ESA CONTRACT REPORT

Contract Report to the European Space Agency

Development of strategies for radar and lidar data assimilation

March 2010

Authors: M. Janisková and O. Stiller

WP-3100 report for ESA contract 1-5576/07/NL/CB:
Project QuARL - Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles

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 Contract Report

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

Development of strategies for radar and lidar data assimilation

Authors: M. Janisková and O. Stiller

WP-3100 report for ESA contract 1-5576/07/NL/CB:
Project QuARL - Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, Berkshire, UK

March 2010

ABSTRACT

Profiting from the enormous wealth of highly vertically resolved data as they are obtained from spaceborne radar and lidar poses a great challenge for the NWP community which, if successful, could lead to great improvements of our forecast accuracy and modelling capabilities. Particularly the large uncertainty of the vertical distribution of clouds is a longstanding problem of current NWP systems which affects both data assimilation and parametrisation developments. Also the vertical structure of aerosols is notoriously under-constrained by current data assimilation systems which contrasts the great interest which these quantities have recently obtained in the context of air quality monitoring and forecasting. This report explores possibilities to exploit these new data types in the context of an NWP system by including them into the 1D-Var system. This approach gives a good indication to which extent variational data assimilation methods can deal with the new data type and also lays the technical foundations for a future inclusion into the full 4D-Var system. More specifically, 1D-Var retrievals of clouds from CloudSat and aerosols from CALIPSO have been performed. Both assimilations were very successful in producing assimilated states whose model-equivalent observations are substantially closer to the real observations. Cloud retrievals substantially incremented both the humidity and the temperature field while aerosol observations updated the aerosol field considerably. Independent observations confirmed that the cloud retrievals bring the model state closer towards that of the true atmosphere.

Contents

1 Introduction

With the introduction of cloud radar and lidar from space a large volume of information on the vertical structure of clouds and aerosols covering all climate regimes is becoming available. Since the current uncertainty in the representation of clouds in large scale models, particularly related to their vertical structure (see Tompkins *et al.*, 2004, for an overview of the ECMWF system), is substantial, these new observations are very important for model evaluation, further model development and model initialization.

Most observational datasets give predominantly a two-dimensional perspective, from the top of the atmosphere or from the surface through vertically integrated quantities. Thus using observations with the vertical structure information provided by cloud radar and lidar observations from space (such as CloudSat, CALIPSO - Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation, EarthCARE - Earth, Clouds, Aerosol and Radiation Explorer) is a big step forward in cloud/aerosol verification (see Crewell *et al.*, 2004; Illingworth *et al.*, 2007, for an overview of ground based profiling). They provide an invaluable source of data to inspire and validate model developments. These observations are also useful for exploring the potential for an assimilation of cloud profiling data.

A large part of forecasting deficiencies is connected with imperfect assimilation of available observational data in the numerical weather prediction (NWP) process. Accurate initial conditions required by NWP models rely on the quality of the observations and the quality of the assimilation schemes. Despite the major influence of clouds and precipitation on the atmospheric water and energy balance, there is still no explicit analysis of clouds in global data assimilation systems. The cloud contributions to the satellite radiances are (mostly) removed from the assimilation systems. In meso-scale models only, cloud analyses based on nudging techniques have been introduced (e.g. Macpherson *et al.*, 1996; Lipton and Modica, 1999; Bayler *et al.*, 2000). The potential increase of forecast skill through a more accurate treatment of clouds and their vertical structure is therefore enormous.

The same is true also for aerosols whose representation in data assimilation systems is notoriously unsatisfactory. Though being only a relatively recent addition to some standard NWP systems, the interest in the prediction of atmospheric aerosols is huge and has given rise to major, strongly funded national and international research programs like GEMS[1] and MACC[2]. The prospect of better vertically resolved data sets gives exciting perspectives to all efforts of air quality monitoring and forecasting.

For this project, feasibility studies have been performed for the definition of an assimilation system for measurements obtained from nadir-pointing radar/lidar instruments (measurements with small spatial coverage, but high vertical resolution). Using such data poses a problem due to the fact that they may not be representative of the corresponding model grid-box mean values. It would be difficult to start such a study in the framework of the full 4D-Var (four-dimensional variational) assimilation system (operationally used at ECMWF) which is very complex and thus quite difficult to interpret. In the past, it was proven that the 1D-Var (one-dimensional variational) 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 1D-Var assimilation studies using ARM (Atmospheric Radiation Measurement) radar data (Benedetti and Janisková, 2004; Janisková, 2004; Lopez *et al.*, 2006) can also be beneficial. Therefore 1D-Var systems for the assimilation of different cloud-related observations from CloudSat and aerosol observations from CALIPSO have been developed and are presented in this report.

Section 2 describes the general methodology of 1D-Var system and provides information about the observation operators (moist parametrization schemes, reflectivity model, cloud optical depth parametrization and

[1] http://data-portal.ecmwf.int/data/d/gems_nrealtime/
[2] http://www.gmes-atmosphere.eu/

lidar backscatter related to aerosols). Details of the setup for assimilation experiments are given in section 3. Section 4 presents results from the cloud assimilation experiments. Results from the aerosol assimilation experiments are presented in section 5.

