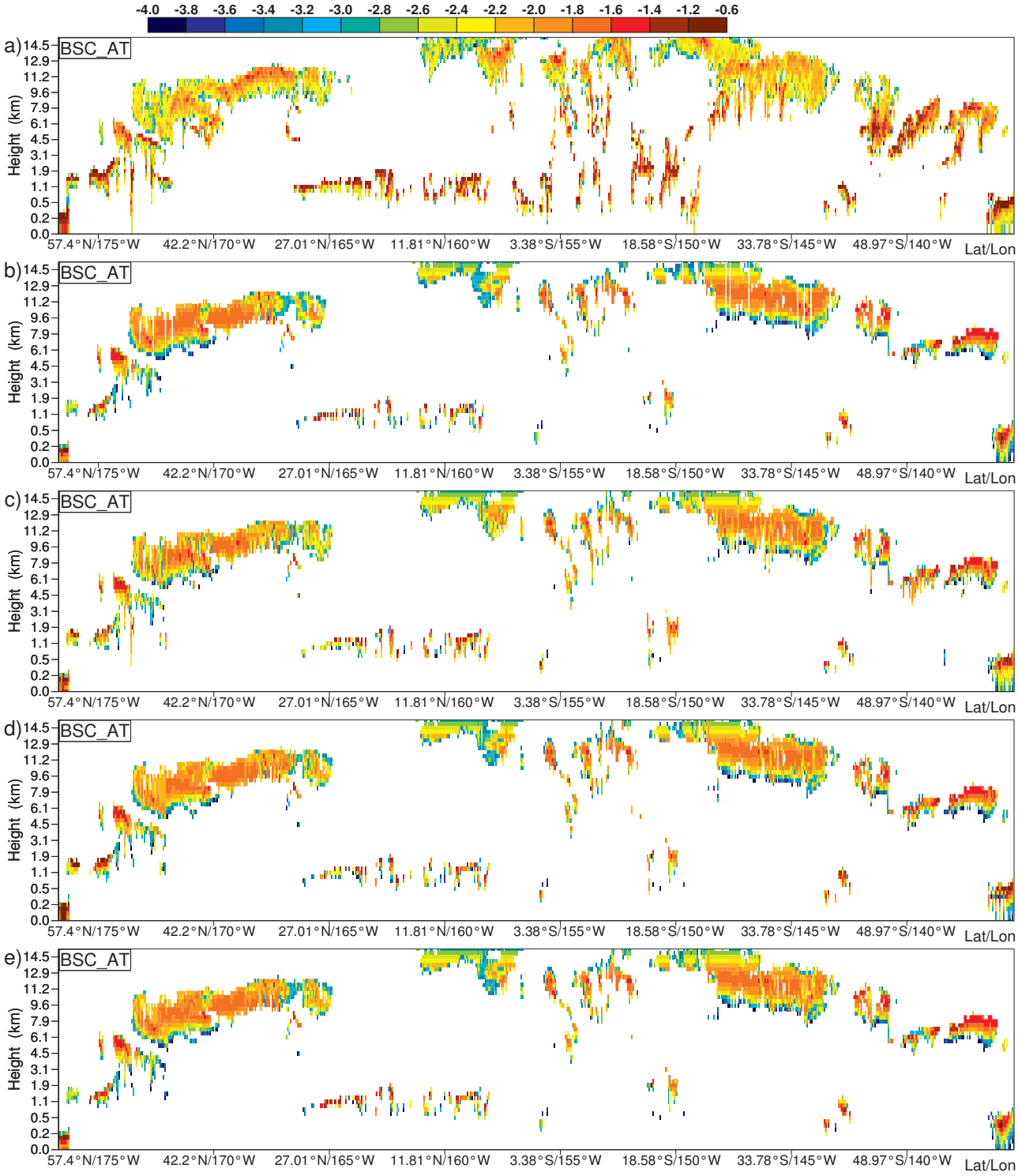

*Figure 4.1: Lidar backscatter due to clouds (in km^{-1} sr^{-1}) for the situation on 24 January 2007 over the Pacific Ocean. (a) CALIPSO observations, (b) model first guess and (c-e) 1D-Var retrievals using cloud lidar backscatter observations with: (c) rough quality control, no bias correction and simple observation error definition (***nreg***), (d) bias correction and quality control (***qcbc***) applied to **nreg** and (e) complex observation error definition applied to **qcbc**.*

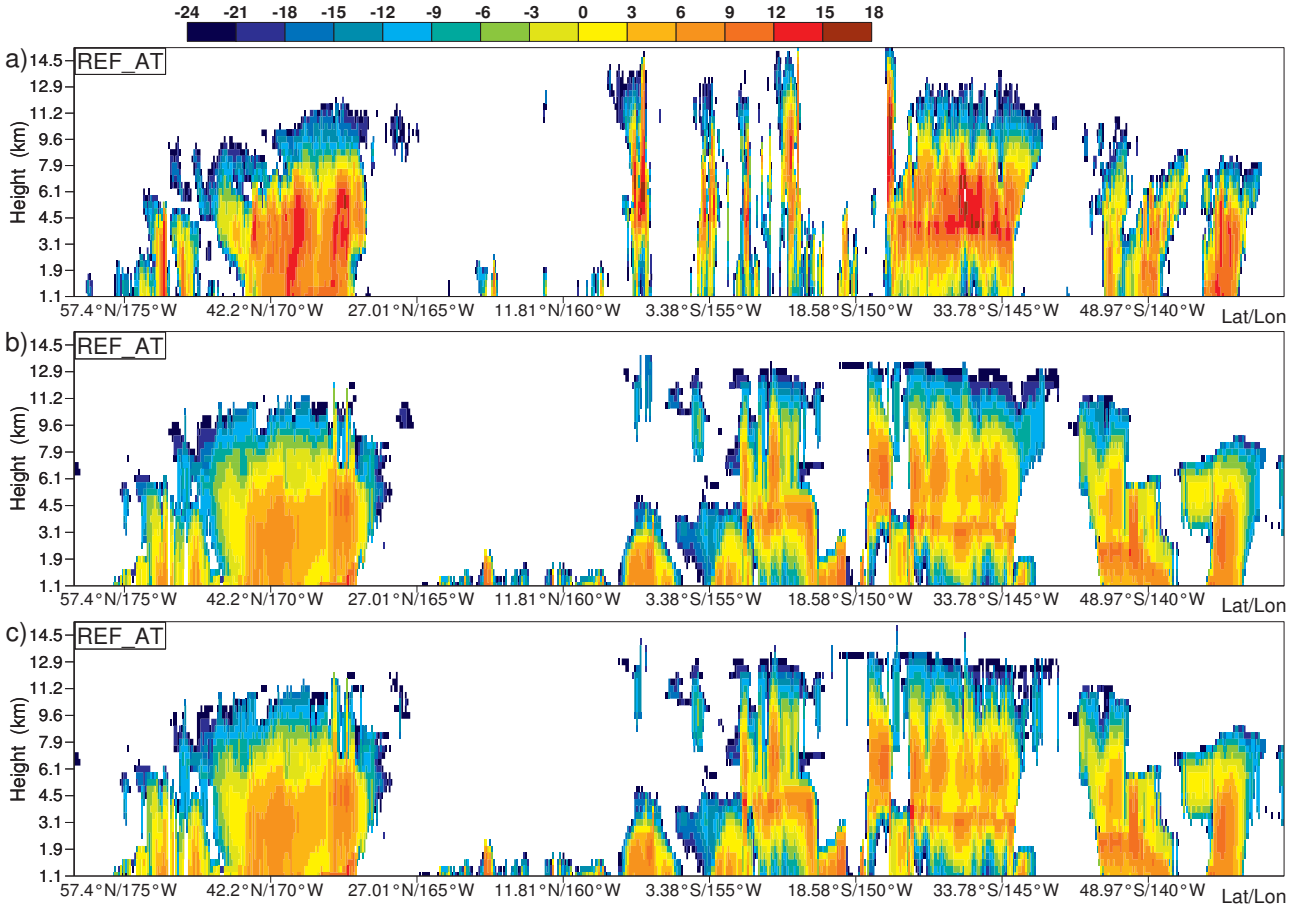

*Figure 4.2: Cloud radar reflectivity (in dBZ) for the same situation as on Fig. 4.1: (a) CloudSat observations, (b) model first guess and (c) analyzed values obtained from the modified profiles of temperature and specific humidity coming from the 1D-Var retrievals using cloud lidar backscatter observations when applying bias correction, quality control and complex observation error definition (***qcbcer***).*

	backscatter				reflectivity			
	bias	stdv	mae	rms	bias	stdv	mae	rms
FG	7.31E-03	3.16E-02	1.41E-02	3.25E-02	0.57	4.31	1.88	4.35
AN-L_nreg	7.30E-03	2.96E-02	1.19E-02	3.05E-02	0.58	4.35	1.90	4.40
AN-L_qcbc	7.12E-03	2.94E-02	1.40E-02	3.02E-02	0.57	4.34	1.90	4.39
AN-L_qcbcer	7.28E-03	3.14E-02	1.39E-02	3.19E-02	0.56	4.30	1.87	4.33

Table 4.1: Bias, standard deviation (stdv), mean absolute error (mae) and root mean square error (rms) of the first guess (FG) and analysis (AN) departures for the different assimilation experiments (see text for experiment description) from the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) and the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) observations. 676 profiles are included in the statistics for the case of 24 January 2007.

	backscatter		reflectivity	
	DIFF	RMSD	DIFF	RMSD
AN-L_nreg	-3.372E-03	-2.734E-03	3.783E-02	5.183E-02
AN-L_qcbc	-1.977E-03	-1.561E-03	1.060E-02	1.352E-02
AN-L_qcbcer	-7.108E-04	-7.124E-04	-3.854E-03	-1.733E-05

Table 4.2: Difference between the absolute value of analysis (AN) minus observation (OBS) and absolute value of the first guess (FG) minus OBS (DIFF), as well as difference of rms errors for the FG and AN departures (RMSD) from the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) and the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) observations. Results are shown for the different assimilation experiments (see text for experiment description). 676 profiles are included in the statistics for the case of 24 January 2007.

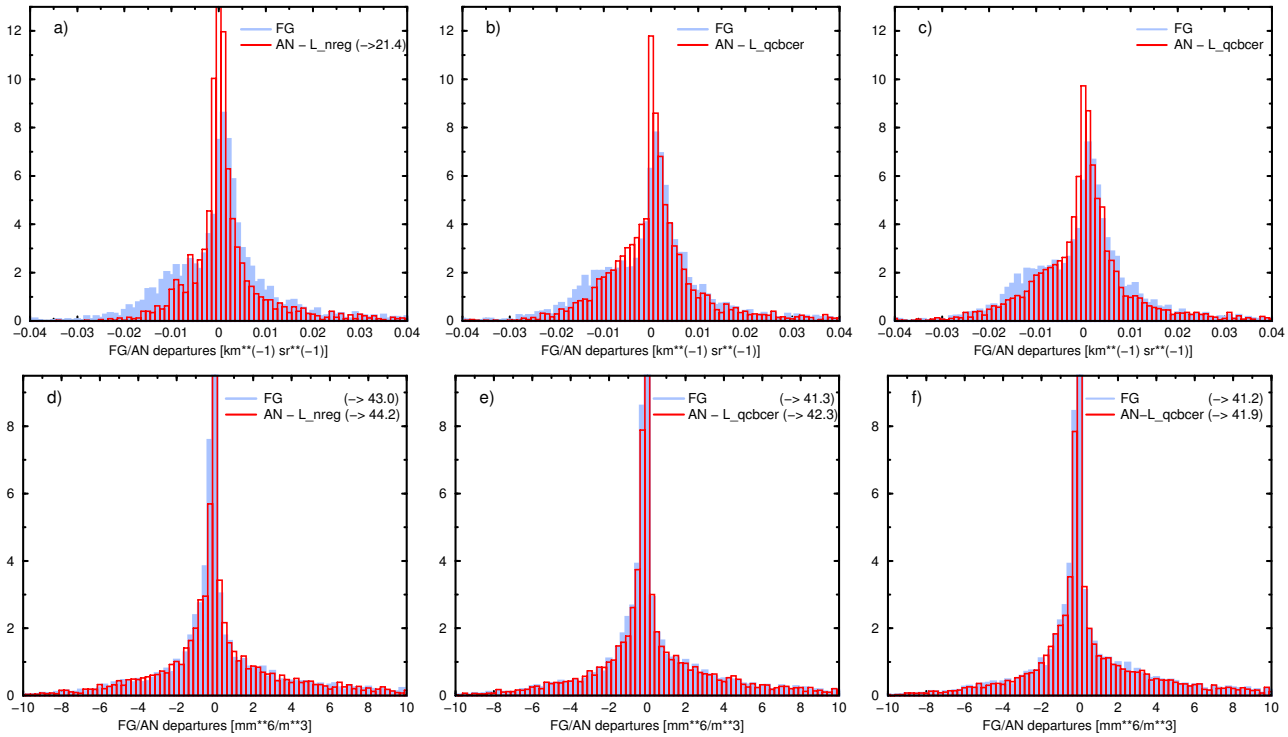

*Figure 4.3: Probability distribution functions of the first-guess (light blue shading) and analysis (red line) departures for (a-c) lidar backscatter due to clouds (in km^{-1} sr^{-1}) and (d-f) cloud radar reflectivity (in mm^6 m^{-3}) coming from 1D-Var retrievals using CALIPSO backscatter observations (a, d) with rough quality control, no bias correction and simple observation error definition (**nreg**), (b, e) applying bias correction and quality control (**qcbc**) on top of **nreg** and (c, f) with complex observation error definition (**qcbcer**) applied on top of **qcbc**. Situation over the Pacific Ocean on 24 January 2007.*

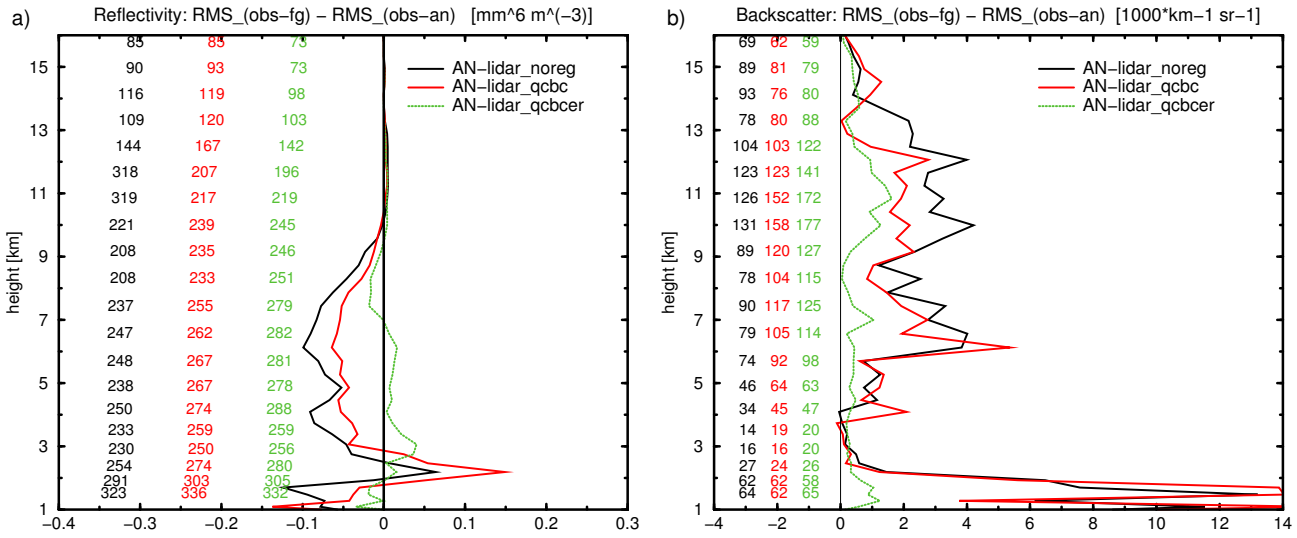

*Figure 4.4: Difference of (a) CloudSat radar reflectivity rms errors (in mm^6 m^{-3}) and (b) CALIPSO lidar backscatter rms errors (in 1000 km^{-1} sr^{-1}) for differences between the first guess (FG) departures and analysis (AN) departures when assimilating observations of lidar backscatter due to clouds. Results are shown for the same experiments (**nreg**, **qcbc** and **qcbcer**) as in Fig. 4.3. Colour numbers on the left side of (a) and (b) indicate number of observations considered for statistics by the different experiments.*

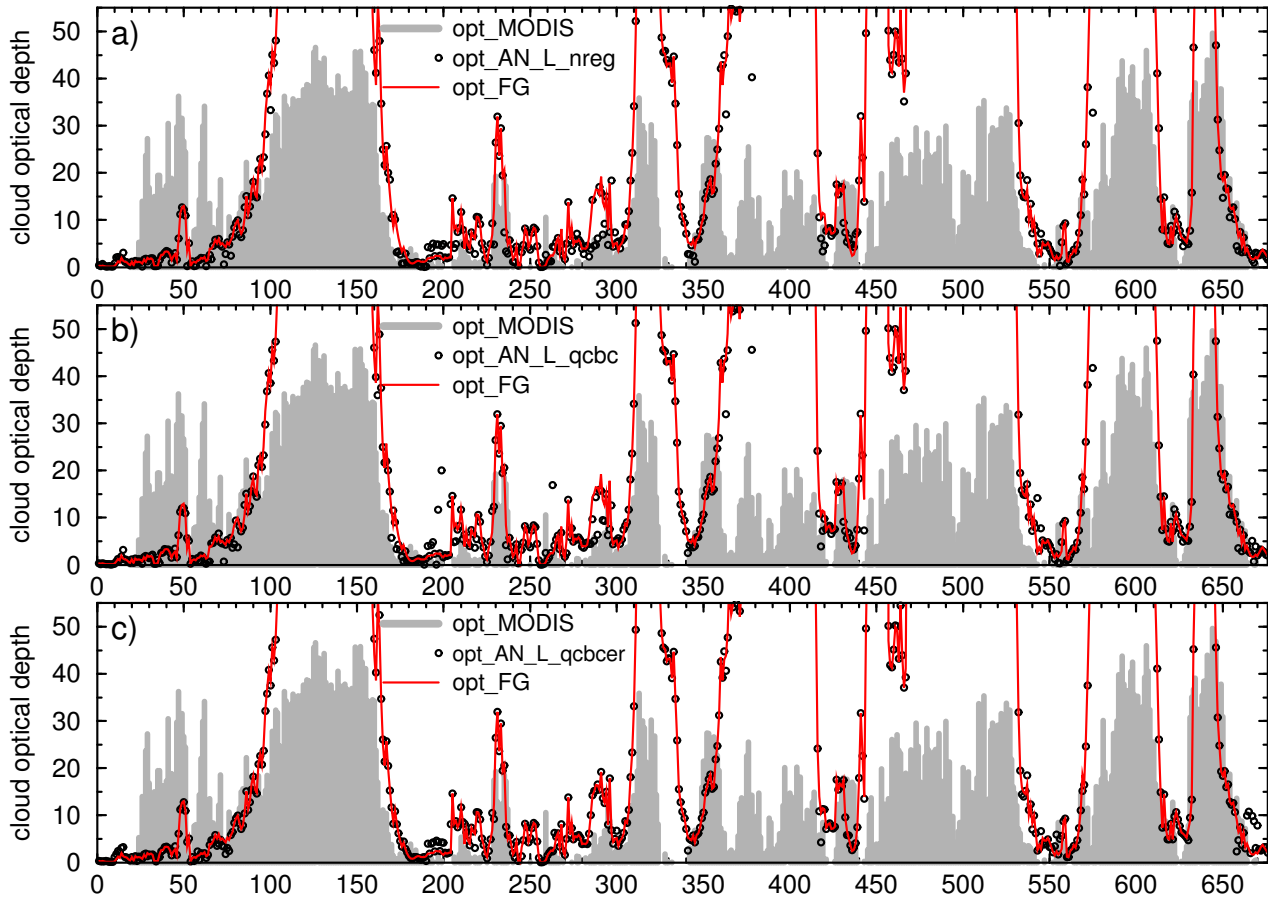

Figure 4.5: Comparison of the model first guess (red line) and the analyzed values of cloud optical depth from 1D-Var retrievals using observations of lidar backscatter due to clouds (black circles) against MODIS cloud optical depth observations (grey shading). Results are shown for the same experiments as in Fig. 4.3: (a) **nreg***, (b)* **qcbc** *and (c)* **qcbcer** *for the situation over the Pacific Ocean on 24 January 2007. The horizontal axis is in grid points.*

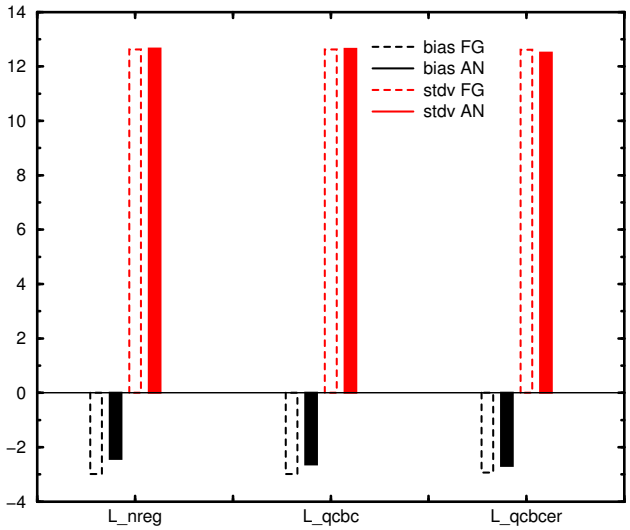

Figure 4.6: Bias (in black) and standard deviation (in red) of the FG (dashed bar) and AN (solid filled bar) departures from MODIS cloud optical depth for the same 1D-Var experiments as in Fig. 4.3. Situation of 24 January 2013.