2 Methodology

2.1 Description of the 1D-Var system

The principle of 1D-Var is similar to that of 4D-Var (Courtier *et al.*, 1994), but the control vector $\mathbf{x}$ represents a single column only and the time dimension is not included. The goal of 1D-Var is to search for the optimal model state $\mathbf{x}_a$, that simultaneously minimizes the distance to the observations $\mathbf{y}^o$, and to a background model state $\mathbf{x}^b$ (i.e., a short-range forecast valid at the time of assimilation). The model state $\mathbf{x}$ consists of the vertical profiles of temperature and specific humidity, which are the control variables of the system. The following cost function is minimized during the assimilation process:

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

where $\mathbf{B}$ is the covariance matrix of background errors taken from the operational ECMWF 4D-Var system. Its role is to provide an appropriate information about the statistical structure of the forecast error. $\mathbf{R}$ is the observation and representativeness error covariance matrix. H is the observation operator which provides the model counterpart to the observations. It can employ physical parametrization schemes, such as moist physics and radar/lidar forward operators, depending on which type of observations will be assimilated.

The minimization requires an estimation of the gradient of the cost function:

$$\nabla\mathscr{J}(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b) + \mathbf{H}^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}^o) \tag{2.2}$$

A simple diagram providing the schematic description of 1D-Var assimilation of CloudSat/ CALIPSO observations is given in Fig. 2.1. For estimating the minimum of $\mathscr{J}$, variational analysis systems use descent algorithms (in our system a limited memory quasi-Newton optimization routine M1QN3 is employed) which updates the model state at each iteration (see bottom lines of Fig. 2.1). As illustrated in the third line from below of Fig. 2.1, one input to such descent algorithms is the gradient of the cost function (Gilbert and Lemaréchal, 1989) which comprises a background and an observational term (corresponding to the first and second terms on the right hand side of Eq.2.2, respectively). The observational part is computed through the adjoint operator ($\mathbf{H}^T$) applied to the innovations (i.e., the difference between the observations and and their model equivalent) normalized by the observation error matrix as input. The background term contains the difference of the model state from the first guess, normalized by the model error statistics.

It should be noted that the variational analysis framework is extremely flexible and can be applied to very different observation types including retrieved cloud and aerosol parameters (in the case of level-2 products) or backscatter cross-sections and reflectivities (in the case of level-1 products). While the basic structure of the 1D-Var system is the same for all observation types, some parts have to be particularly developed or adjusted with regard to the respective observations which are being assimilated. In Fig.2.1, boxes which require major observation-specific developments are framed in blue while green boxes indicate observation specific input data to the 1D-Var system. Most of the effort when adjusting a 1D-Var system to a new observation type is generally related to the development of the observation operator and its adjoint (blue framed boxes in Fig. 2.1).

In principle, the adjoint operator computes adjoint sensitivities for all the model parameters which the forward operator requires as input. In theory all these fields can therefore be updated during the minimisation process. In practice, however, some adjoint sensitivities may be much stronger than others or some model fields may have a much larger uncertainty than other fields so that updating them together in the same minimisation cycle would not be of much benefit for the better constrained (i.e., less uncertain) fields.

The observation operators H used in this report are described in section 2.2. For cloud assimilation from CloudSat the main inputs to the observation operator are the background fields of temperature and humidity. By using the 1D-Var technique, these fields are updated and analyzed values of temperature and humidity are obtained. For aerosol assimilation, the observation operator also depends on temperature and humidity as well as various other fields. The only fields which have been chosen to be updated through the 1D-Var technique are however aerosol mixing ratios so that only for these quantities analyzed values are obtained.

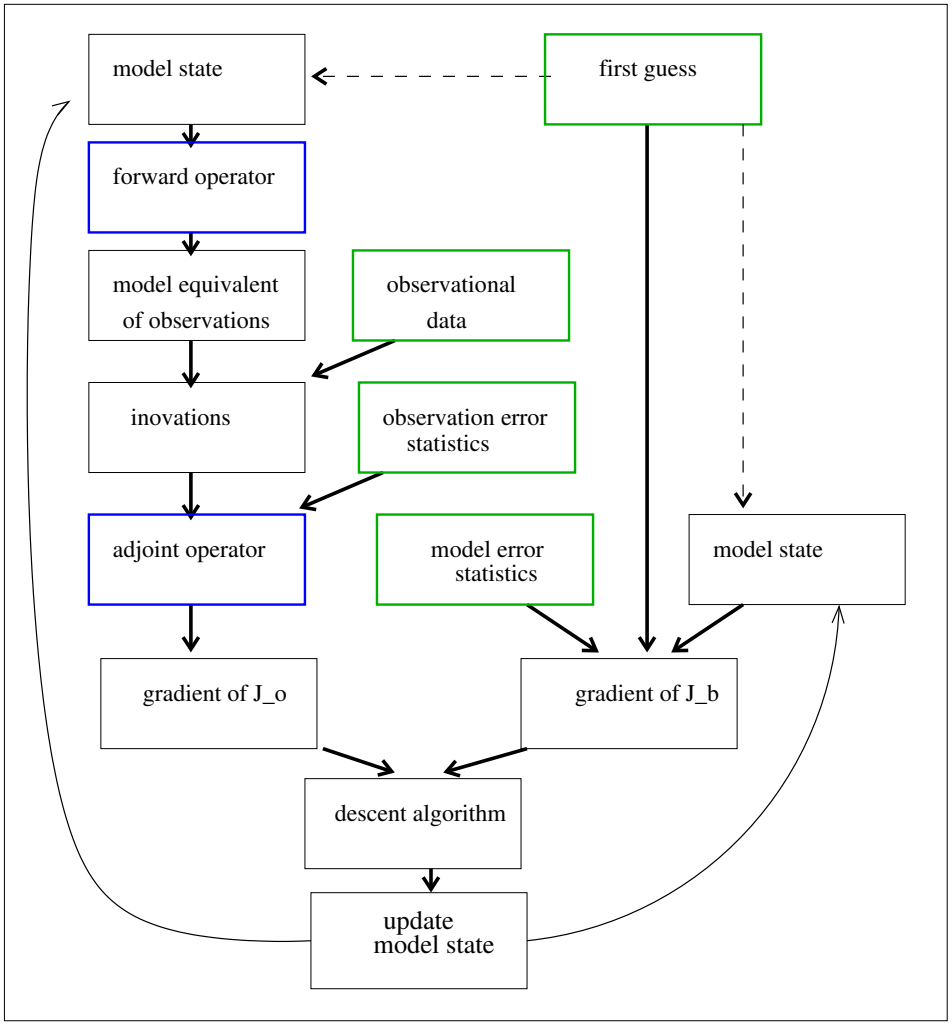


Figure 2.1: *Flow diagram of the functioning of the 1D-Var system.*

2.2 Observation operator

In the assimilation of cloud properties, the operator H consists of two simplified parametrizations of moist atmospheric processes: a convection scheme (Lopez and Moreau, 2005) and a cloud scheme simulating large-scale condensation and precipitation processes (Tompkins and Janisková, 2004). The assimilation of cloud radar reflectivity observations further requires a radar forward model to convert model fields into reflectivities. Input cloud and precipitation fields to the radar forward model are computed by the moist physics parametrisations whose main input variables are temperature and humidity fields.

For aerosol assimilation, the operator H consists of a single forward operator that simulates lidar backscattering. No moist parametrisation scheme is needed as aerosol assimilation can only be performed in cloud

free regions. For the cloud screening, observational data from CALIPSO level-2 products are used.

As explained above, an efficient computation of the gradient of the cost function during the minimization requires the linearized (tangent-linear and adjoint) versions of the observation operators. The development and testing of the tangent-linear and adjoint versions of the new observation operators (radar and lidar) was an important part of this project.

2.2.1 Cloud scheme

This scheme (Tompkins and Janisková, 2004) diagnoses cloud cover and cloud water (liquid+ice) from the input profiles of the temperature and humidity control variables. It describes four processes which are the generation of stratiform cloud through large-scale motions, convective generation of cloud, precipitation generation and evaporation. These were shown to dominate the sensitivity of the Tiedke's prognostic scheme (Fillion and Mahfouf, 2003). Therefore, a simplified scheme based on this reduced set of processes should be able to reproduce most of the sensitivity of the full complex prognostic scheme (Tiedtke, 1993).

(a) Stratiform cloud properties
The scheme is based on statistical ideas assuming subgrid scale fluctuations of total water and/or temperature allowing the in-cloud liquid water and cloud cover to be determined by integrating the saturated portion of the grid box. The uniform distribution is used to describe a probability density function (PDF) of subgrid-scale fluctuations for simplicity and for consistency with the full prognostic cloud scheme, which uses the same assumption for the clear-sky vapour fluctuations (Tiedtke, 1993). Currently the variance of the distribution is fixed, implying that cloud cover and liquid water can be reduced to formulations based on relative humidity. In order to account for the asymmetric nature of the cloud formation process, the distribution width is a function of relative humidity (RH) and is reduced linearly from its maximum value at a critical relative humidity RH_c (at which cloud formation sets in) to its minimum value when $RH = 1$.

(b) Convective contribution
One of the most significant sources of cloud in the tropics and mid-latitude summers in the Tiedtke (1993) cloud scheme is the detrainment from deep convection (e.g. Teixeira, 2001). This information comes from a convection scheme which has been developed particularly for the requirements of data assimilation (Lopez and Moreau, 2005). For cloud cover this scheme adopts assumption from Tiedtke (1993) that convective clouds randomly overlap with existing stratiform clouds. Unlike in the prognostic scheme used in the forecast model, no memory exists for cloud water between consecutive timesteps. Thus the additional assumption is made that any convectively detrained cloud water that is not converted to precipitation during the timestep, evaporates. This is necessary to prevent convectively active regions from artificially drying during the forward integration.

(c) Precipitation production
Simple auto-conversion terms are used for the generation of rain and snow (based on Sundqvist *et al.*, 1989), which fall out of the model column during one timestep. The Bergeron-Findeisen or collection processes are not explicitly represented in this cloud scheme.