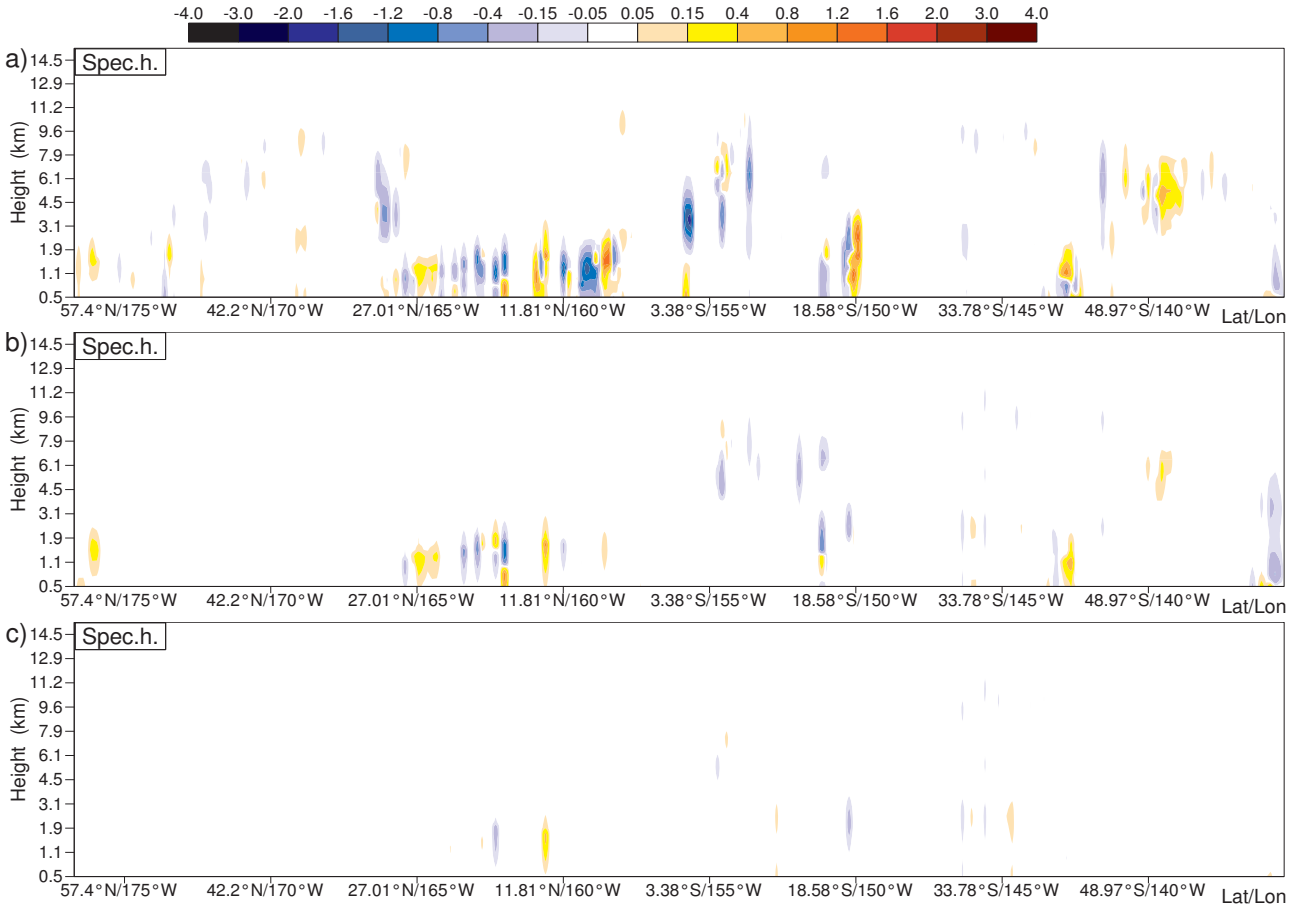

*Figure 4.7: Analysis increments for specific humidity in g kg^{-1} from 1D-Var assimilation of CALIPSO backscatter due to clouds (shown in Fig. 4.1): (a) using rough quality control, no bias correction and simple observation error definition (***nreg***), (b) applying bias correction together with the quality control (***qcbc***) to* ***nreg*** *and (c) applying complex observation error definition on top of* ***qcbc****. Situation over the Pacific Ocean on 24 January 2007.*

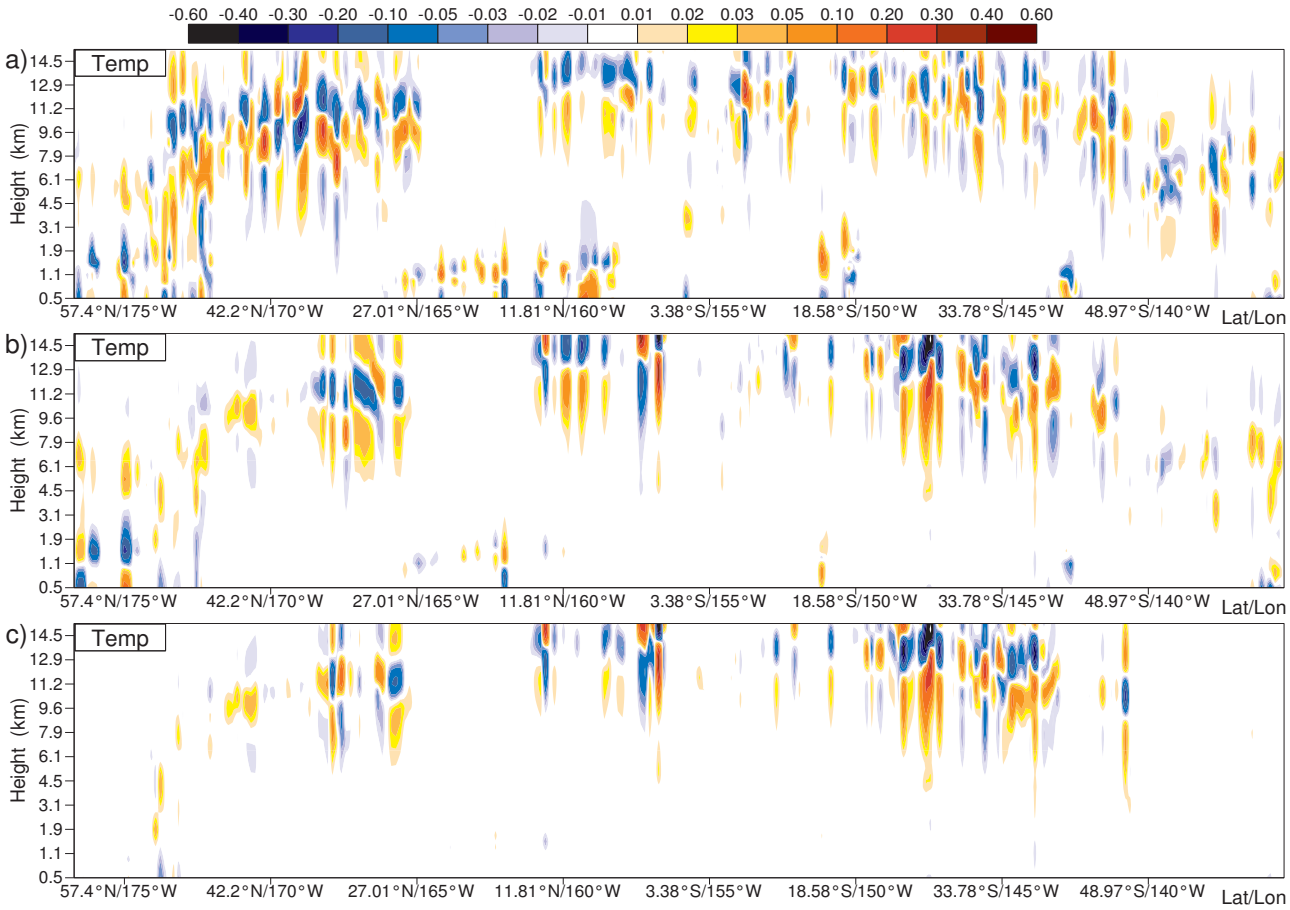

Figure 4.8: Same as Fig. 4.7, but for temperature in K.

5 1D-Var experiments for combined radar and lidar observations

5.1 Experimental setup

Assimilation of the combined cloud radar reflectivity and lidar backscatter observations in the 1D-Var system for the same situation as in Sections 3 and 4 has been done for the different observation treatments describe in Section 2.5. In this case, 707 model profiles at T799 horizontal resolution have been matched with the observations averaged to the model grid for the situation over the Pacific Ocean on 24 January 2007. Due to a slight delay of the CALIPSO track, there are some points which are not covered by both observations at the same time step. However, for most of the points, both observations are available simultaneously.

5.2 Results

Figure 5.1 and 5.2 show a comparison of the first-guess and 1D-Var retrieved radar reflectivity and lidar backscatter, respectively, against the corresponding observations over the Pacific Ocean. For cloud radar reflectivity, similar features to the reflectivity only 1D-Var experiments (Section 3) are also observed in the case of combined observations, e.g removing or reducing clouds more than indicated by the observations when the 1D-Var is run without a stricter control of observations and a proper observation error estimate (**nreg**). This is obvious above 4.5 km in the area between 167° and 173°W or above 9 km between 144° and 148°W. The applied quality control and bias correction reduces adjustment of analysis when the differences between the first guess and observations are too large or they are a consequence of the model bias. The experiment **qcbcer** with the complex error definition reduces further adjustment in the areas with representativity issues between observations and the model, e.g. in the tropical area (the most significant between 156° and 160°W). On other hand, there are cases when the analysis is closer to the observations for **qcbc** or **qcbcer** experiments than for **nreg** experiment, such as increased reflectivity below 4 km around 173°W, between 155° and 156°W or around 138°W which corresponds better to the observations. The differences between the analyzed values of cloud lidar backscatter (Fig. 5.2) for the different 1D-Var experiments are smaller than for the reflectivity.

From the results of 1D-Var assimilation experiments summarized in Tables 5.1 and 5.2 it is not straightforward to conclude which experiment (i.e. treatment of observations) provides the analysis closest to both types of the assimilated observations. Some of them are getting closer to the reflectivity observations and some of them to the backscatter observations. This is also obvious from the vertical profiles of differences between rms errors of the FG departures and rms errors of the AN departures (Fig. 5.4) from both CloudSat radar reflectivity and CALIPSO lidar backscatter.

PDFs of the analysis departures (Fig. 5.3) are narrower than for the first-guess departures in all experiments, though PDFs of the analyses departures for **qcbcer** experiment are slightly more symmetric for both observation types.

Only a comparison against the independent MODIS observation of cloud optical depth (Fig. 5.5-5.6) indicates that **qcbcer** experiment performs better than others. Smaller differences between experiments with the different treatment of observations when assimilating the combined radar and lidar observations rather than a single observation type may be due to the fact that using two observation types on its own already provides some level of control to the 1D-Var performance since analysis must find compromise between adjustments to the both observation types.

An evaluation of analysis increments of specific humidity (Fig. 5.7) and temperature (Fig. 5.8) reveals similar features as for 1D-Var experiments of cloud radar reflectivity and lidar backscatter separately, i.e. reduced increments for **qcbcer** experiment with respect to **qcbc** and especially to **nreg** experiments. Close comparison indicates that the analysis increments from the assimilation of combined observations actually result from

the superimposition of increments coming from adjustment to the radar reflectivity and the lidar backscatter observations.

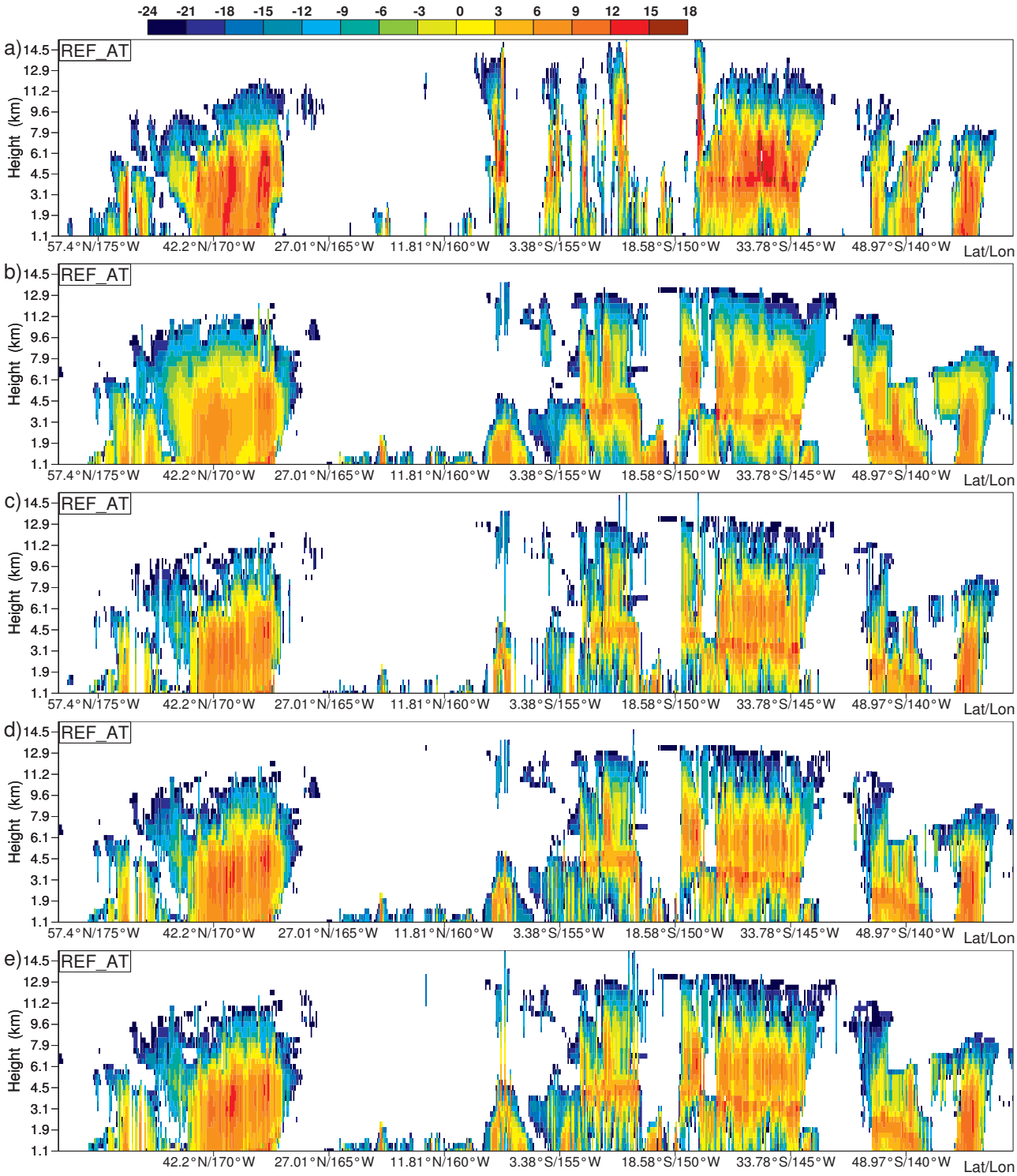

*Figure 5.1: Cloud radar reflectivity (dBZ) for the situation on 24 January 2007 over the Pacific Ocean. (a) CloudSat observations, (b) model first guess (FG) and (c-e) 1D-Var retrievals using both cloud radar reflectivity and lidar backscatter observations with: (c) rough quality control, no bias correction and simple observation error definition (***nreg***), (d) bias correction and quality control (**qcbc**) applied to **nreg** and (e) complex observation error definition applied to **qcbc**.*

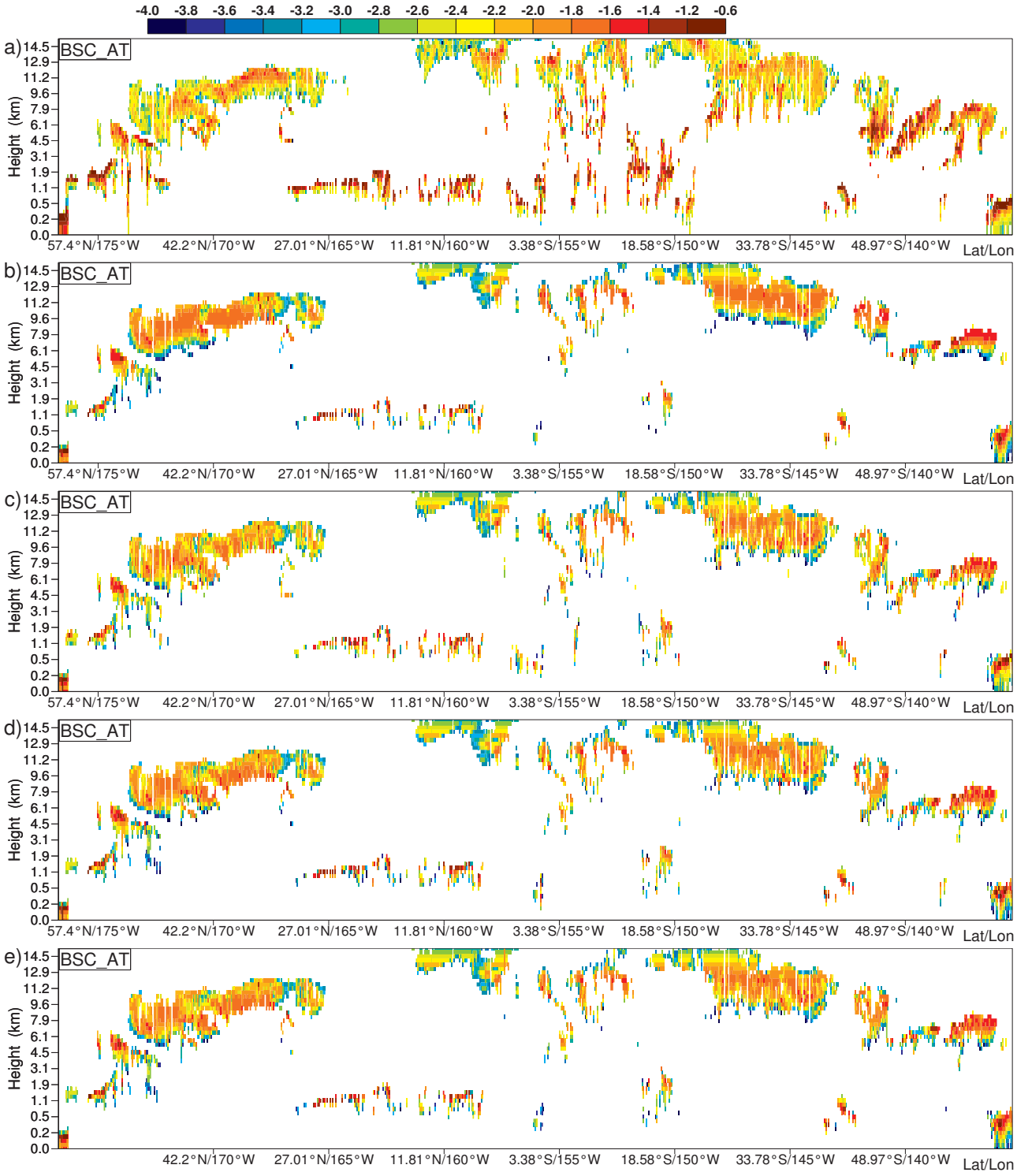

Figure 5.2: Same as Fig. 5.1, but for lidar backscatter due to clouds (in km^{-1} sr^{-1}).

	reflectivity				backscatter			
	bias	stdv	mae	rms	bias	stdv	mae	rms
FG	0.55	4.26	1.86	4.30	7.35E-03	3.17E-02	1.41E-02	3.25E-02
AN-C_nreg	0.78	3.74	1.39	3.82	6.95E-03	2.90E-02	1.17E-02	2.99E-02
AN-C_qcbc	0.59	3.76	1.40	3.81	7.24E-03	3.09E-02	1.35E-02	3.20E-02
AN-C_qcbcer	0.53	3.70	1.42	3.80	7.29E-03	3.12E-02	1.37E-02	3.21E-02

Table 5.1: Bias, standard deviation (stdv), mean absolute error (mae) and root mean square error (rms) of the first guess (FG) and analysis (AN) departures for the different assimilation experiments (see text for experiment description) from the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) and the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) observations. 707 profiles are included in the statistics for the case of 24 January 2007.

	reflectivity		backscatter	
	DIFF	RMSD	DIFF	RMSD
AN-C_nreg	-0.613	-0.664	-2.450E-04	-2.176E-04
AN-C_qcbc	-0.619	-0.692	-8.263E-04	-5.385E-04
AN-C_qcbcer	-0.596	-0.689	-8.583E-04	-6.012E-04

Table 5.2: Difference between the absolute value of analysis (AN) minus observation (OBS) and absolute value of the first guess (FG) minus OBS (DIFF), as well as difference of rms errors for the FG and AN departures (RMSD) from the CloudSat cloud radar reflectivity (in $mm^6\ m^{-3}$) and the CALIPSO lidar cloud backscatter (in $km^{-1}\ sr^{-1}$) observations. Results are shown for the different assimilation experiments (see text for experiment description). 707 profiles are included in the statistics for the case of 24 January 2007.