(d) Rainfall evaporation
As with the Tiedtke (1993) scheme, precipitation falls to the surface within one time-step and evaporates during its descent. This process is based on the Newtonian relaxation formulation of Kessler (1969). For simplicity, the diagnostic scheme uses a maximum overlap assumption for the precipitation fraction (the proportion of a grid cell through which precipitation is falling). Precipitation evaporation during the descent partially accounts for the overlap of precipitation with the subgrid clear sky distribution of humidity fluctuations in the scheme. This removes the requirement of artificial and discrete process thresholds, which could have a detrimental effect on minimization performance in the assimilation system.

2.2.2 *Convection scheme*

In the simplified convection scheme (Lopez and Moreau, 2005), all types of convection (shallow, mid-level, and deep) are included. In particular, the link between the model control variables and the cloud base mass flux (the so-called closure assumption) is based on the release of convective available potential energy (CAPE) in time for deep and mid-level convection and on low-level moisture convergence for shallow convection. In contrast with the operational convection scheme used in the forecast model, the equations that describe the vertical evolution of the updraught mass flux and of the updraught thermodynamic characteristics (dry static energy and total water), are uncoupled. The uncoupling allows the removal of the iterative calculations involved in the operational code for updating the cloud base mass flux, thereby leading to an easier development of the adjoint code.

Convection is assumed to be activated only if the bulk convective updraught vertical velocity remains positive at cloud base. The updraught is assumed to originate from the surface only if its initial vertical velocity is positive, which is calculated from the surface heat fluxes using the formulation of Holtslag and Moeng (1991). The initial temperature and specific humidity departures of the updraught from the environment are assumed to be proportional to the surface sensible and latent heat fluxes. If convection cannot be initiated from the surface, the convective ascent may originate from higher levels provided relative humidity exceeds 80%, in which case the initial vertical velocity of the updraught is somewhat arbitrarily set to 1 m s^{-1}. The vertical evolution of its kinetic energy is computed following Simpson and Wiggert (1969), which involves the buoyancy as well as the entrainment of environmental air into the updraught. The convective ascent is assumed to stop at the level where updraught vertical velocity becomes negative.

Simplified calculations of downdraughts and convective momentum transport based on the operational scheme (Tiedtke, 1989) are also included in the new parametrization. More importantly, the precipitation formation rate is made proportional to the updraught cloud water content as in Tiedtke (1989).

2.2.3 *Reflectivity model*

The reflectivity model for variational assimilation (ZmVAR) is designed to meet the requirements of an assimilation system, i.e. to be computationally efficient and to allow the coding of its adjoint counterpart. A detail description of ZmVAR can be found in the WP-1000 report. 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 particles density is a function of their diameter. Hydrometeorological size distributions are exponential for rain and snow. A Gamma distribution is used for cloud ice, while cloud liquid water follows a lognormal distribution. 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.

Since ZmVar was originally developed in WP-1000 (Forward operator developments) additional testing of this operator has been performed which led to the definition of a modified Mie-table displayed in Fig. 2.2. This table improves relationships between cloud liquid and reflectivity especially.

The adjoint of the reflectivity model has been coded for this study to make 1D-Var assimilation experiments computationally more efficient. The correctness of the adjoint code has been tested.

2.2.4 *Cloud optical depth*

The model optical depths are calculated at 0.55 μm using the observation operator described in detail in Benedetti and Janisková (2008). The routine for cloud optical depth uses the Slingo (1989) parametrization

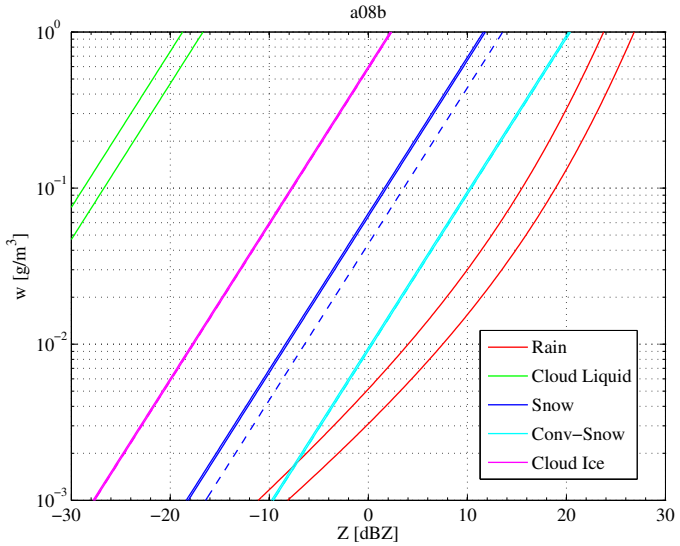


Figure 2.2: *Hydrometeor content - reflectivity Z relationships for each of the hydrometeors involved (modified Mie-table for ZmVAR operator). Double lines for cloud liquid and cloud ice used: one for atmospheric temperature set to the minimum value in the Mie-table (lower Z) and one for temperature set to maximum (for ice - min=-70°C, max=0°C and for liquid - min= -40°C, max=30°C). The dashed line for snow denotes the melting case. Conv-Snow stands for convective snow.*

for the optical properties of liquid water clouds and the Fu (1996) formulation for ice clouds. For liquid water clouds, the effective radius (r_e) is derived from the cloud liquid water content following Martin *et al.* (1994), with the concentration of cloud condensation nuclei fixed at 50 cm^{-3} over oceans and 900 cm^{-3} over continents. For ice clouds, the effective size of particles is a function of temperature following Ou and Liou (1995).