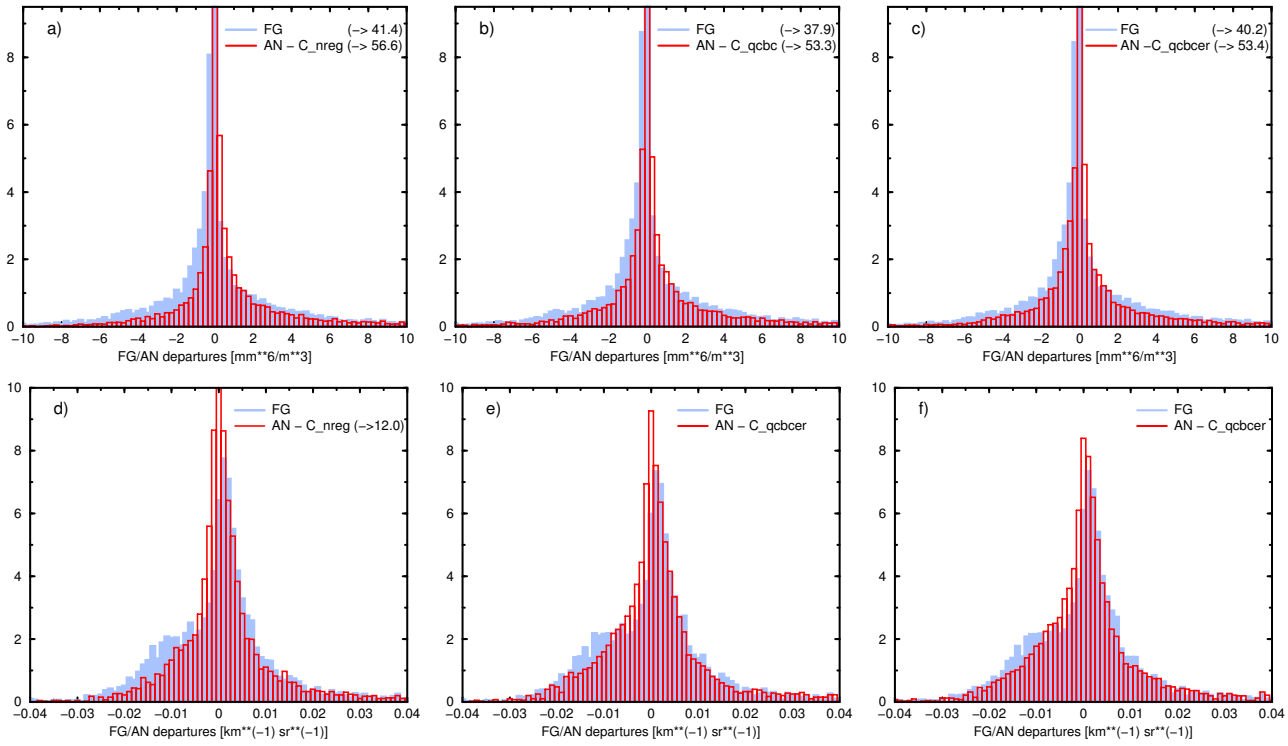

*Figure 5.3: Probability distribution functions of the first-guess (light blue shading) and analysis (red line) departures for (a-c) cloud radar reflectivity (in $mm^{-6}\ m^{-3}$) and (d-f) lidar backscatter due to clouds (in $km^{-1}\ sr^{-1}$) coming from 1D-Var retrievals using combined CloudSat reflectivity and CALIPSO backscatter observations (a, d) with rough quality control, no bias correction and simple observation error definition (**nreg**), (b, e) applying bias correction and quality control (**qcbc**) on top of **nreg** and (c, f) with complex observation error definition (**qcbcer**) applied on top of **qcbc**. Situation over the Pacific Ocean on 24 January 2007.*

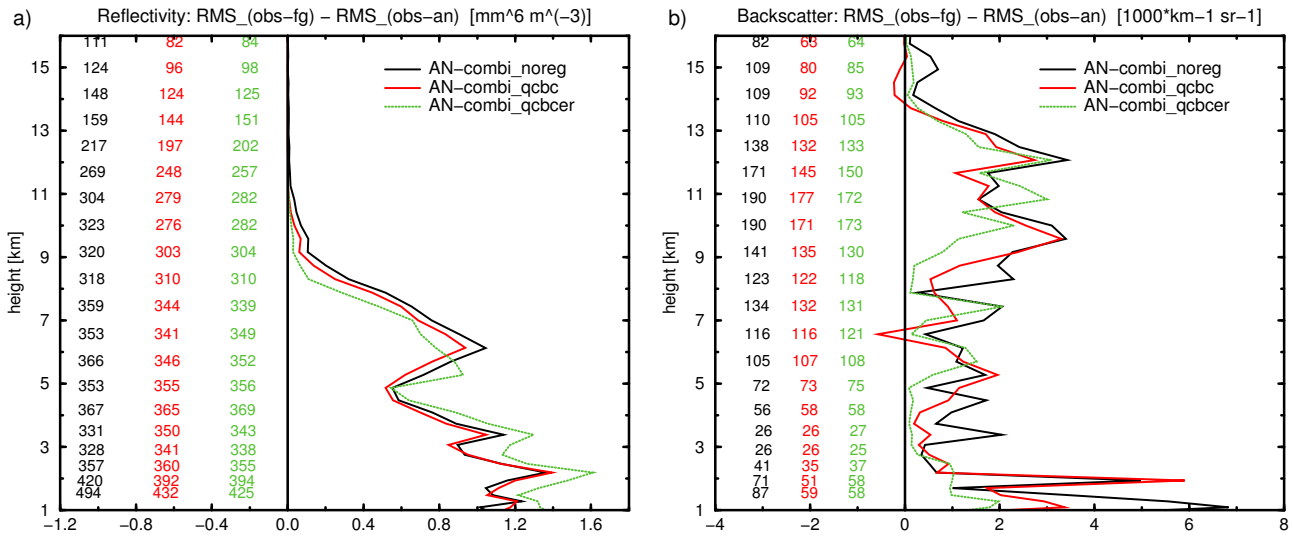

Figure 5.4: Difference of (a) CloudSat radar reflectivity rms errors (in $mm^6\ m^{-3}$) and (b) CALIPSO lidar backscatter rms errors (in 1000 $km^{-1}\ sr^{-1}$) for differences between the first guess (FG) departures and analysis (AN) departures when assimilating combined observations of cloud radar reflectivity and lidar backscatter due to clouds. Results are shown for the same experiments (**nreg**, **qcbc** *and* **qcbcer***) as in Fig. 5.3. Colour numbers on the left side of (a) and (b) indicate number of observations considered for statistics by the different experiments.*

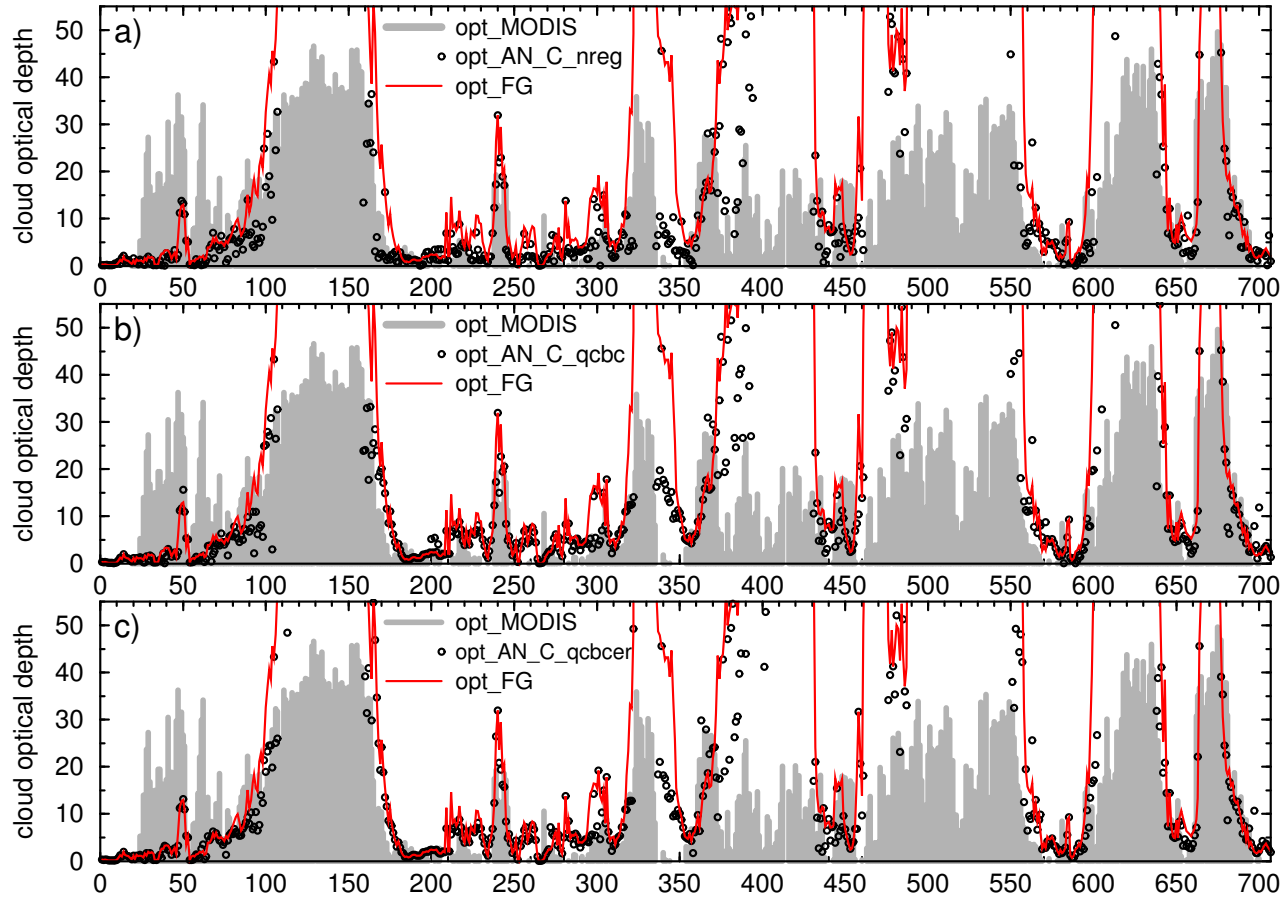

Figure 5.5: Comparison of the model first guess (red line) and the analyzed values of cloud optical depth from 1D-Var retrievals using observations of both cloud radar reflectivity and lidar backscatter due to clouds (black circles) against MODIS cloud optical depth observations (grey shading). Results are shown for the same experiments as in Fig. 5.3: (a) **nreg**, *(b)* **qcbc** *and (c)* **qcbcer** *for the situation over the Pacific Ocean on 24 January 2007. The horizontal axis is in grid points.*

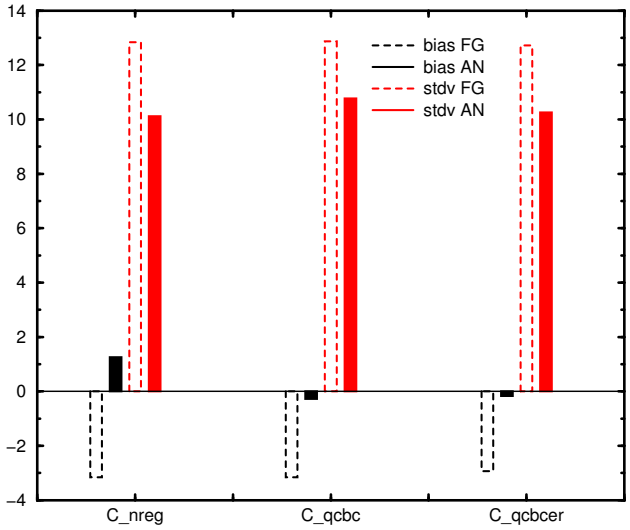

Figure 5.6: Bias (in black) and standard deviation (in red) of the FG (dashed bar) and AN (solid filled bar) departures from MODIS cloud optical depth for the same 1D-Var experiments as in Fig. 5.3. Situation of 24 January 2013.

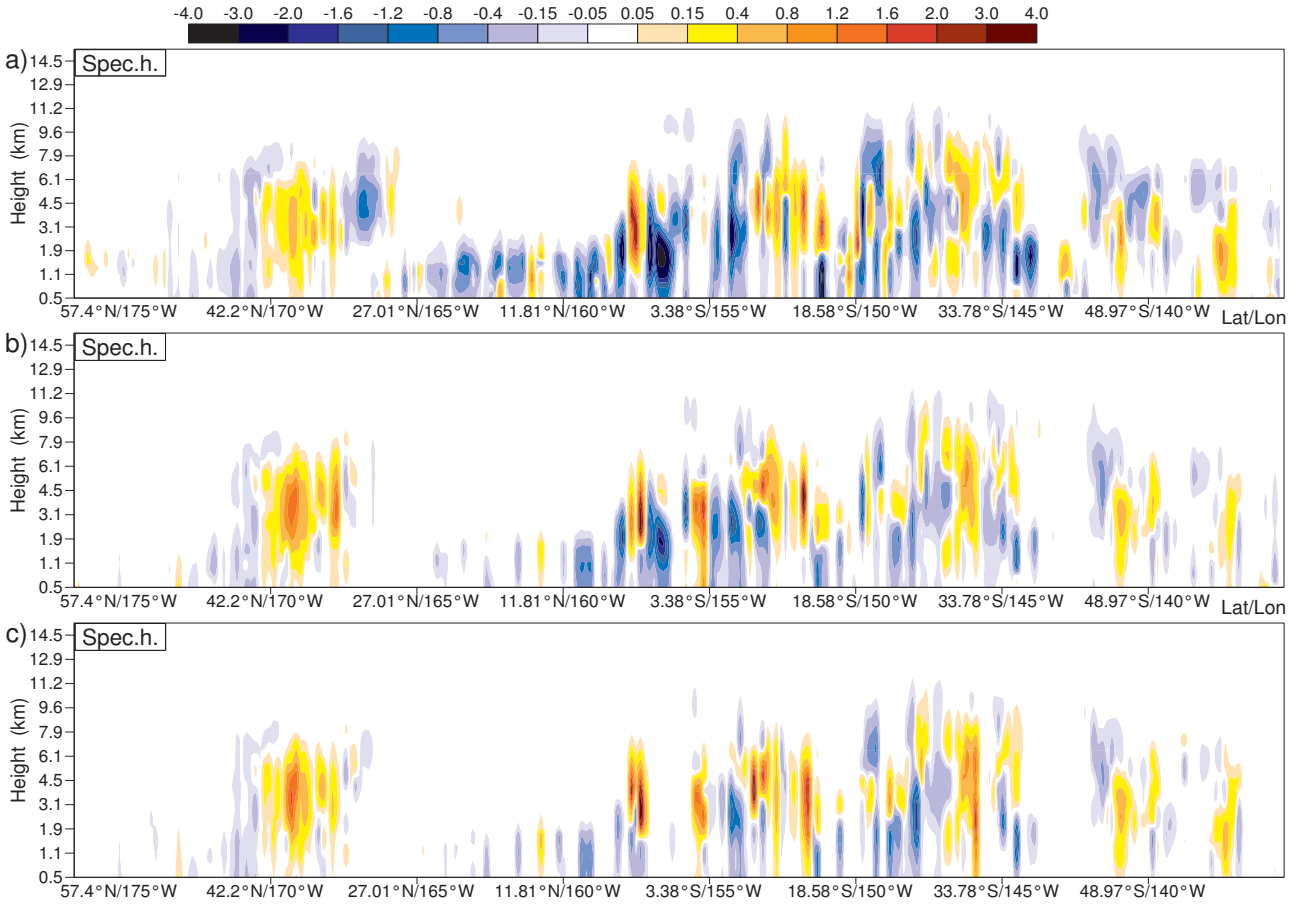

*Figure 5.7: Analysis increments for specific humidity in g kg^{-1} (a, b) from 1D-Var assimilation of combined CloudSat reflectivity and CALIPSO backscatter due to clouds (shown in Fig. 5.1): (a) using rough quality control, no bias correction and simple observation error definition (**nreg**), (b) applying bias correction together with the quality control (**qcbc**) to **nreg** and (c) applying complex observation error definition on top of **qcbc**. Situation over the Pacific Ocean on 24 January 2007.*

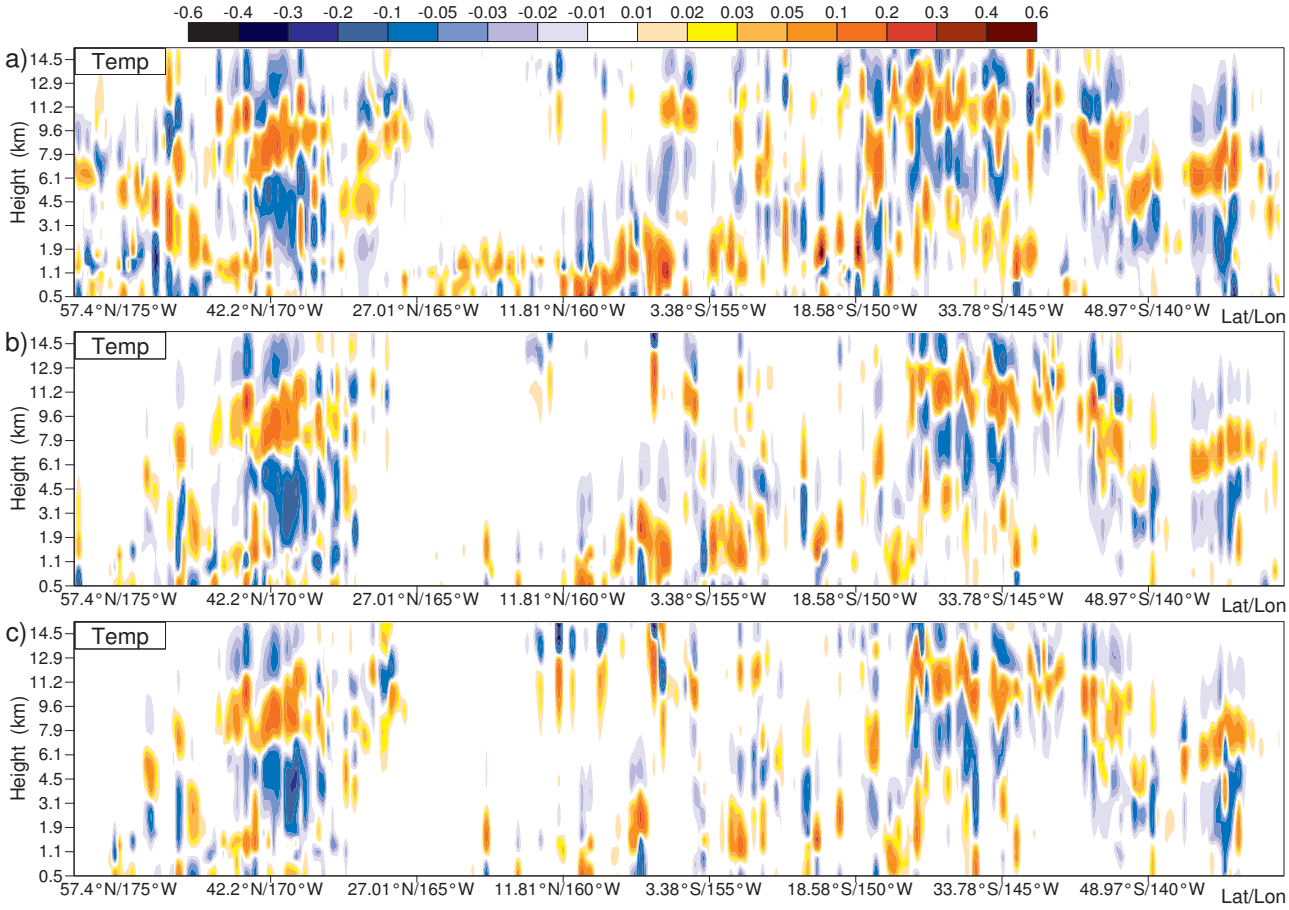

Figure 5.8: Same as Fig. 5.7, but for temperature in K.

6 Summary of 1D-Var experiments

6.1 Single track

For a better cross-comparison between assimilation experiments using different observations, the results in this section are shown in the figures combining the results from 1D-Var assimilation of radar reflectivity, lidar backscatter, separately and in combination. All presented results in this section are coming from the experiments with the quality control, bias correction and the complex error definition applied to observations.

Figure 6.1 shows the probability distribution functions of the first-guess with respect to the analysis departures for the different experiments. Assimilation of either cloud radar reflectivity (a,d) or lidar backscatter (b,e) observations leads to more narrow PDFs of the analysis departures not only for the assimilated (a,e,c,f) but also for the independent observations (d,b), though assimilation of the lidar backscatter provides only a small adjustment to the cloud radar reflectivity. By combining both the radar and lidar observations, a better adjustment (i.e. more narrow distribution for PDF of the analysis departure) is obtained than when assimilating the radar reflectivity only.

Diagrams displaying the rms differences between the first-guess and analysis departures for both CloudSat radar reflectivity and CALIPSO lidar backscatter (Fig. 6.2) demonstrate that although the 1D-Var analyses for any used observations get closer to both assimilated and independent observations, the impact of cloud radar reflectivity is larger than the one of the lidar backscatter. This is also confirmed in Fig. 6.3 which compares the background and analysis departures for MODIS cloud optical depth.

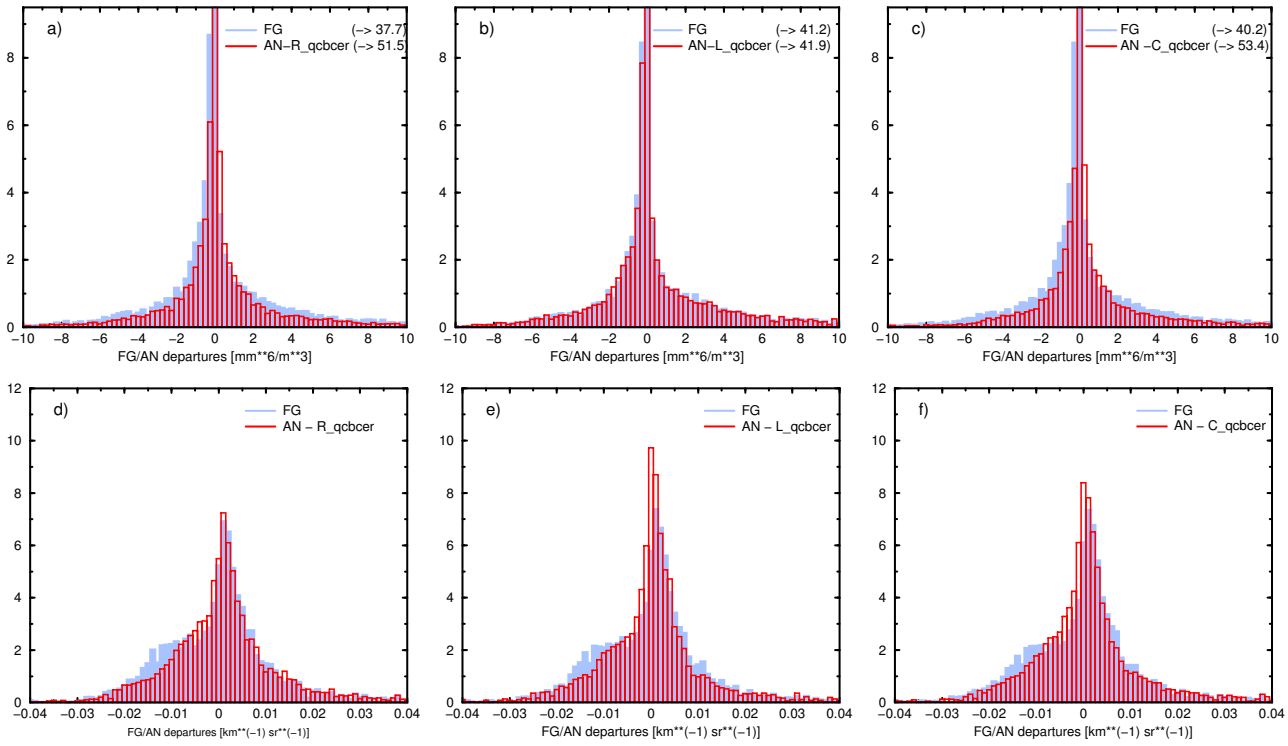

Figure 6.1: Probability distribution functions of the first-guess (light blue shading) and analysis (red line) departures for (a-c) cloud radar reflectivity (in $mm^6\ m^{-3}$) and (d-f) lidar backscatter due to clouds (in $km^{-1}\ sr^{-1}$) coming from 1D-Var retrievals using (a,d) CloudSat reflectivity **R** *and (b,e) CALIPSO backscatter* **L** *observations either separately, or (c,f) in combination* **C**. *Situation over the Pacific Ocean on 24 January 2007.*

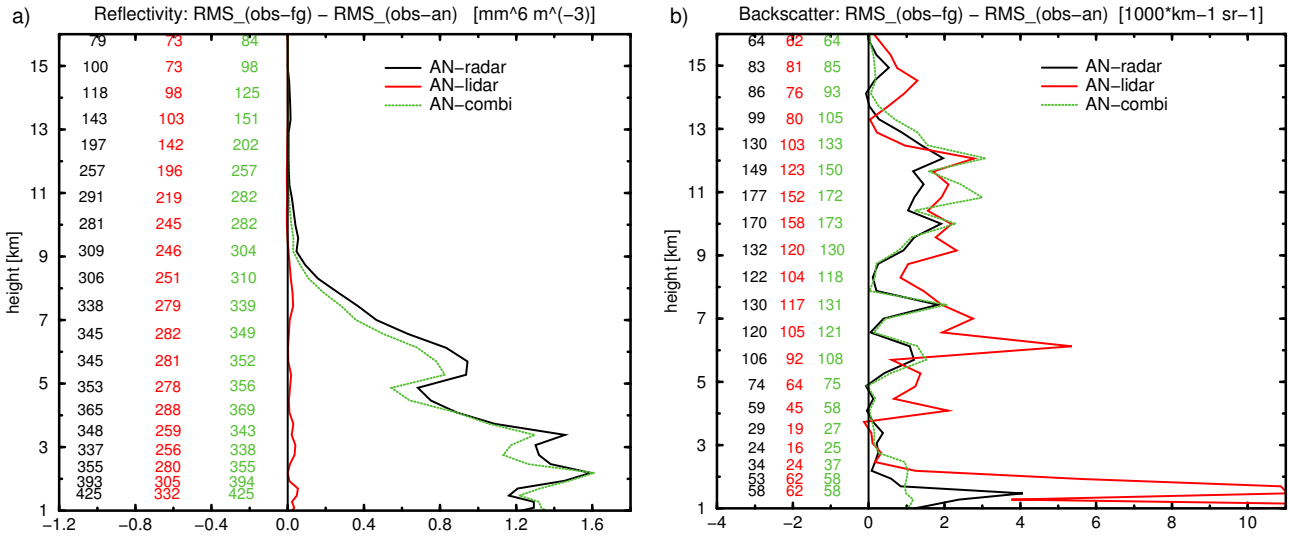

*Figure 6.2: Difference of (a) CloudSat radar reflectivity rms errors (in $mm^6\ m^{-3}$) and (b) CALIPSO lidar backscatter rms errors (in 1000 $km^{-1}\ sr^{-1}$) for differences between the first guess (FG) departures and analysis (AN) departures when assimilating cloud radar reflectivity (**radar**) or lidar backscatter (**lidar**) either separately, or in combination (**combi**). Colour numbers on the left side of (a) and (b) indicate number of observations considered for statistics by the different experiments.*

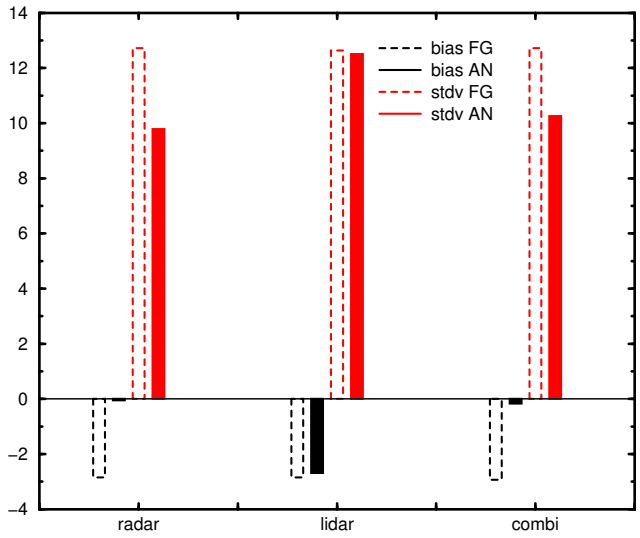

Figure 6.3: Bias (in black) and standard deviation (in red) of the FG (dashed bar) and AN (solid filled bar) departures from MODIS cloud optical depth for the same 1D-Var experiments as in Fig. 6.1. Situation of 24 January 2013.

An evaluation of analysis increments for temperature (Fig. 6.4) and for the derived relative humidity (Fig. 6.5 reveals that comparing to radar, the lidar increments occur at higher altitudes and are therefore complimentary. At altitudes, where both radar and lidar observations are available, the increments are consistent. From distribution of the increments it is possible to identify which increments are due to the radar reflectivity (e.g. between 168° and 175°W, around 150°-155°W or between 135° and 138°W) and which are mostly due to the lidar backscatter (e.g. between 165° and 168°W or between 157° and 161°W).

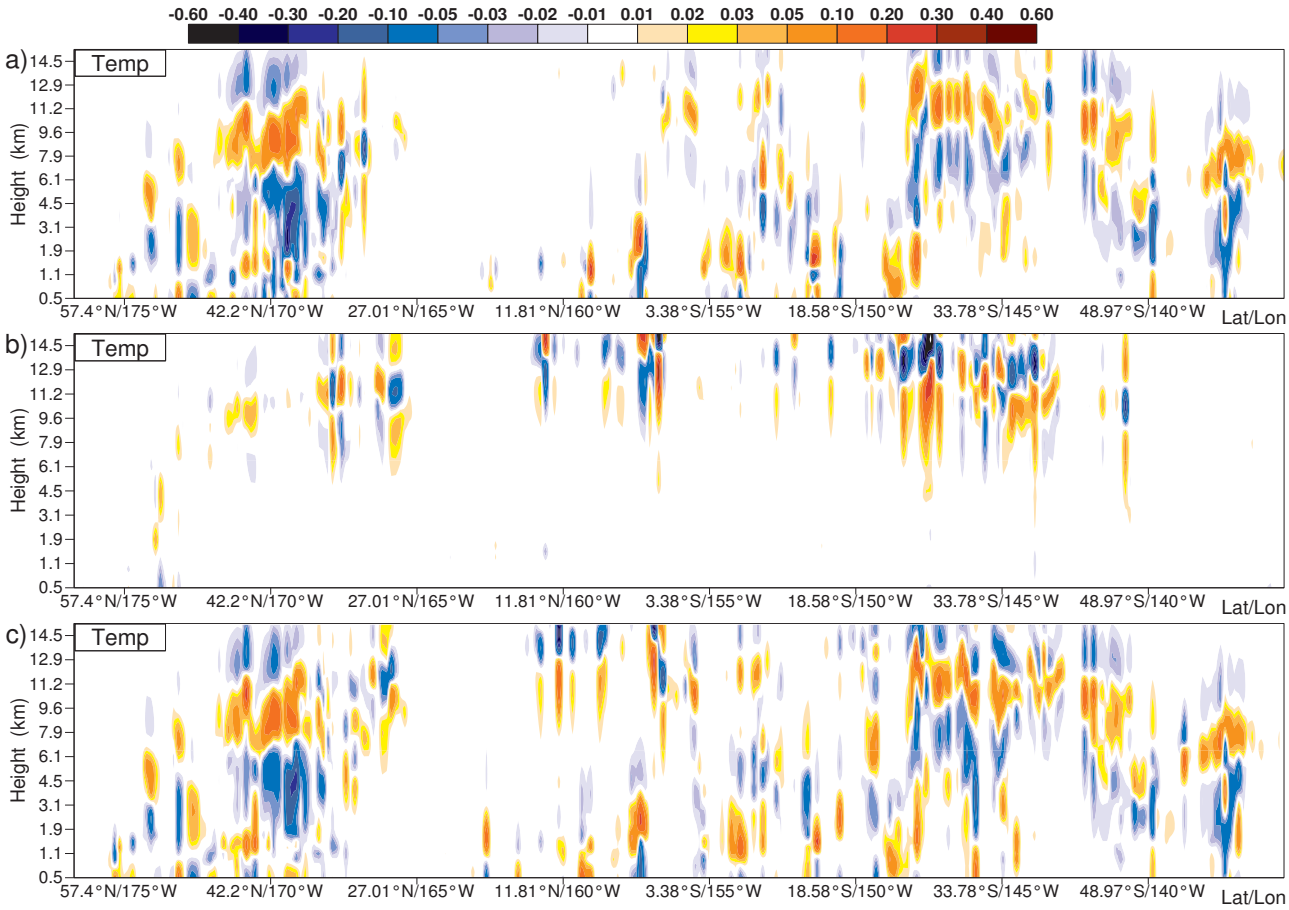

Figure 6.4: Analysis increments for temperature in K from 1D-Var assimilation of (a) CloudSat reflectivity or (b) CALIPSO backscatter either separately or (c) in combination. Situation over the Pacific Ocean on 24 January 2007.

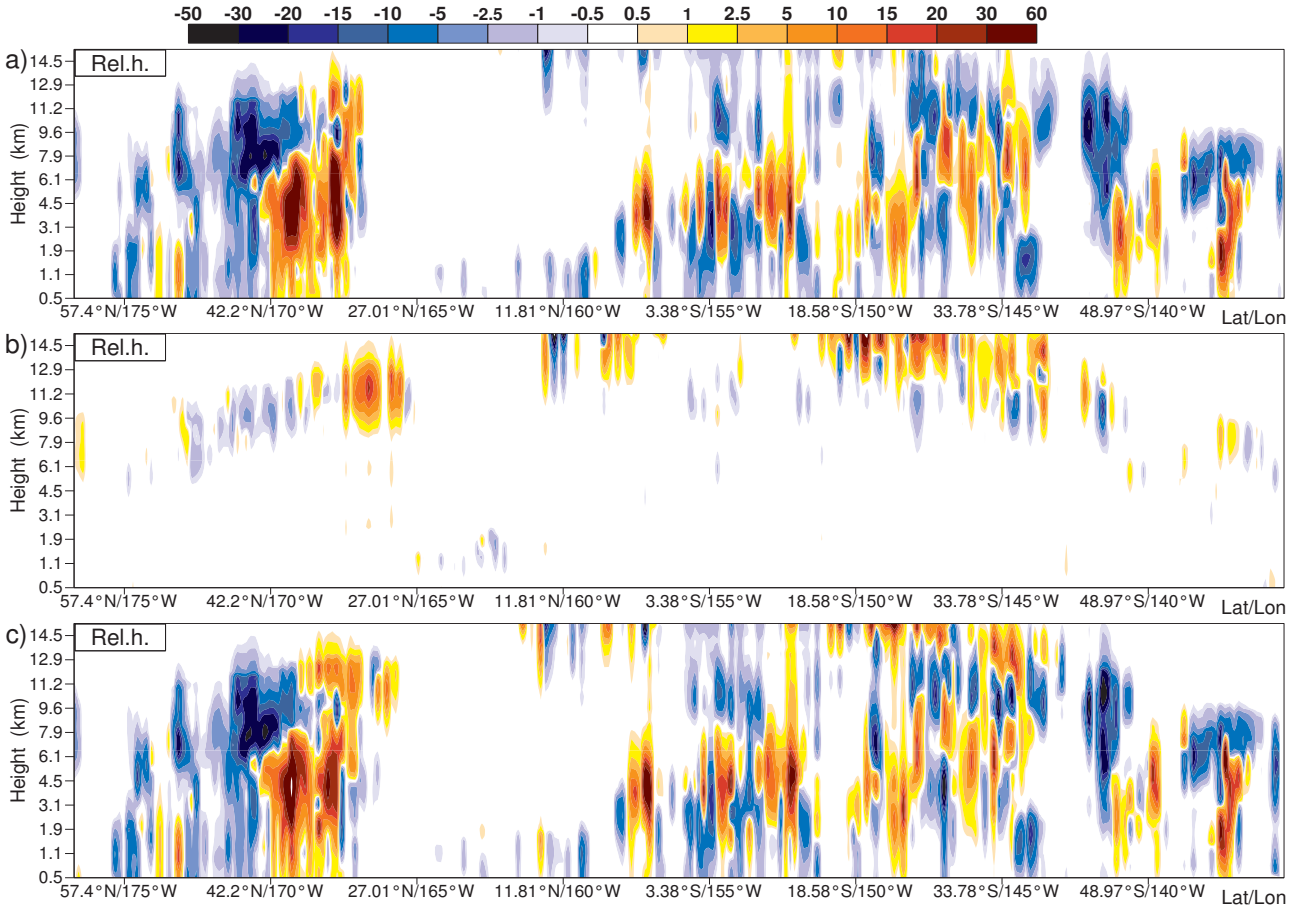

Figure 6.5: Same as Fig. 6.4, but for relative humidity in %.

6.2 Multiple tracks in 12-hour interval to be used in 4D-Var

Summary of the results for 12-hour period corresponding to the length of 4D-Var assimilation window at ECMWF is provided in Tables 6.1 - 6.2 and in Fig. 6.6 - 6.8. The satellite tracks covering the analyzed period between 23 January 2007 21:00 UTC and 24 January 2007 09:00 UTC are displayed in Fig. 8.1.

The probability distribution functions (Fig. 6.6) of the first guess with respect to the analysis departures from the cloud radar reflectivity (a-c) and the lidar backscatter due to clouds (d-f) confirms the results observed from the experiments for the single satellite track over the Pacific Ocean (Section 6.1). The PDFs of the analysis departures from 1D-Var experiments assimilating the cloud radar reflectivity and lidar backscatter either separately or in combination are more narrow for both assimilated and independent observations, with the impact of lidar backscatter being significantly smaller than the one of cloud radar reflectivity. Such difference in impact is more obvious from Fig. 6.7 displaying differences between the rms errors of the first guess departures and the rms of the analysis departures for the CloudSat radar reflectivity and the CALIPSO lidar backscatter. For the lidar backscatter, the largest improvement (i.e. smallest rms errors of the analysis departures) is achieved by assimilation of the combined radar and lidar observations indicated by the positive values in Fig. 6.7b and by the largest negative values in Table 6.2 displaying an error reduction. For the lidar radar reflectivity, assimilation of the combined radar and lidar observations provides better results above 13 km than using the radar reflectivity alone, below that altitude there are very small differences between these two experiments.

A general error reduction of the analysis departures with respect to the first-guess departures for all assimilated experiments is also demonstrated in Table 6.1. Using MODIS cloud optical depths as validation data (Fig. 6.8) shows that the 1D-Var analyses get closer to these independent observations, though the impact of lidar backscatter is very small indeed.

	reflectivity					backscatter				
	bias	stdv	mae	rms	NOBS	bias	stdv	mae	rms	NOBS
FG-R_qcbcer	-0.223	7.260	1.360	7.264	103105	1.42E-02	4.49E-02	2.01E-02	4.71E-02	45873
AN-R_qcbcer	0.133	6.989	0.816	6.990		1.32E-02	4.35E-02	1.91E-02	4.55E-02	
FG-L_qcbcer	0.074	8.664	1.255	8.664	84715	1.27E-02	4.65E-02	2.01E-02	4.71E-02	43448
AN-L_qcbcer	0.072	8.663	1.252	6.663		1.24E-02	4.62E-02	1.91E-02	4.55E-02	
FG-C_qcbcer	-0.209	7.419	1.394	7.422	100121	1.45E-02	4.61E-02	2.05E-02	4.84E-02	45435
AN-C_qcbcer	0.155	7.149	0.849	7.151		1.35E-02	4.44E-02	1.93E-02	4.64E-02	

Table 6.1: Bias, standard deviation (stdv), mean absolute error (mae) and root mean square error (rms) of the first guess (FG) and analysis (AN) departures for the different assimilation experiments (see text for experiment description) from the CloudSat cloud reflectivity (in mm^6 m^{-3}) and CALIPSO cloud backscatter (in km^{-1} sr^{-1}) observations. NOBS indicate number of observations used in the statistics for 12-hour period from 23 January 2007 21:00 UTC to 24 January 2007 09:00 UTC.

	reflectivity		backscatter	
	DIFF	RMSD	DIFF	RMSD
AN-R_qcbcer	-4.134	-4.354	-7.829E-04	-8.686E-04
AN-L_qcbcer	-0.079	-0.059	-5.928E-04	-6.706E-04
AN-C_qcbcer	-4.062	-4.290	-1.143E-03	-1.521E-04

Table 6.2: Difference between the absolute value of analysis (AN) minus observation (OBS) and absolute value of the first guess (FG) and OBS (DIFF), as well as difference of rms errors for the FG and AN departures (RMSD) from the CloudSat cloud reflectivity (in dBZ) and the backscatter (in km^{-1} sr^{-1}) observations. Results are shown for the different assimilation experiments (see text for experiment description) for 12-hour period from 23 January 2007 21:00 UTC to 24 January 2007 09:00 UTC.