2.2.5 *Aerosol backscatter*

The lidar forward operator is based on a new more general and flexible software which has been recently developed and included within the ECMWF IFS (Morcrette *et al.*, 2009). This code is quite flexible and can simulate lidar signals at the wavelengths commonly used by lidar systems (355, 532 and 1064 nm) and that is applicable to both, to ground based (as from EARLINET) and space borne (as from CALIPSO) measurements. In the work presented here, the operator is only used for the CALIPSO 532 nm channel.

The operator uses extinction-to-backscatter ratios (Ackermann, 1998) for the eleven prognostic aerosol variables included in the GEMS/MACC version of the ECMWF IFS (Morcrette *et al.*, 2009) which have been computed with a standard Mie code. The lidar signal (*LS*) is then simulated according to Huneeus and Boucher (2007), from the mass mixing ratios of these eleven prognostic aerosols (namely, three bins of sea salt, three bins of dust, organic and black carbon, both in their hydrophilic and hydrophobic types, and sulphate) following the original formulation of Klett (1985)

$$LS(z,\lambda) = C_{LS}\,[\beta(\lambda)\omega(\lambda)\sigma(z,\lambda)]\;exp(-2\int_{z}^{TOA}\sigma(z',\lambda)\,dz') \tag{2.3}$$

where λ is the wavelength, z the height, $\beta(\lambda)$ the phase function at 180^o (proportional to the backscattering coefficient), $\omega(\lambda)$ the single-scattering albedo, and $\sigma(z,\lambda)$ the extinction coefficient. C_{LS} is a calibration

constant set to one. In the equation above, account is also taken of Rayleigh scattering and potential gaseous absorption by O_3, NO_2, CO_2 and O_2 in the computation of $\sigma(z,\lambda)$.

In this project, we employ this operator only for the 532nm chanel which is not sensitive to CO_2 absorption. Since the impact of variations of NO_2 absorption on this study is expected to be comparably weak, the NO_2 field has been approximated by a constant value (i.e., $NO_2 = 5.10^{-6}$) for this feasibility study.

Having adjoint versions of all the above described observation operators used by the 1D-Var system reduces computational cost significantly (approximately by 60 times) compared to the finite-difference approach, thus allowing to run more experiments in order to define the most suitable approach for the assimilation of new type observations.

2.3 Background error statistics

The background error covariance matrix **B** provides the appropriate information about the 1D (vertical) statistical structure of the forecast errors to the variational analysis. For cloud assimilation, the covariance matrix of the background errors 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 up to around 800 hPa, then slowly decreases 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 the local 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 a vertical profile of standard deviations for temperature and humidity is illustrated on Fig. 2.3.

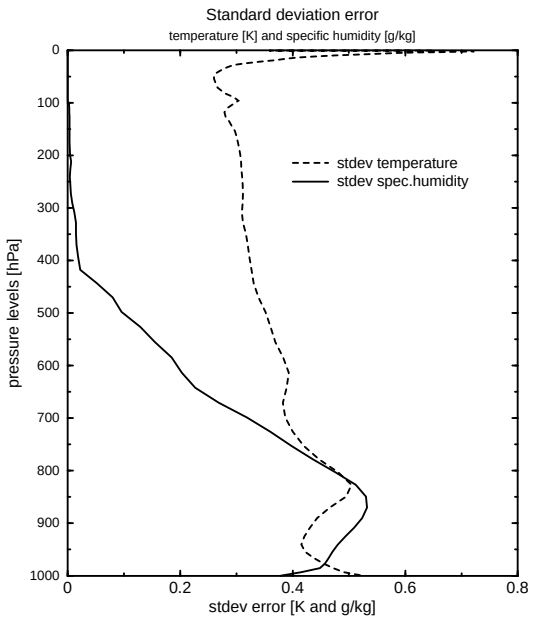


Figure 2.3: *Vertical profiles of typical values of the standard deviation of the ECMWF model background errors for temperature (dashed line) and specific humidity (solid line). Units are in K and g kg^{-1} respectively.*

For aerosol assimilation, background error statistics have been generated from GEMS near real-time experiments via the NMC method (Benedetti and Fisher, 2007). These GEMS experiments are performed with a

60 level system. To apply them to the 91 level configuration used in this study a simple linear interpolation has been applied. In the troposphere the employed aerosol standard deviations decreases roughly exponentially with height from about 6.10^{-8} in the boundary layer to 2.10^{-9} near the tropopause while substantially smaller values ($\approx 10^{-11}$) are used in most of the stratosphere.

3 Experimental framework

1D-Var experiments have been carried out first using the cloud radar data only and were then combined with column-integrated measurements in order to define an assimilation system for measurements obtained from nadir-pointing radar instruments (measurements with small spatial coverage, but high vertical resolution). Using these profile measurements may pose a problem due to the fact that they may not be representative of the corresponding model grid-box values. The aim of this study has been therefore to study approaches how to use such measurements for assimilation. Another purpose has been to show the ability of 1D-Var to modify dynamical variables, namely temperature and specific humidity, since an important aspect of the cloud assimilation is to achieve a consistency between cloud parameters and dynamics. Experiments have shown that when dynamical fields are inconsistent with cloud profiles, the analyzed cloud information is lost within a few model time steps.