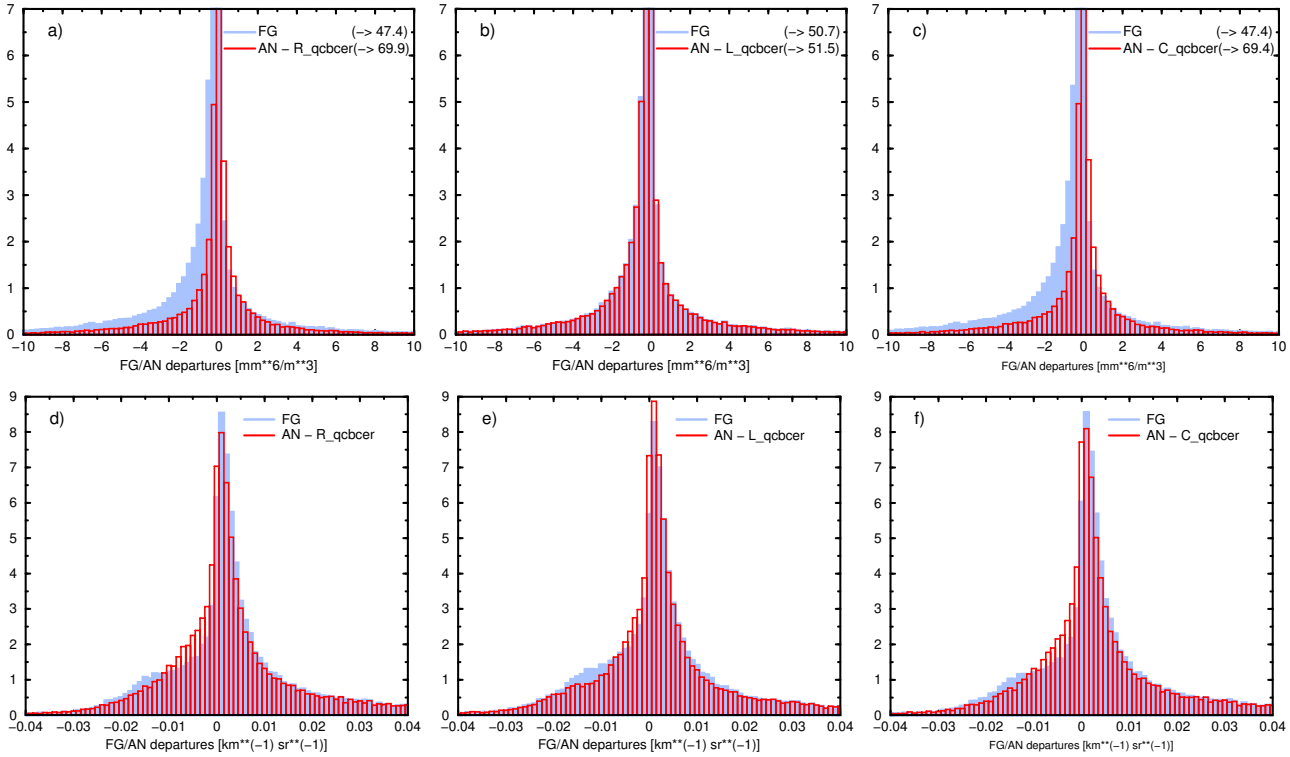

Figure 6.6: Probability distribution functions of the first-guess (light blue shading) and analysis (red line) departures for (a-c) cloud radar reflectivity (in $mm^6\ m^{-3}$) and (d-f) lidar backscatter due to clouds (in $km^{-1}\ sr^{-1}$) coming from 1D-Var retrievals using (a,d) CloudSat reflectivity **R** *and (b,e) CALIPSO backscatter* **L** *observations either separately, or (c,f) in combination* **C**. *12-hour period from 23 January 2007 21:00 UTC to 24 January 2007 09:00 UTC.*

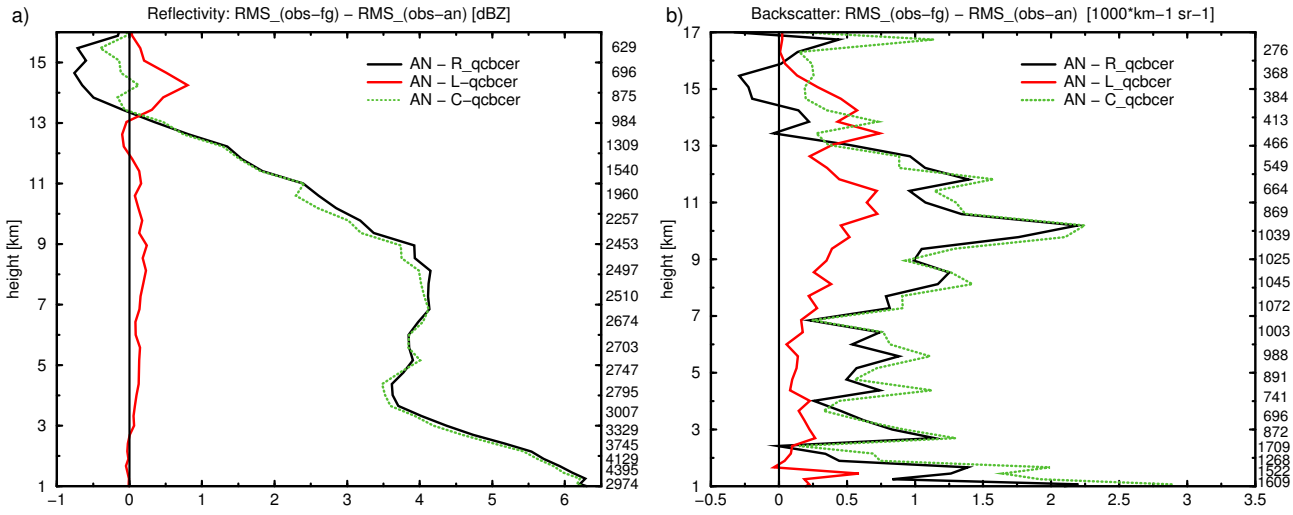

Figure 6.7: Difference of (a) CloudSat radar reflectivity rms errors (in dBZ) and (b) CALIPSO lidar backscatter rms errors (in 1000 $km^{-1}\ sr^{-1}$) for differences between the first guess (FG) departures and analysis (AN) departures when assimilating cloud radar reflectivity (black solid line) or lidar backscatter (red solid line) either separately, or in combination (green dotted line). Numbers on the right side of (a) and (b) indicate an average number of observations considered for statistics by the different experiments for 12-hour period from 23 January 2007 21:00 UTC to 24 January 2007 09:00 UTC.

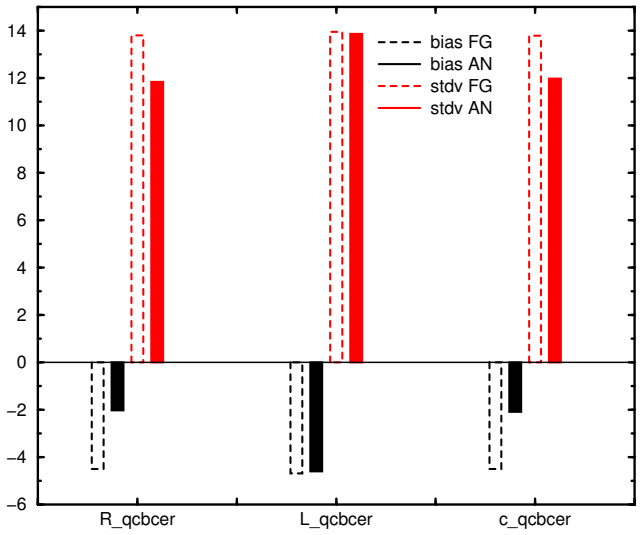

Figure 6.8: Bias (in black) and standard deviation (in red) of the FG (dashed bar) and AN (solid filled bar) departures from MODIS cloud optical depth for the same 1D-Var experiments as in Fig. 6.6. 12-hour period from 23 January 2007 21:00 UTC to 24 January 2007 09:00 UTC.

7 1D+4D-Var technique

In the past years, a 1D+4D-Var approach has been developed for the assimilation of rain affected observations at ECMWF (Bauer et al., 2006b; Lopez and Bauer, 2007). This technique has also been used in the study of Janisková et al. (2012) which demonstrated the positive impact that the 4D-Var assimilation of temperature and humidity pseudo-observations derived from CloudSat has on the analysis and forecast of temperature, humidity and winds. As an extension of this study, lidar observations or combined radar/lidar observations have been also used in 1D-Var to produce the pseudo-observations.

7.1 Description of the 1D+4D-Var approach

The simple diagram (Fig. 5) provides the schematic description of 1D+4D-Var assimilation. In the first step, a 1D-Var assimilation technique is used to assimilate CloudSat reflectivity and CALIPSO backscatter due to clouds either separately or in combination. The 1D-Var assimilation is performed with the aim of adjusting the model temperature and specific humidity profiles. The second step consists in the assimilation of specific humidity (q) and temperature (T) profiles retrieved from off-line 1D-Var assimilation of radar and lidar data as pseudo-observations into the 4D-Var system in order to study the impact of the new observations on 4D-Var analyses and subsequent forecasts.

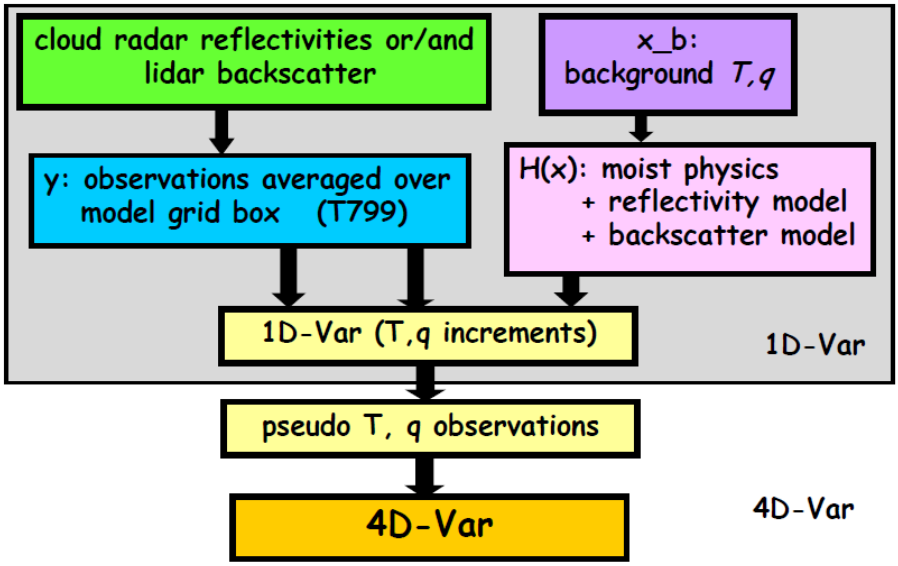

Figure 7.1: Schematic description of 1D+4D-Var system for assimilation of cloud related observations.

7.1.1 4D-Var system

Since November 1997 the operational data assimilation system at ECMWF is a four-dimensional variational (4D-Var) system (Rabier et al., 2000; Mahfouf and Rabier, 2000). This system is based on the incremental formulation proposed by Courtier et al. (1994).

4D-Var seeks an optimal balance between observations and the dynamics of the atmosphere by finding a model trajectory $\mathbf{x}(t)$ which is as close as possible, in a least-square sense, to the observations available during a given time period $[t_0, t_n]$. The model trajectory $\mathbf{x}(t)$ is completely defined by the initial state $\mathbf{x}_0$ at time t_0.

The misfit to given observations $\mathbf{y}^o$ and to an *a-priori* model state $\mathbf{x}^b$ called the background, which is usually provided by a short-range forecast, is measured by the objective cost-function defined as:

$$\mathscr{J}(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{i=0}^{n}(H_i[\mathbf{x}(t_i)] - \mathbf{y}_i^o)^T \mathbf{R}_i^{-1}(H_i[\mathbf{x}(t_i)] - \mathbf{y}_i^o) \tag{7.1}$$

where at any time t_i,

- $\mathbf{y}_i^o$ is the vector of observations;
- $\mathbf{x}_0^b$ is the background model state at the initial time and in the current 4D-Var system it consists of temperature, humidity, vorticity, divergence and surface pressure;
- H_i is the operator providing the equivalent of the observed data from the model variable $\mathbf{x}(t_i)$, it also describes the spatial interpolations to the observation locations and the action of the forecast model to propagate the initial state $\mathbf{x}_0$ to the time of observation;
- $\mathbf{R}_i$ is the observation error covariance matrix (including measurement and representativeness errors);
- $\mathbf{B}$ 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.

7.2 Pseudo-observations and their errors used in 4D-Var

The performed 1D-Var assimilation experiments have revealed that the assimilated cloud related observations have modified both control variables of the 1D-Var system, i.e temperature and specific humidity, as shown by the size of analysis increments. Therefore pseudo-observations of both temperature (T) and specific humidity (q) profiles from 1D-Var are included in the 4D-Var system.

The determination of T and q observation errors is not straightforward since these pseudo-observations come from the 1D-Var retrievals and therefore the observation errors correspond to the 1D-Var retrieval errors. These depend on the background error assumed for the 1D-Var control variables, the observation errors (either for reflectivity or for backscatter) and the accuracy of the observation operator used by the 1D-Var system. As shown by Rodgers (2000) and also used in 1D+4D-Var approaches described by Bauer et al. (2006b), Benedetti et al. (2006) or Lopez and Bauer (2007), the error on the retrieved variables can be directly derived from the 1D-Var analysis error covariance matrix, $\mathbf{A}$, defined as follows:

$$\mathbf{A} = \left[\mathbf{B} + \mathbf{K}^T(\mathbf{x})\mathbf{R}^{-1}\mathbf{K}(\mathbf{x})\right]^{-1} \tag{7.2}$$

where $\mathbf{K} = \left[\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}}\right]$ represents the Jacobian matrix of the partial derivatives of the 1D-Var observation operator H with respect to the control variable $\mathbf{x}$. This error computation could be influenced by the problem of possible non-linearity of the observation operator. However, based on the convergence performance of the 1D-Var system, it seems that the linearity of the observation operator is ensured for the majority of cases. Since the dimension of the observation vector is quite high when using profiling observations, the computation of error is quite expensive even when using the adjoint to derive the Jacobian matrix. But due to the lack of other error estimation options, Eq. 7.2 has been used as an affordable approach for non-operational applications (i.e. feasibility studies performed for this project).

The errors for T and q produced by Eq. 7.2 are in average approximately twice as small as those used for the radiosonde temperature and specific humidity measurements. Since the observation errors partly influence the impact of observations in data assimilation, experiments have also been performed with increased errors compared to those which were directly derived from the 1D-Var analysis error covariance matrix $\mathbf{A}$.

8 1D+4D-Var experiments for combined CloudSat and CALIPSO cloud related observation

8.1 Experimental setup

Several 4D-Var experiments have been performed for the 12-hour period (i.e. the current length of 4D-Var assimilation window at ECMWF) between 23 January 2007 21:00 UTC and 24 January 2007 09:00 UTC (Fig. 8.1) using the full system of regularly assimilated observations:

- **ref**: reference run, i.e. run without pseudo-observations included in the 4D-Var system;
- **R**: 4D-Var experiment assimilating T and q pseudo-observations retrieved from 1D-Var with cloud radar reflectivity, using observation errors derived from the 1D-Var analysis covariance matrix (Eq. 7.2);
- **R_2err**: same as **R**, but using observation errors twice as large as the computed ones;
- **L**: 4D-Var experiment assimilating T and q pseudo-observations retrieved from 1D-Var with lidar backscatter due to clouds, using observation errors derived from the 1D-Var analysis covariance matrix;
- **L_2err**: same as **L**, but using observation errors twice as large as the computed ones;
- **C**: 4D-Var experiment assimilating T and q pseudo-observations retrieved from 1D-Var with the combined cloud radar reflectivity and lidar backscatter observations , using observation errors as derived from Eq. 7.2;
- **C_2err**: same as **C**, but using observation errors twice as large as the computed ones.

A special experiment with a reduced data set, i.e. limited to the satellite track over the Pacific Ocean, has also been run using q pseudo-observations only. The results from 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using lidar backscatter only are not presented since the impact of this type of observations on its own is very small, as already indicated by 1D-Var experiments (Sections 4 and 5).

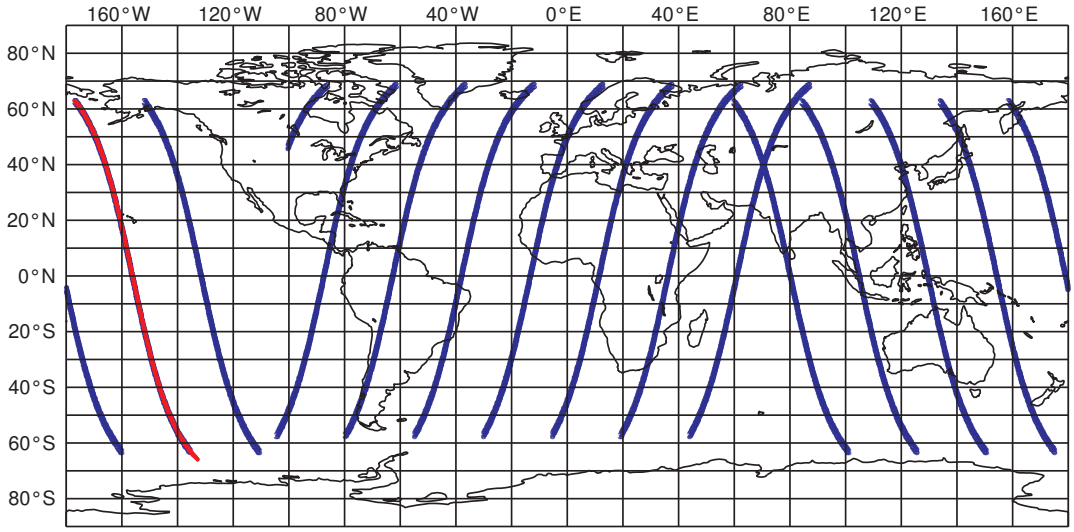

Figure 8.1: CloudSat/CALIPSO data coverage for 12-hour period between 23 January 2007 21:00 UTC and 24 January 2007 09:00 UTC. Red track highlights the situation over the Pacific Ocean used for extended 1D-Var experimentation.

1D+4D-Var assimilation experiments have been run over one assimilation cycle of 12 hours and 10 day forecasts have been run from the analyses to study the impact of the new observations on 4D-Var analyses and on

the subsequent forecasts. Only profiles with non-zero increments of temperature and specific humidity obtained from 1D-Var assimilation have been used in 4D-Var. Therefore a coverage of pseudo-observations of temperature and specific humidity is slightly reduced (as indicated in Fig. 8.2) compared to the full CloudSat/CALIPSO data coverage. Zero increments can be a consequence of (i) the quality control blacklisting observations when differences between observations and the model equivalents exceed a certain threshold, (ii) failed 1D-Var convergence due to non-linearities in observation operator or (iii) setting increments to zero when they exceed the typical values of the standard deviation of the ECMWF model background (Fig. 2.1) by five times (a kind of additional quality control for 1D-Var system).

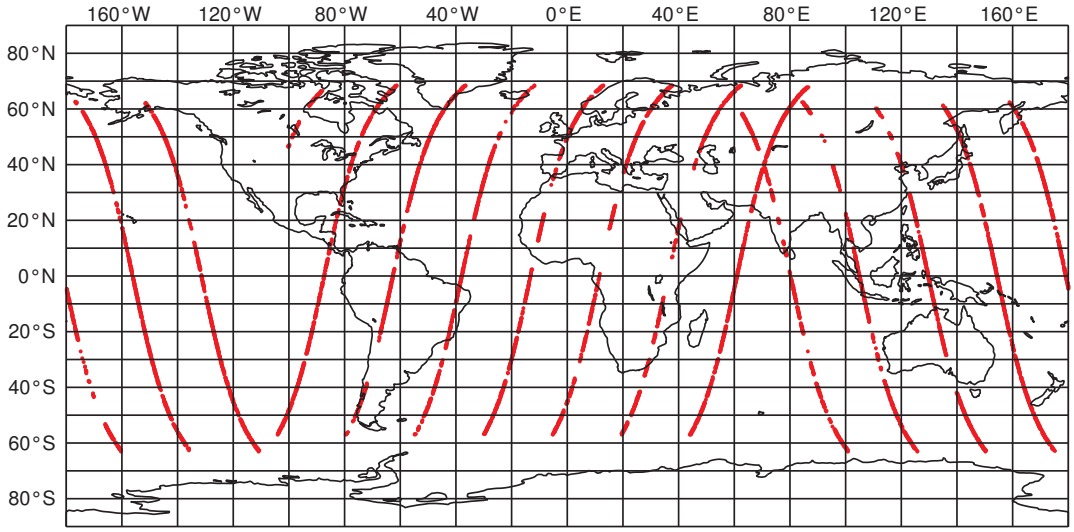

Figure 8.2: Coverage of pseudo-observations of temperature and specific humidity used in 4D-Var for 12-hour period between 23 January 2007 21:00 UTC and 24 January 2007 09:00 UTC.

8.2 Results

(a) Impact on 4D-Var analyses

Profiles of bias and standard deviation of temperature and specific humidity for the first-guess and analysis departures for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity either separately or in combination with the lidar backscatter are shown in Fig. 8.3-8.5. Though bias and, between levels 550 and 850 hPa, standard deviation of the analysis departures are reduced for specific humidity, the analysis departures for temperature are (surprisingly) larger that the first-guess ones. This is a common feature for the experiments using pseudo-observations coming from 1D-Var assimilation experiments of separate radar or lidar observations as well as of combined observations. When using observation errors for T, q pseudo-observations as derived from the 1D-Var analysis covariance matrix, only the analysis departures for temperature are reduced more than when using observation errors twice as large as the computed ones. In general, bias and standard deviation of the analysis departures are smaller for the experiment assimilating pseudo-observations from 1D-Var of the combined observations than the separate ones, which clearly indicates a small but positive impact coming from the use of lidar backscatter on top of radar reflectivity.

In order to understand why the analysis departures, especially for temperature, are larger than the first guess ones, Fig. 8.6 displays profiles of the departures for particular steps of the 4D-Var assimilation cycle. The incremental 4D-Var assimilation system of ECMWF uses three so called inner loops, in which minimization

is performed at a lower horizontal resolution than in trajectory computations (i.e. forecast or trajectory update runs) starting from the coarsest resolution (T95 in the first inner loop, T159 in the second and T225 in the third loop). Each minimization run is followed by the trajectory run at high resolution (T799 in this study). The first-guess departures followed by the departures after each trajectory update and finally the analysis departures displayed in Fig. 8.6 indicate that it is the first minimization (1st inner cycle) which is at the end responsible for the final analysis departures being larger than the first-guess ones. This first minimization is done at a coarse horizontal resolution (T95 corresponding to approximately 200 km) and in general it is supposed to adjust mostly the model large-scale features. On the contrary, the information retrieved from the cloud related observations represents rather small scales. In addition, the first-guess departures for temperature are rather small and comparable with typical background errors for temperature. These two possible reasons may be responsible for the fact that the departures after the first inner loop are much larger than the first-guess ones. During the subsequent inner loops, the departures are reduced thus indicating that information from the pseudo-observations is being taken into account by 4D-Var analysis.

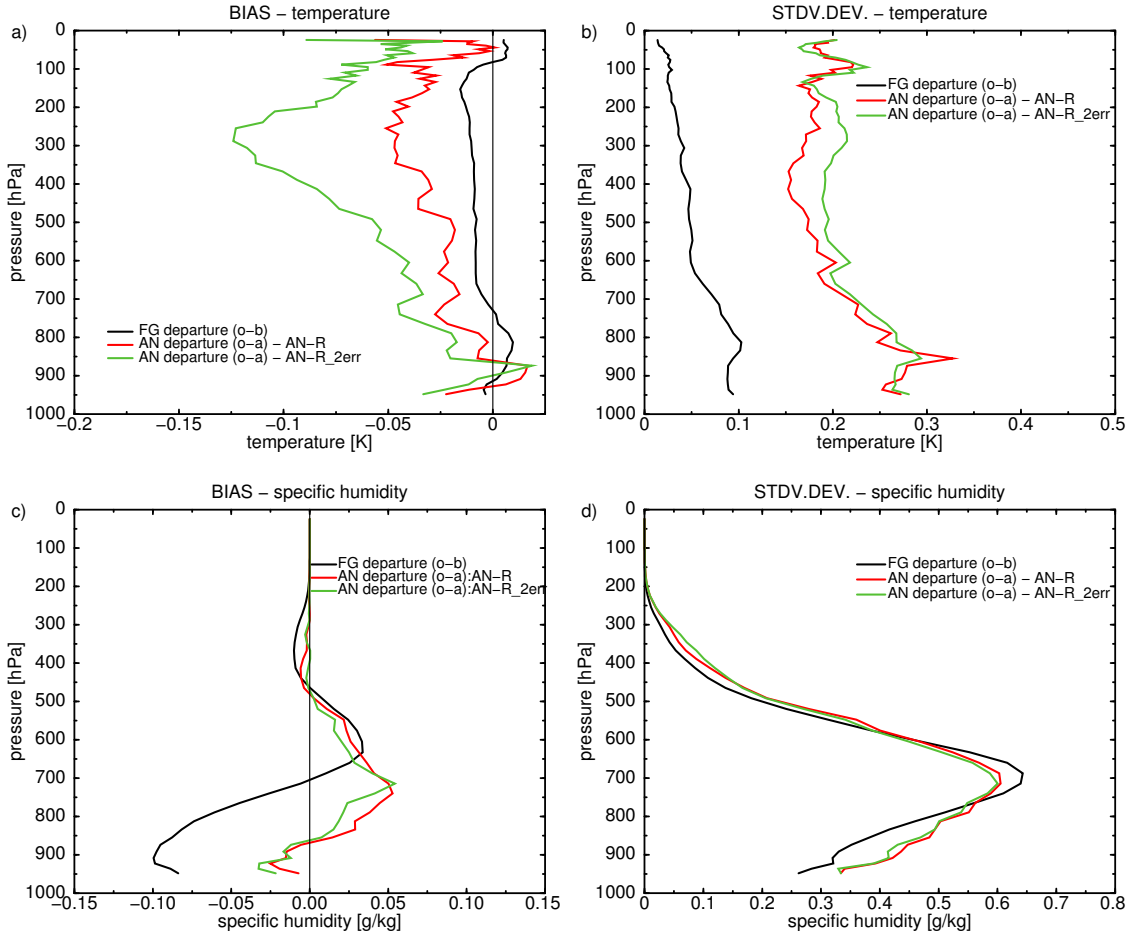

Figure 8.3: Profiles of (a,c) bias and (b,d) standard deviation of (a,b) temperature (in K) and (c,d) specific humidity (in g.kg^{-1}) for the first guess (black solid line) and analysis departures for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity with observation errors either as computed from the 1D-Var analysis covariance matrix (red solid line) or twice as large as the computed ones (green solid line). Case of analysis valid on 24 January 2007 00:00 UTC.

To confirm that indeed pseudo-observations of temperature play a role in 4D-Var assimilation despite the fact that the analyses are not closer to them than the first guess, an experiment has been run where only q pseudo-observations retrieved from 1D-Var have been included in 4D-Var. The results from such experiment are shown

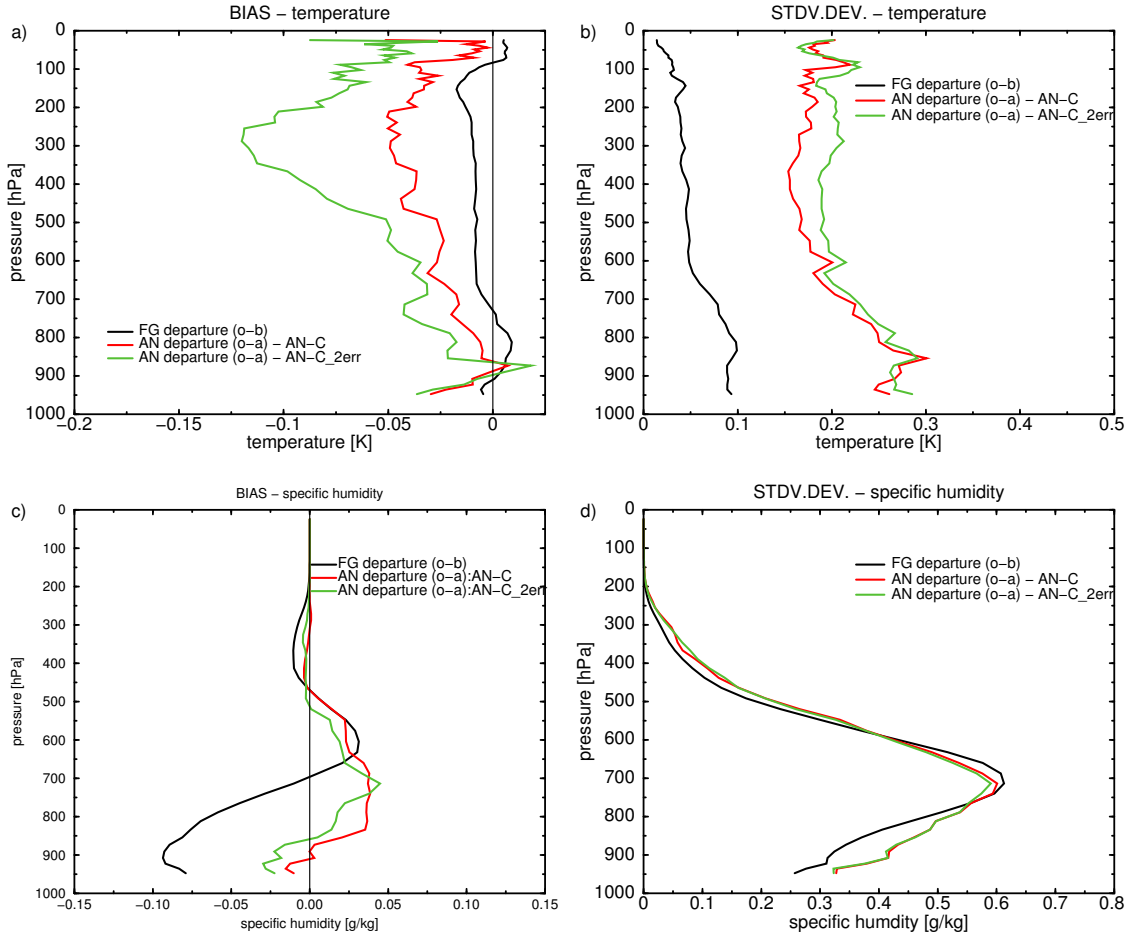

Figure 8.4: Same as Fig. 8.3, but for 4D-Var experiments assimilating pseudo-observations retrieved from 1D-Var using observations of both cloud radar reflectivity and lidar backscatter.

in Fig. 8.7 together with the 4D-Var experiment assimilating both T and q pseudo-observations. These experiments have been performed using the single satellite track over the Pacific Ocean (Fig. 8.1) and pseudo-observations have been taken from 1D-Var with cloud radar reflectivity. One can see that analysis departures from T pseudo-observations when not assimilated in 4D-Var are significantly larger than in the case of their assimilation, while the analysis departures from q pseudo-observations remain comparable. This indicate that T pseudo-observations should also be used. This contrasts with the 1D+4D-Var precipitation assimilation of Bauer et al. (2006b) where only humidity information in the form of the total column water vapour retrieved from 1D-Var of precipitation-affected microwave radiances was included in the 4D-Var system.

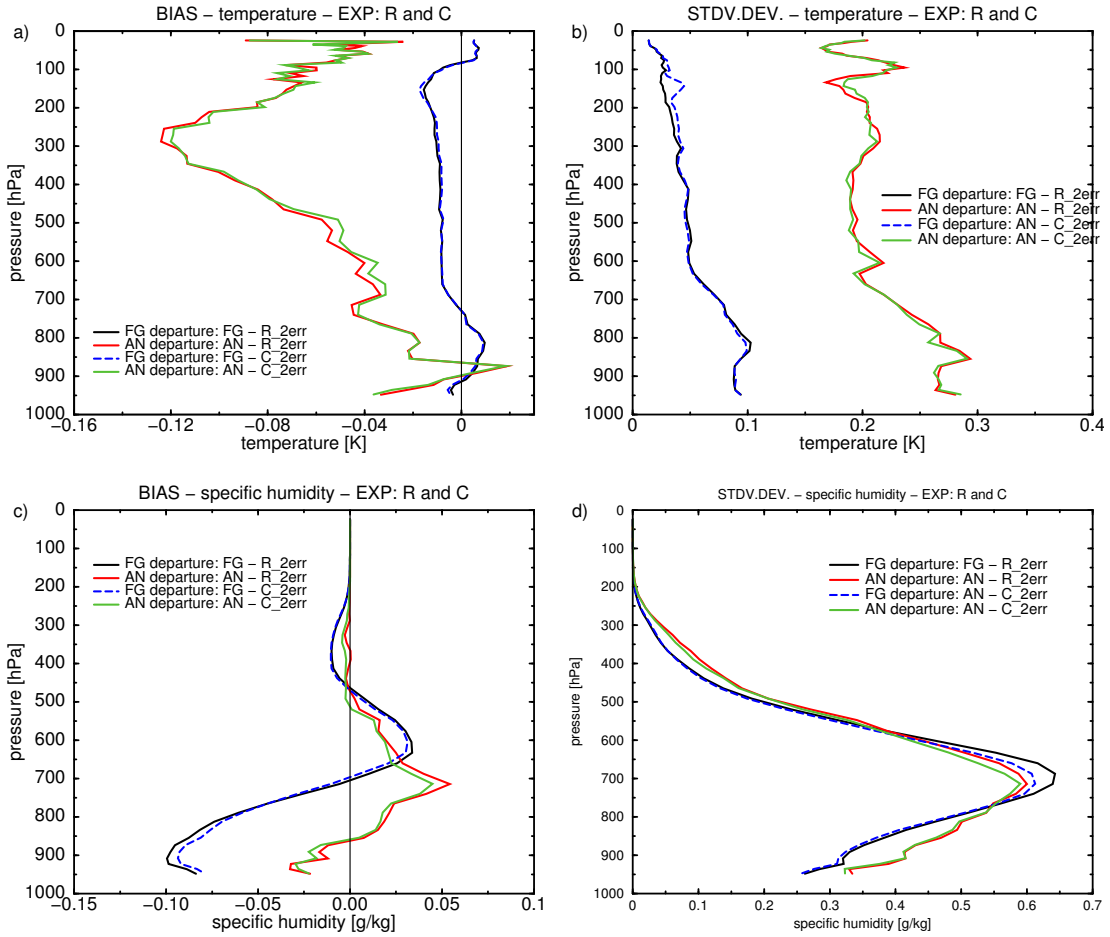

Figure 8.5: Profiles of (a,c) bias and (b,d) standard deviation of (a,b) temperature (in K) and (c,d) specific humidity (in $g.kg^{-1}$) for the first guess and analysis departures for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of either cloud radar reflectivity separately or in combination with lidar backscatter applying observation errors twice as large as the computed ones from the 1D-Var analysis covariance matrix. Different colour lines are explained by the figure legends. Case of analysis valid on 24 January 2007 00:00 UTC.

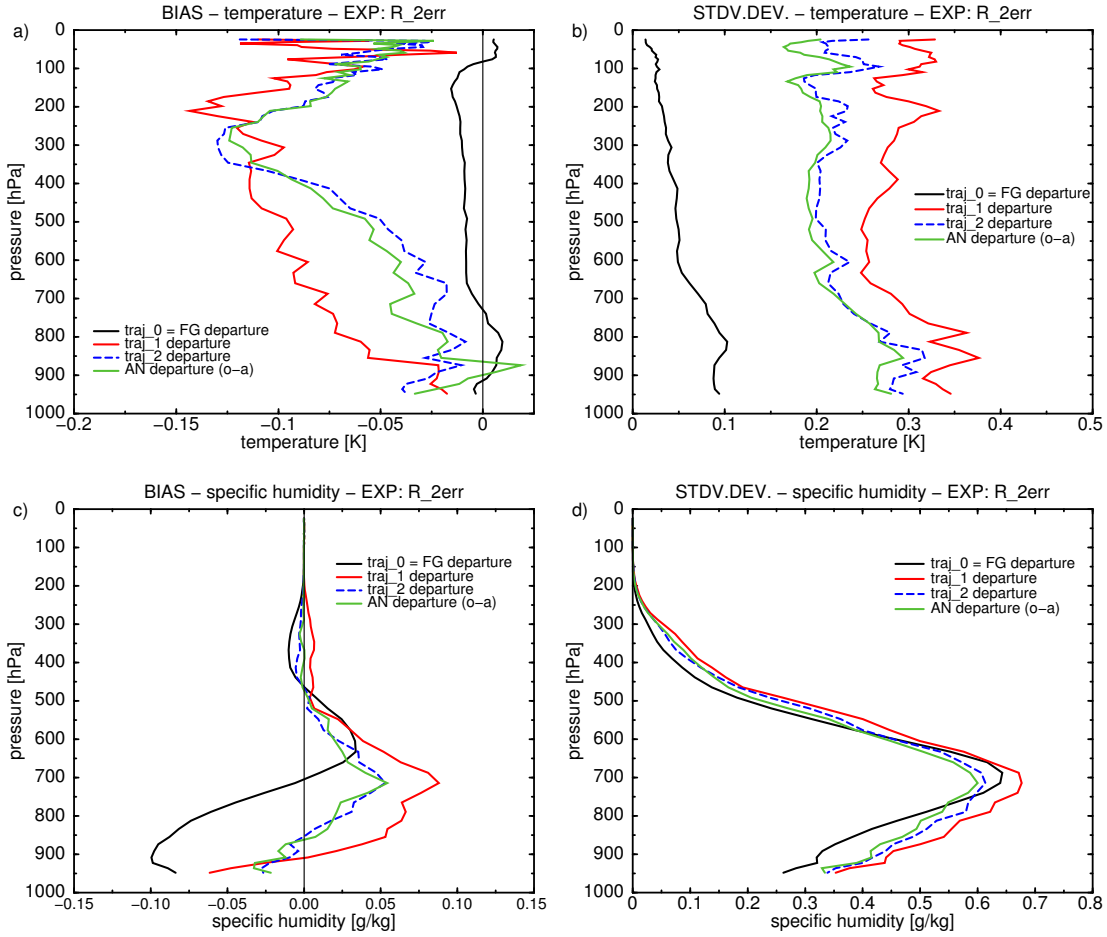

Figure 8.6: Profiles of (a,c) bias and (b,d) standard deviation of (a,b) temperature (in K) and (c,d) specific humidity (in g.kg^{-1}) of the model departures for the first trajectory run corresponding to the first guess (black solid line), the trajectory runs after the first (red solid line) and the second (blue dashed line) minimizations as well as for the analysis departures (green solid line) of 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity with observation errors twice as large as the computed ones from the 1D-Var analysis covariance matrix. Case of analysis valid on 24 January 2007 00:00 UTC.

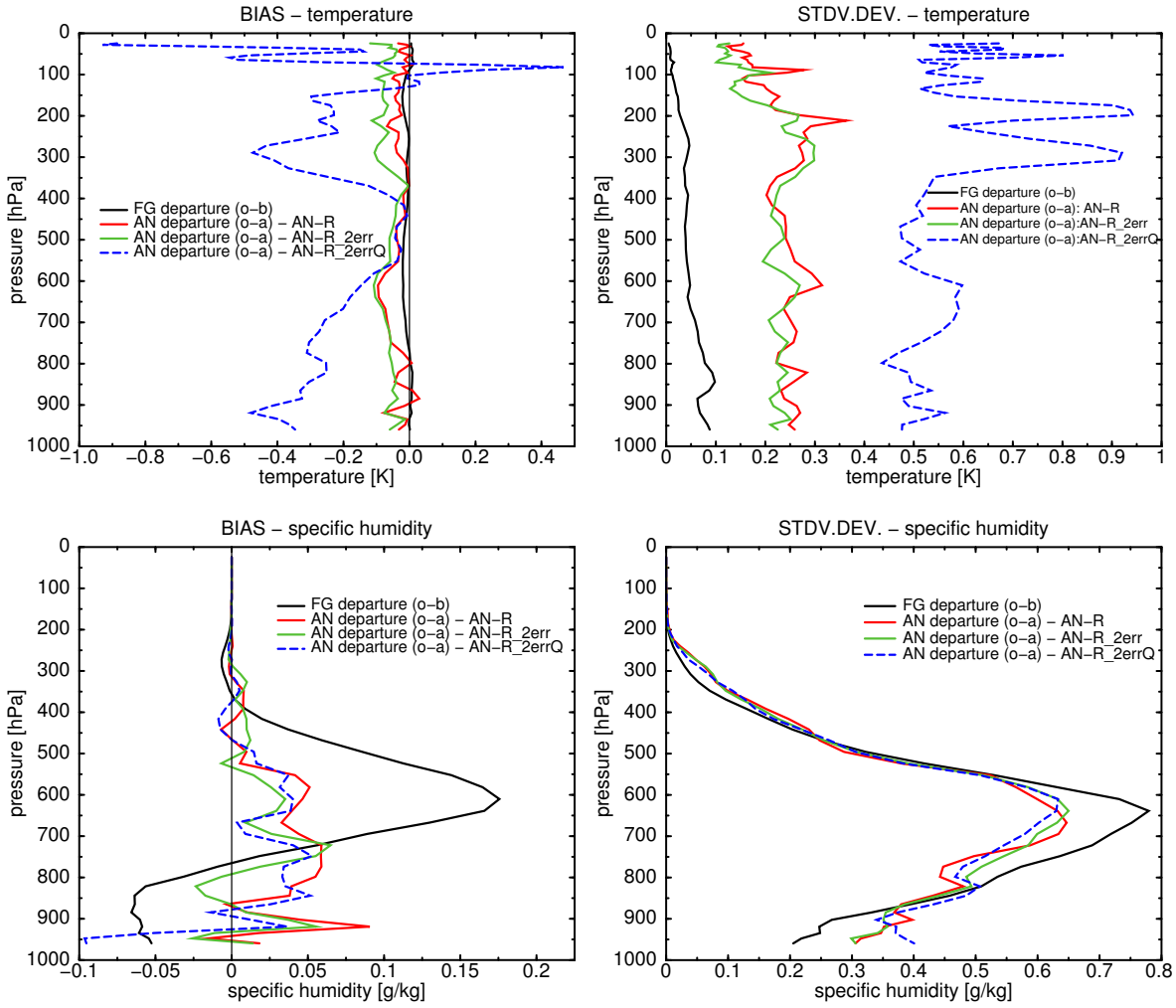

Figure 8.7: Profiles of (a,c) bias and (b,d) standard deviation of (a,b) temperature (in K) and (c,d) specific humidity (in g.kg^{-1}) for the first guess (black solid line) and analysis departures for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity with observation errors either as computed from the 1D-Var analysis covariance matrix (red solid line) or twice as large as the computed ones (green solid line). Dashed blue line represents analysis departure for 4D-Var experiment assimilating only q pseudo-observations with observation errors twice as large as computed. Case of 24 January 2007 over the Pacific Ocean.