3.1 Background values

For cloud assimilations, the background values have been taken from a 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 the same way as in the operational 4D-Var system where all observations are split to 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 above) to compute cloud properties (cloud cover, ice and liquid-water contents) and precipitation fluxes. When assimilating reflectivities, a radar observation operator is then applied to the model fields to obtain the equivalent model reflectivity.

For aerosol assimilation, the employed background data consist of the following fields: surface pressure, temperature, specific humidity, ozone and the 11 aerosol species which are used in the GEMS system. These fields have been taken from a 12-hour forecast of the ECMWF model with T511 spectral truncation (corresponding to approximately 40 km). Initial conditions were taken from the end of a 1 month trial run with free wheeling aerosols as described in (Morcrette *et al.*, 2009). In this trial run, all model fields apart from aerosols were updated 12 hourly by the corresponding operational ECMWF analysis. Aerosols which are passive tracers in the ECMWF system were spun up from zero (initial conditions at day 1) through the physical representation of sources and sinks. These runs did not involve assimilation of aerosol data.

3.2 Observations

Measurements of cloud radar reflectivity, converted to $mm^6\ m^{-3}$ (level-1 product), or the cloud liquid and ice water contents in $kg\ m^{-3}$ (level-2 products) from the CloudSat 94 GHz radar have been used in these feasibility studies in order to determine which of the two approaches performs best. The above profile observations have also been combined with column-integrated measurements (namely cloud optical depths from MODIS - Moderate Resolution Imaging Spectroradiometer) with larger spatial coverage as a possible

way to eliminate the ambiguity and representativeness problem when using measurements with small spatial coverage of CloudSat. As defined and explained by Benedetti and Janisková (2008), a logarithmic cloud optical depth has been used as quantity to be assimilated in order to optimize the optical depth departure distribution (more Gaussian distribution of observation-model departures).

Experiments have been performed using either observations closest to the model grid-point or averaged over the model grid-box.

To validate the 1D-Var performance, MODIS cloud optical depths (at the standard reference wavelength of 0.55 μm) have been used as independent observations for the cases when they were not assimilated. For the situations over the USA, the NEXRAD (NEXt-generation RADar) precipitation data have been used for validation in addition. These data come from NCEP stage IV radar and gauge precipitation analysis with 4-km resolution and hourly accumulation (Baldwin and Mitchell, 1996; Fulton *et al.*, 1998).

For aerosols, the CALIPSO level 1 product "Total Attenuated Backscatter" at 532 nm has been assimilated. The preprocessing of these data comprised of two steps:

- Cloud screening.
- Averaging all data within the corresponding forecast model grid box.

Cloud screening was performed using the CALIPSO level 2 cloud layer product whereby all data were rejected below the highest cloud top. As explained below, screening with the 1 km cloud layer product appeared to be insufficient and therefore the more restrictive 5 km product has been employed for the results presented in this report.

3.3 Observation errors

The impact of any type of observations in data assimilation is partly determined by the errors that are assigned to them. These errors should take into account not only instrumental error, but also representativeness error.

As a first approach to error definition, the error on observed reflectivities and backscatter has been fixed to 25% of the observed values at all levels.

Observation errors for the derived cloud liquid and ice water contents have been used as percentage from the contents themselves as defined in the CloudSat level-2 products. These errors could be quite often as high as 75%, especially for the ice water content, thereby giving a small weight to these observations in 1D-Var system.

The cloud optical depth errors have been used as defined in the products, but converted to logarithmic space according to the formula

$$r_{\log} = \sqrt{\log_{10}(1+\varepsilon_r^2)} \tag{3.1}$$

where $r_{\log}$ is the error variance of the logarithmic optical depth and ε_r is the relative error of physical optical depth.

In order to account in some way for the representativeness problem related to the small spatial coverage of CloudSat measurements, the errors for these observations have been increased depending on the CloudSat scene variability product (based on the MODIS data) as follows:

- highly uniform scene - no increase,
- uniform scene - 10% increase,

- weakly variable scene - 25% increase,
- variable scene - 50% increase,
- highly variable scene - 80% increase.

In addition, observations have been blacklisted when:

- data quality flag is not zero,
- cloud mask indicates likely bad data or ground clutter,
- reflectivity is smaller than -40 dBZ or larger than 50 dBZ,
- liquid/ice water uncertainty is larger than 250%.

4 1D-Var experiments for CloudSat observations

Several 1D-Var experiments have been run for the different selected meteorological situations using the observations mentioned in section 3.2.

4.1 Experimental setup

The observations used in the experiments have been either **matched** with the model data (using observations closest to the model grid-point) or **averaged** over the model grid-box.

The experiments using the following combinations of observations have been done:

- **refl** - cloud radar reflectivity,
- **reflopt** - cloud radar reflectivity combined with cloud optical depth,
- **liwc** - cloud liquid and ice water contents,
- **liwcopt** - cloud liquid and ice water contents combined with cloud optical depth,
- **opt** - cloud optical depth (only for situation on 23 January 2007),
- **liwc4err** - cloud liquid and ice water contents with 4-times smaller observation error for ice water content (only for situation on 23 January 2007).