Verification of the performed assimilation runs has also been carried out against other assimilated observation types and the results are presented in Fig. 8.8-8.12. First an impact of the pseudo-observation error size has been studied. As indicated by bias and standard deviation of the background and the analysis departures for some selected observations (Fig. 8.8-8.10), when using pseudo-observation errors as derived from the 1D-Var analysis covariance matrix (Eq. 7.2), standard deviation of the analysis departures is systematically larger than in the case with pseudo-observation errors twice as large as the computed ones or than in the reference run. For bias, the results are more mixed, but on average, biases are smaller when the double pseudo-observation errors are used.

While for wind observations (such as TEMP radiosonde observations - Fig. 8.8 or PILOT observations - Fig. 8.9 over the Northern hemisphere) bias of the analysis departures is generally larger in the experimental runs than in the reference one, for temperature observations (such as TEMP radiosondes over the Northern hemisphere and Tropics - Fig. 8.10) bias of the experimental runs (mainly with the double pseudo-observation errors) is smaller than of the reference run for most of the pressure levels.

Getting systematically larger standard deviations in the experimental runs compared to the reference run, while bias being more comparable, indicates that the new cloud related observations bring more variability into the system which may not be so surprising as clouds often represent small scale features.

Since the aim of 4D-Var is to find an initial atmospheric state which provides the best compromise between the model and all assimilated observations, using the double error for pseudo-observations (i.e. reducing the weight given to these observations in the assimilation cycle) seems to be a logical choice based on the results shown in Fig. 8.8-8.10 and discussed above.

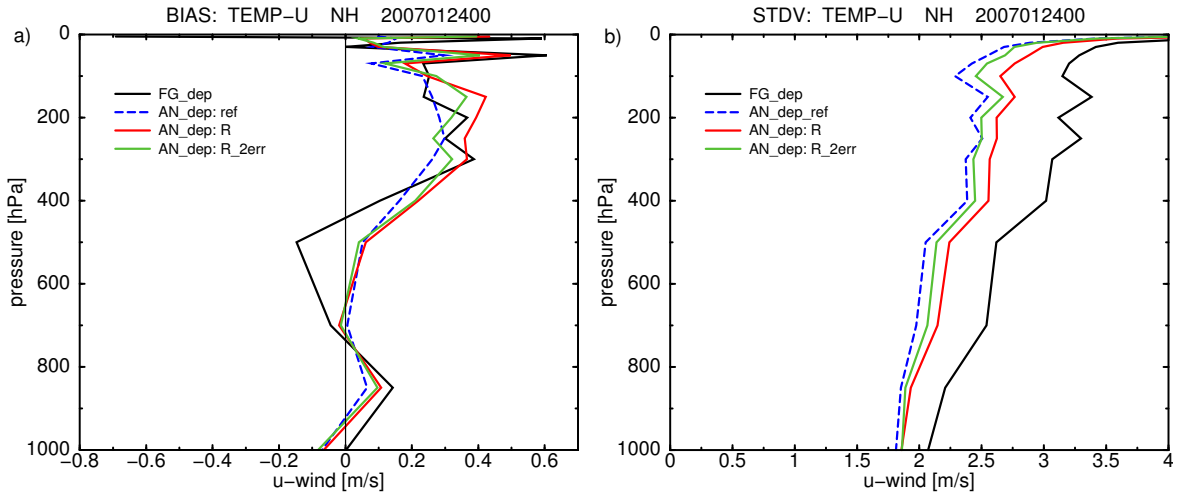

Figure 8.8: Bias (a) and standard deviation (b) of background departures (black solid line) and analysis departures with respect to TEMP observations of zonal wind over the Northern hemisphere for the reference 4D-Var run (blue dashed line) and for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity with observation errors either as computed from the 1D-Var analysis covariance matrix (red solid line) or twice as large as the computed ones (green solid line). Case of analysis valid on 24 January 2007 00:00 UTC.

A comparison of the results from the 4D-Var experiments assimilating T, q pseudo-observations retrieved from 1D-Var using the cloud radar reflectivity separately and in combination with the lidar backscatter (Fig. 8.11-8.12) clearly demonstrates the small but systematic improvements coming from the lidar observations when combined with the radar.

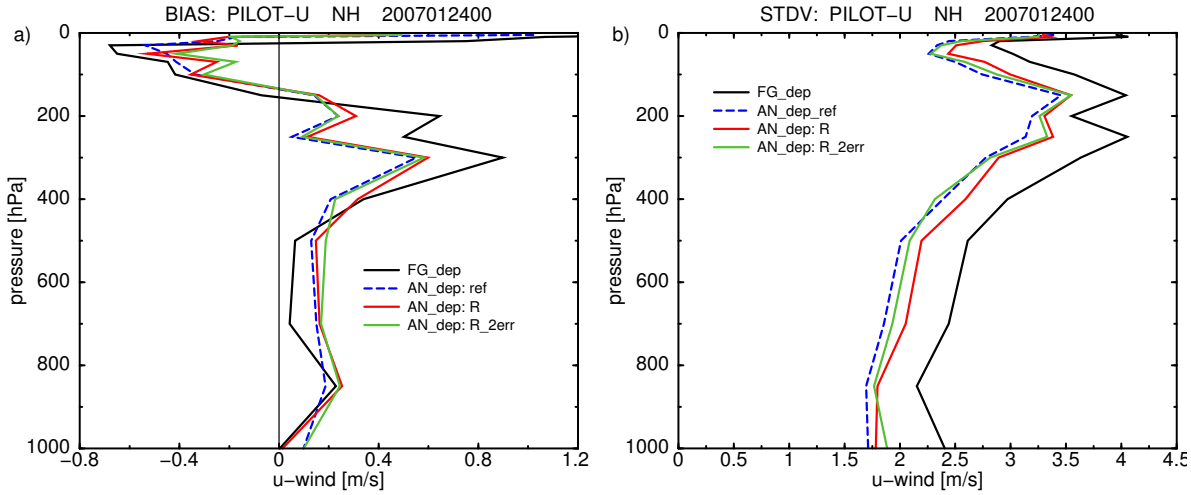

Figure 8.9: Same as Fig. 8.8, but for PILOT observations of zonal wind.

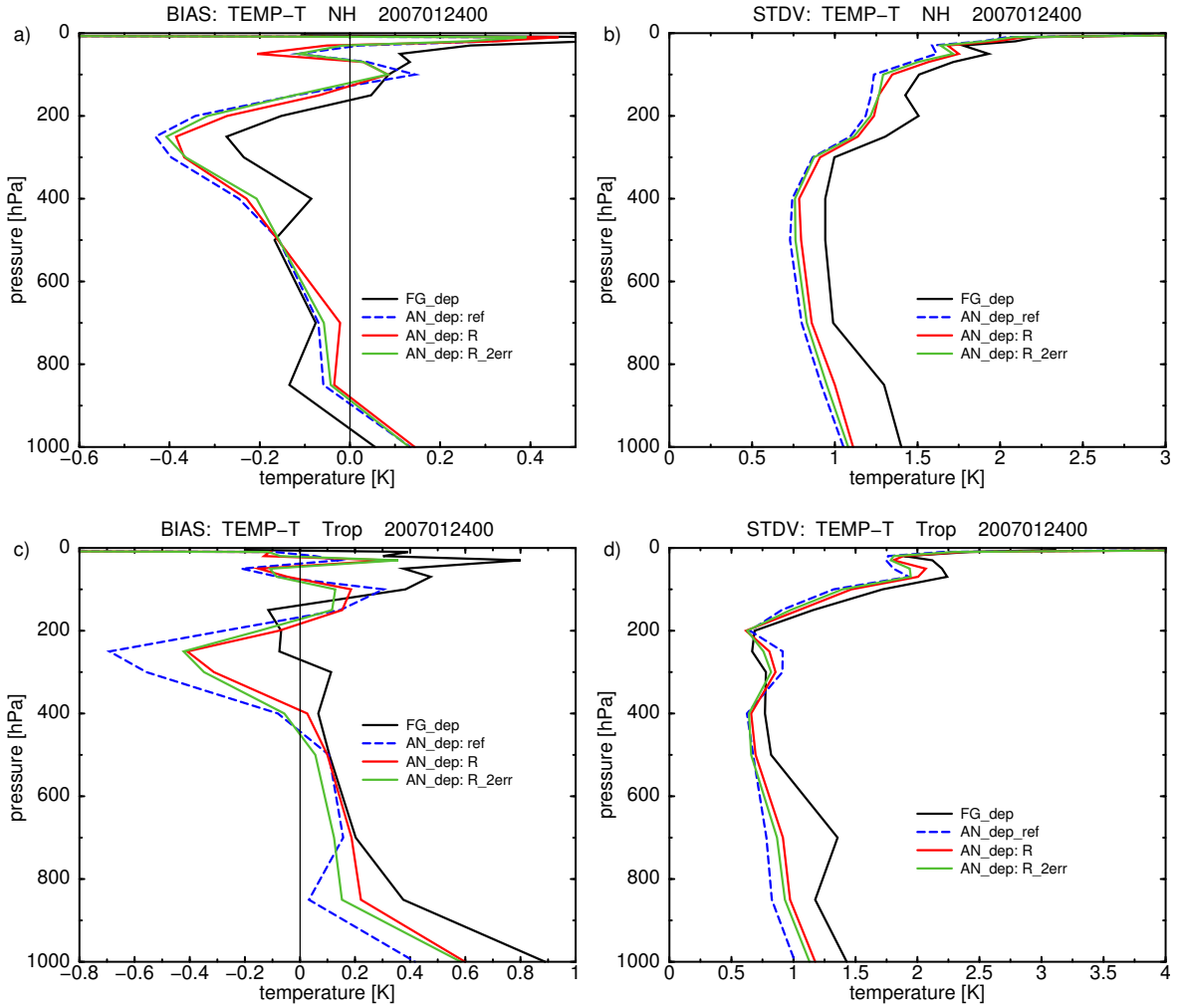

Figure 8.10: Same as Fig. 8.8, but for TEMP observations of temperature over (a,b) the Northern hemisphere and (c,d) Tropics.

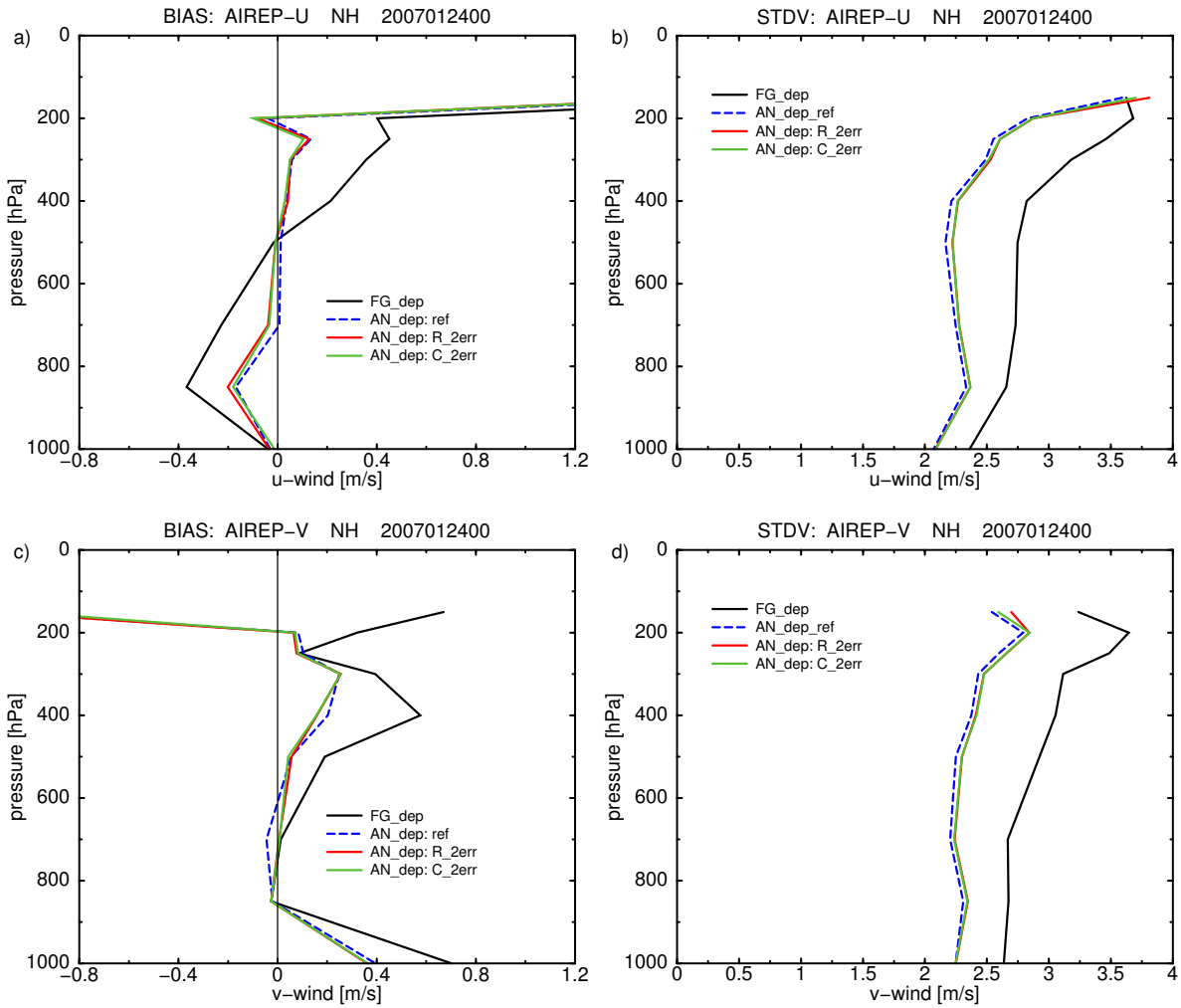

Figure 8.11: Bias (a,c) and standard deviation (b,d) of background departures (black solid line) and analysis departures with respect to AIREP observations of (a,b) zonal and (c,d) meridional wind over the Northern hemisphere for the reference 4D-Var run (blue dashed line) and for 4D-Var experiments assimilating T and q pseudo-observations retrieved from 1D-Var using observations of cloud radar reflectivity either separately (red solid line) or in combination with lidar backscatter (green solid line) applying observation errors twice as large as the computed ones from the 1D-Var analysis covariance matrix. Case of analysis valid on 24 January 2007 00:00 UTC.

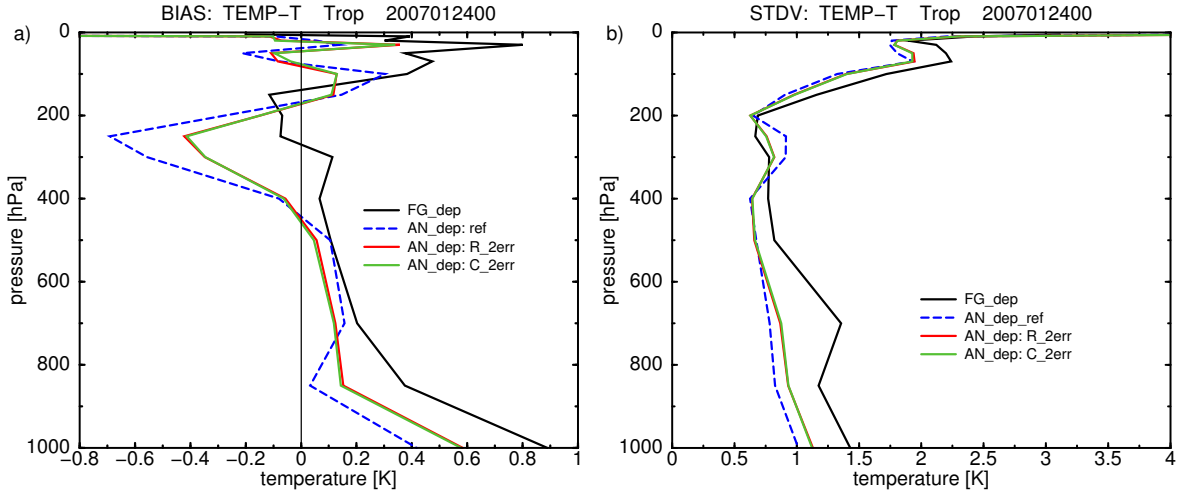

Figure 8.12: Same as Fig. 8.11, but for TEMP observations of temperature in the Tropics.

For all other types of observations assimilated in 4D-Var, no significant changes have appeared when considering observation-minus-background and observation-minus-analysis departure statistics. It should be mentioned that achieving a significant improvement between the experimental runs and the reference run over a domain well covered by a large amount of other measurements is not easy, so any improvement is encouraging since it indicates a potential benefit from assimilating cloud information.

(b) Impact of 1D+4D-Var assimilation on the subsequent forecast

An evaluation of the impact of the 1D+4D-Var system for assimilation of the cloud radar and lidar observations on the subsequent forecast has been done by comparing zonal means of the differences in rms errors for the differences between the forecasts starting from analysis created by 4D-Var assimilation of T and q retrieved from the 1D-Var of either cloud radar reflectivity separately or in combination with the lidar backscatter and the operational analysis, and between the forecast starting from the reference analysis and the operational analysis (i.e. zonal means of *rms(FC_exp - AN_oper) - rms(FC_ref - AN_oper)*, where *exp* indicates experimental run, *ref* reference run and *oper* indicates operational analysis. The comparisons have been done for specific humidity (Fig. 8.13), temperature (Fig. 8.14 and wind (Fig. 8.15). They are presented for 12-, 24- and 48-hour forecasts to show how the signal of reduced , resp. increased, rms errors from the experimental run propagates in time. Generally, rms errors are reduced in the experimental runs compared to the reference run, which is also demonstrated in Table 8.1 providing rms errors over the whole globe.

When comparing the results from the forecast starting from analysis created by 4D-Var assimilating T and q pseudo-observations retrieved from 1D-Var of cloud radar reflectivity alone against those starting from the analysis where both radar and lidar observations were used in 1D-Var, it is possible to see more improvements (i.e. larger reduction of rms errors) at higher altitudes for the experiment with the combined observations. An investigation of the analysis increments in Sections 4 - 6 shows that the lidar increments occur at higher altitudes compared to radar, which indicates that these improvements come from the lidar backscatter due to clouds. Such improvement is the easiest to spot for temperature above 200 hPa (Fig. 8.14), though it also appears below that altitude, for instance between 30° and 60° N. For wind (Fig.8.15), an additional improvement in the combined run is also noticeable between 200 and 400 hPa in the southern hemisphere (30° - 60° S).

Generally, rms errors are mainly reduced in the zonal band between 30° N and 30° S for specific humidity (up to 48-hour forecast) and temperature (up to 24-hour), while error reductions for wind are more scattered in space and time. Even though the positive impact of the new assimilated observations on the subsequent forecast decreases in time, it is still noticeable up to 48-hour forecast. This indicates a potential benefit from the assimilation of cloud information in a 4D-Var system.

	specific humidity			temperature			wind		
	T+12	T+24	T+48	T+12	T+24	T+48	T+12	T+24	T+48
RMS (FC_ref - AN)	0.6392	0.7692	0.8684	1.8305	2.3053	2.8800	5.3078	7.0468	8.6276
RMS (FC_R - AN)	0.6313	0.7627	0.8630	1.8113	2.2975	2.8833	5.2866	7.0325	8.6296
RMS (FC_C - AN)	0.6307	0.7629	0.8631	1.8087	2.2963	2.8832	5.2863	7.0299	8.6271

*Table 8.1: Rms errors for the differences: (**ref**) between the forecast starting from the reference analysis and the operational analysis, (**R**) or (**C**) between the forecast starting from the experimental 4D-Var assimilation using T and q pseudo-observations over the whole 12-hour assimilation window coming from 1D-Var of cloud radar reflectivity separately or in combination with the lidar backscatter and the operational analysis. Statistics is computed over 12-, 24- and 48-hour forecasts over the whole globe.*

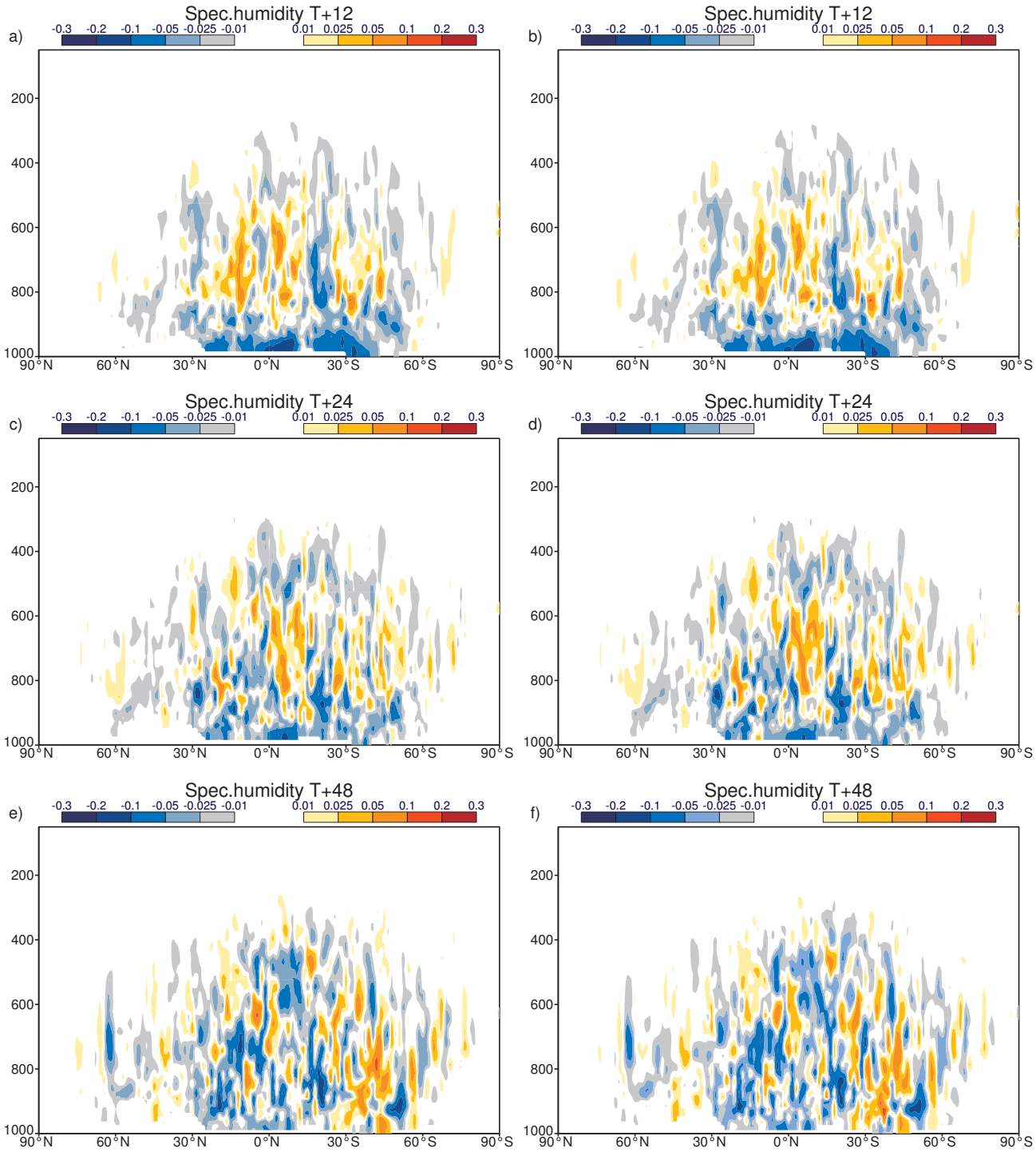

Figure 8.13: Zonal mean of differences of specific humidity rms errors (in g.kg^{-1}) for the differences between the forecasts starting from analysis created by 4D-Var assimilation of T and q pseudo-observations retrieved from 1D-Var of either (a,c,e) cloud radar reflectivity separately or (b,d,f) in combination with lidar backscatter using observation errors twice as large as the computed ones from the 1D-Var analysis covariance matrix and the operational analysis and between the forecast starting from the reference analysis and the operational analysis. (a,b) 12-hour, (c,d) 24-hour and (e,f) 48-hour forecasts. Reduction (resp. increase) of rms errors for the experimental run is shown with blue (resp. red) shadings.

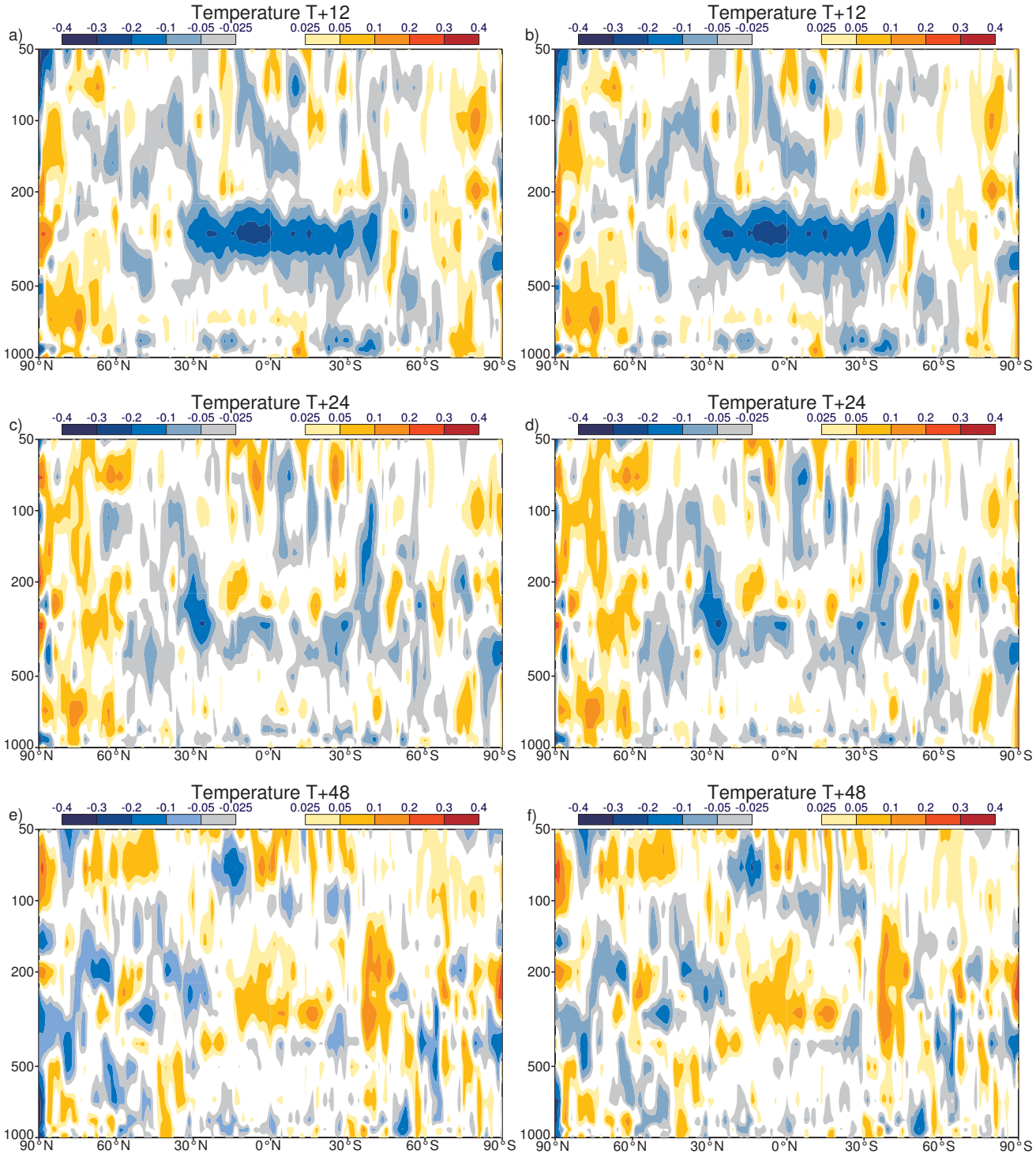

Figure 8.14: Same as Fig. 8.13, but for temperature (in K).

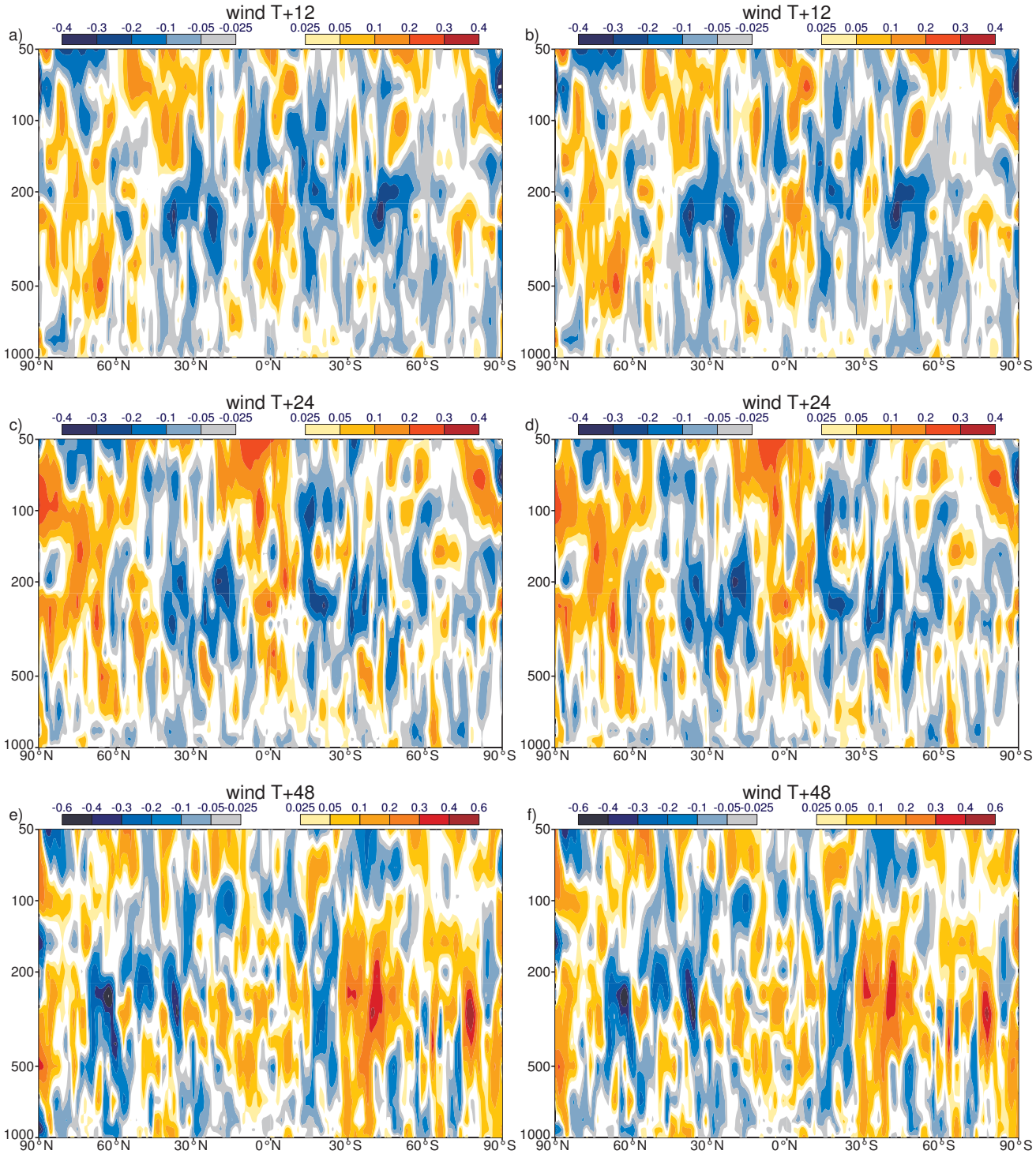

Figure 8.15: Same as Fig. 8.13, but for wind.

9 Conclusions and perspectives

1D-Var assimilation experiments have been performed using observations of cloud radar reflectivity and lidar backscatter, either separately or in combination. The results indicate that the 1D-Var analyses get closer not only to assimilated, but also to independent observations. However, the impact of lidar backscatter due to clouds is smaller than that of cloud radar reflectivity. The performed experiments also indicate how important it is to apply an appropriate quality control, bias correction and error estimate to the used observations for an optimal 1D-Var performance. In the case of 1D-Var with the lidar backscatter, analysis only gets closer to the independent CloudSat radar reflectivity observations when a proper treatment of lidar observations is performed.

Analysis increments of temperature and specific humidity have also been evaluated since they can provide information about the impact of assimilated observations on the control variables of 1D-Var system. The evaluation revealed that both increments are modified by the assimilation of cloud radar reflectivity and/or lidar backscatter. Therefore the pseudo-observations of both temperature and specific humidity profiles from 1D-Var retrievals have been included into the 4D-Var system. Comparing to radar, the lidar increments occur at higher altitudes and are therefore complimentary. At altitudes where both radar and lidar observations are available, the increments are consistent.

Several 1D+4D-Var experiments have been run over one assimilation cycle of 12-hours (i.e. the current length of 4D-Var assimilation window at ECMWF) assimilating T and q pseudo-observations retrieved from the 1D-Var with cloud radar reflectivity and lidar backscatter either separately or in combination. From these analyses, 10 day forecasts have been run to study the impact of these new observations not only on 4D-Var analyses, but also on the subsequent forecasts. Only profiles with non-zero increments of temperature and specific humidity obtained from 1D-Var assimilation have been used in 4D-Var.

1D+4D-Var reduces analysis departures for q pseudo-observations and provides analysis departures closer to T pseudo-observations than would be obtained if temperature pseudo-observations were not used. Verification of the performed assimilation runs has also been carried out against other assimilated observation types. The results indicated that mainly for conventional observations (such as TEMP radiosonde, PILOT or AIREP observations) standard deviations of the analysis departures are systematically larger in the experimental runs compared to the reference run not using pseudo-observations, while bias of these departures is more comparable. This indicates that the new cloud related observations bring more variability into the system which may be related to the fact that clouds often represent small scale features. Small, but systematic improvements coming from the lidar observations when combined with the radar have also been noticed. For all other types of observations assimilated in 4D-Var, no significant changes have appeared when considering observation-minus-background and observation-minus-analysis departure statistics. It is usually not easy to obtain a significant improvement between the experimental runs and the reference run over a domain well covered by a large amount of other measurements, therefore any improvement is encouraging since it indicates a potential benefit from assimilating cloud information.

The impact on the forecasts started from the 1D+4D-Var analyses obtained by assimilating pseudo-observations retrieved from the cloud radar and lidar observations has also been evaluated. Generally, a positive impact of the new observations on the subsequent forecast has been observed. Even though this positive impact decreases in time, it is still noticeable up to 48-hour forecasts.

The performed feasibility studies documented in this report provide some hints on what one could expect from assimilating cloud information from active sensors. The results are encouraging, but real assimilation of such measurements would still require a substantial amount of work to fully benefit from these observations. More experiments and statistical evaluation of the model equivalents to the observations need to be done for different situations to refine data control and error definition usage. The 1D+4D-Var approach requires to define

errors for pseudo-observations retrieved from 1D-Var, which are computed from the 1D-Var analysis covariance matrix. This is quite expensive for profiling observations and only affordable for non-operational application. Therefore using a direct 4D-Var assimilation of cloud related observations should be considered for any future operational implementation.

Acknowledgements

The NASA CloudSat Project is kindly acknowledged for providing the CloudSat data. The authors are also grateful to the NASA Langley Research Center - Atmospheric Science Data Center for making the CALIPSO data available. Thanks are due to Anton Beljaars, Philippe Lopez and Stephen English for helpful advice and review of the document.

List of Acronyms

1D-Var	One-Dimensional Variational Assimilation
4D-Var	Four-Dimensional Variational Assimilation
AIREP	AIRcraft Weather REPort
ARM	Atmospheric Radiation Measurements
ATLID	ATmospheric LIDar
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CloudSat	NASA's cloud radar mission
CPR	Cloud Profiling Radar
CTH	Cloud Top Height
EarthCARE	Earth, Clouds, Aerosols and Radiation Explorer
ECMWF	European Centre for Medium Range Weather Forecasts
ESA	European Space Agency
FG	First Guess
GCM	Global Circulation Model
IFS	Integrated Forecasting System of ECMWF
mae	mean absolute error
NASA	National Aeronautics and Space Administration
NWP	Numerical Weather Prediction
PDF	Probability density function
PSD	Particle Size Distribution
rms	root-mean square error
SMOS	Soil Moisture and Ocean Salinity mission
SNR	Signal-to-Noise ratio
stdv	standard deviation
STSE	Support-to-Science-Element
Z	Radar reflectivity
ZmVar	Z (reflectivity) Model for Variational assimilation of ECMWF

References

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2006a: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1D-Var, *Q. J. R. Meteorol. Soc.*, **132**, 2277–2306.

Bauer, P., P. Lopez, A. Salmond, D.and Benedetti, S. Saarinen, and M. Bonazzola, 2006b: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var, *Q. J. R. Meteorol. Soc.*, **132**, 2307–2332.

Benedetti, A. and M. Janisková, 2004: Advances in cloud assimilation at ECMWF using ARM radar data, *Extended abstract for ICCP, Bologna*.

Benedetti, A., P. Lopez, P. Bauer, and E. Moreau, 2006: Experimental use of TRMM precipitation radar observations in 1D+4D-Var assimilation, *Q. J. R. Meteorol. Soc.*, **131**, 2473–2495.

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.

Di Michele, S., E. Martins, and M. Janisková, 2014a: Observation operator and observation processing for cloud lidar, WP-1200 report for the project Support-to-Science-Element STSE Study - EarthCARE Assimilation, 4000102816/11/NL/CT, ECMWF, 40 pp.

Di Michele, S., E. Martins, and M. Janisková, 2014b: Observation operator and observation processing for cloud radar, WP-1100 report for the project Support-to-Science-Element STSE Study - EarthCARE Assimilation, 4000102816/11/NL/CT, ECMWF, 59 pp.

Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson, 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis, *Q. J. R. Meteorol. Soc.*, **119**, 1427–1463.

Field, P., A. Heymsfield, and A. Bansemer, 2007: Snow size distribution parameterization for midlatitude and tropical ice clouds, *J. Atmos. Sci.*, **64**(12), 4346–4365.

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.

Gérard, E. and R. W. Saunders, 1999: Four-dimensional assimilation of Special Sensor Microwave/Imager total column water vapor in the ECMWF model, *Q. J. R. Meteorol. Soc.*, **125**, 3077–3101.

Illingworth, A. and T. Blackman, 2002: The need to represent raindrop size spectra as normalized gamma distributions for the interpretation of polarization radar observations, *J. Applied Meteor.*, **41**(3), 286–297.

Janisková, M., 2004: Impact of EarthCARE products on Numerical Weather Prediction, *Contract report to the European Space Agency, 59 pp*.

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

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

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

Liu, Z., W. Hunt, M. Vaughan, C. Hostetler, M. McGill, K. Powell, D. Winker, and Y. Hu, 2006: Estimating random errors due to shot noise in backscatter lidar observations, *Applied Optics*, **45**(18), 4437–4447.

Lopez, P. and P. Bauer, 2007: "1D+4D-Var" assimilation of NCEP Stage IV Radar and gauge hourly precipitation data at ECMWF, *Mon. Weather Rev.*, **135**, 2506–2524.

Lopez, P., A. Benedetti, P. Bauer, M. Janisková, and M. Köhler, 2006: Experimental 2D-Var assimilation of ARM cloud and precipitation observations, *Q. J. R. Meteorol. Soc.*, **132**, 1325–1347.

Lopez, P. and E. Moreau, 2005: A convection scheme for data assimilation: Description and initial tests, *Q. J. R. Meteorol. Soc.*, **131**, 409–436.

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

Marécal, V. and J.-F. Mahfouf, 2000: Variational retrieval of temperature and humidity profiles from TRMM precipitation data, *Mon. Weather Rev.*, **128**, 3853–3866.

Marécal, V. and J.-F. Mahfouf, 2002: Four-dimensional variational assimilation of total column water vapour in rainy aeras, *Mon. Weather Rev.*, **130**, 43–58.

Moreau, E., P. Lopez, P. Bauer, A. Tompkins, M. Janisková, and F. Chevallier, 2004: Variational retrieval of temperature and humidity profiles using rain rates versus microwave brightness temperatures, *Q. J. R. Meteorol. Soc.*, **130**, 827–852.

Phalippou, L., 2005: Variational retrieval of humidity profile, wind speed and cloud liquid-water path with the SSM/I: Potential for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, **131**, 409–436.

Rabier, A., F.and McNally, E. Andersson, P. Courtier, P. Unden, J. Eyre, A. Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of the three dimensional variational assimilation (3D-Var). Part II: Structure functions, *Q. J. R. Meteorol. Soc.*, **124**, 1809–1829.

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.

Rodgers, C. D., 2000: Inverse methods for atmospheric sounding. Theory and practice. *Series on atmospheric, oceanic and planetary physics*, Vol. 2, World Scientific, Singapore, New Jersey, London, Hong Kong, 238 pp.

Stiller, O., 2010: A flow-dependent estimate for the sampling error, *J. Geophys. Res.*, **115**(D22).

Sundqvist, H., E. Berge, and J. E. Kristjánsson, 1989: Condensation and cloud paranetrization studies with a mesoscale numerical weather prediction model, *Mon. Weather Rev.*, **117**, 1641–1657.

Tompkins, A. and M. Janisková, 2004: A cloud scheme for data assimilation: Description and initial tests, *Q. J. R. Meteorol. Soc.*, **130**, 2495–2517.