4.2 Situations used in experimentation

A variety of situations has been selected for 1D-Var experimentations. The situations with a good coverage of both the profiling information from the cloud radar and the column-integrated measurements represented by the cloud optical depth from MODIS have been only used for this study. The selected situations (displayed in Fig. 4.1) are:

- **20070123** - the track between 23:50 UTC on 23 January 2007 and 0:26 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 or extra-tropical cyclone in the north),
- **20070603** - the track between 7:56 and 8:03 UTC on 3 June 2007 crossing a large system south-west of Australia,
- **20070913** - the track between 7:31 and 7:41 UTC on 13 September 2007 passing over tropical convection over the Indian Ocean,
- **20070916** - the track between 4:48 and 4:54 UTC on 16 September 2007 crossing the typhoon Wipha in the west Pacific,

- **20080207** - the track between 3:37 and 3:47 UTC on 7 February 2008 passing over a frontal system in the Atlantic Ocean,
- **20080424** - the track between 19:13 and 19:23 UTC on 24 April 2008 over a cloud system (with precipitation in its middle) in USA,
- **20080428** - the track between 7:45 and 7:52 UTC on 28 April 20088 over the eastern part of the USA.

4.3 Results

4.3.1 Situation on 23 January 2007 over the Pacific Ocean

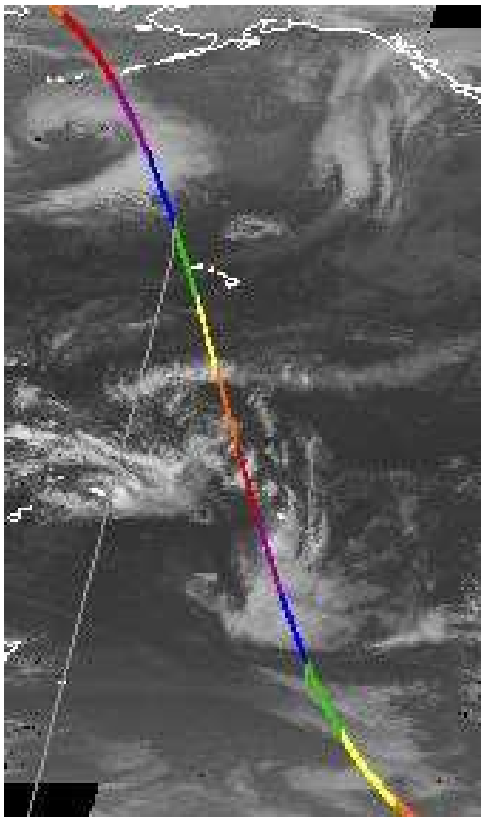
The satellite track for this situation is displayed in Fig. 4.1 and passes over 705 model grid points (for the model with T799 resolution). This longest track used in our experimentation crosses a large extra-tropical cyclone in the north; tropical convection between 10°N and 20°S; a large, vertically 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.

(a) Comparisons of the first-guess and analysis against assimilated observations

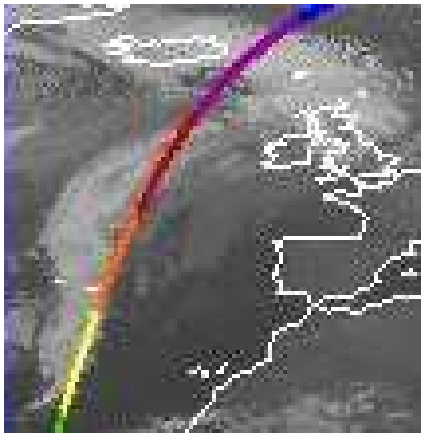
Figure 4.3.1 ((a) for observations matched with the model data, (b) for observations averaged over the model grid-box) shows the cloud radar reflectivity derived from the CloudSat 94 GHz cloud radar along the track between 23:50 UTC on 23 January 2007 and 00:26 UTC on 24 January 2007. The model equivalent to the observations (model first-guess), which is displayed in 4.3.1c, shows an ability of the model to produce cloud large-scale structures similar to the observations. The model clouds appear mostly at the right locations, even though there are differences in their vertical structure with respect to the observations. The cloud liquid water and ice contents derived from the cloud radar observations (level-2 products) and their model equivalents are shown in Fig. 4.3 and 4.3.1. One can see that there are not a lot of differences between matched and averaged observations in the case of cloud radar reflectivity and cloud ice water content. The averaged observations are naturally smoother, which should provide smoother model-observation departures and, therefore be profitable for an assimilation system. Averaging also brings more information about the liquid water content where observations (as provided by the CloudSat level-2 product) are often missing and therefore very sparse coverage is obtained when these observations are matched with the model data at approximately 25-km resolution.

Results from the different assimilation experiments (described in subsection 4.1) are provided in Fig. 4.5 - 4.15 and Tables 4.1 - 4.2.

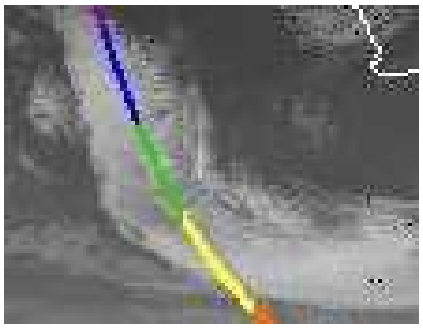
Figures 4.5 - 4.7 show how the analysed cloud radar reflectivity is changed by assimilating different types of observations averaged over the model grid-box. When the cloud radar reflectivity is assimilated, the 1D-Var analysis is closer to the observations for most profiles. However, one can notice that convective clouds between 8°N and 8°S still remain closer to the first guess values than to observations. Assimilation of MODIS cloud optical depth in combination with cloud radar reflectivity (Fig. 4.3.1a) brings small additional adjustments of the analysis with respect to cloud reflectivity, mainly in the lower part of the atmosphere. When assimilating level-2 products (cloud liquid and ice water contents) in 1D-Var (Fig. 4.3.1b), the analysis over the tropical area (20°N and 20°S), as well as around 55°N and between 45°S and 50°S gets closer to the observed values of the cloud radar reflectivity than in the case of cloud reflectivity assimilation. Those areas are linked to the better coverage of the cloud liquid water content product (see Fig. 4.3). However, the adjustment in the levels with the cloud ice water content is quite small and the analysis remains close to the first-guess values above approximately 8 km, especially in the extra-tropics. Combining the level-2 products with cloud optical depth (Fig. 4.3.1c) modifies the analyzed values of cloud reflectivity in the lower part of



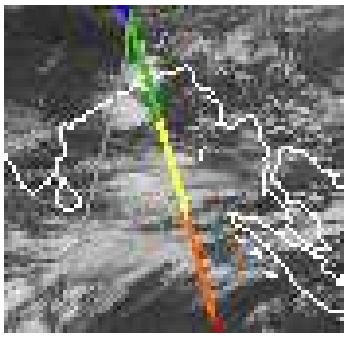
20070123 over Pacific



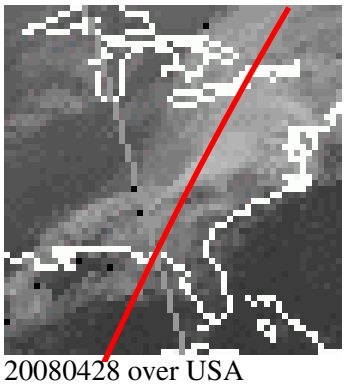
20080207 over Atlantic



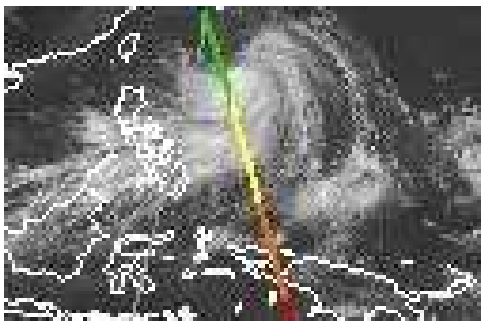
20070603 SW of Australia



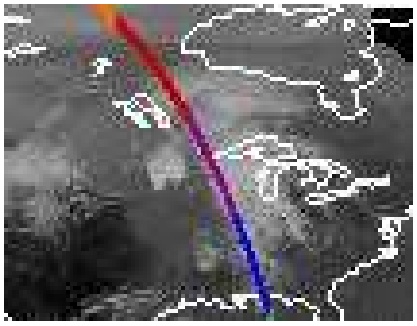
20070913 Indian Ocean



20080428 over USA



20070916 Wipha cyclone



20080424 over USA

Figure 4.1: *Overview of the situations used in 1D-Var experimentation for cloud related observations.*

the atmosphere which seems sometimes to bring analysis back to the first-guess values (as for the systems around 50°S or between 60 and 50°N) or even to drive the analysis further away from the observations compared to the first-guess (e.g. around 30°S). This type of behaviour may indicate discrepancies between the CloudSat level-2 and the MODIS optical depth products (i.e. one indicating that the cloud should have smaller liquid/ice water contents and the other that they should be optically thicker or vice-versa). However, there are also some improvements (e.g. between 10°and 20°S) in the analyzed cloud reflectivity when combining cloud liquid and ice water contents with optical depth.

Generally, the impact of cloud optical depth on the analysis is larger when combined with the level-2 products than with the cloud radar reflectivity. There can be several reasons for this behaviour, such as missing cloud liquid water content data and/or larger observation errors for liquid and ice water contents than for cloud optical depth. The errors of cloud reflectivity and optical depth may be more balanced and also observational coverage for these quantities is better.

The impact on cloud ice water content coming from the assimilation of the different types of observations is displayed in Fig. 4.3.1 - 4.9. Even though the analysis gets closer to the CloudSat cloud ice water product, the adjustments seem to be smaller than for the reflectivity field. Larger modifications are observed when assimilating cloud radar reflectivity, but those modifications are mainly towards a reduction of the model ice water content when it was larger than the observed values.

Results from the assimilation experiments using either matched or averaged observations are also summarized in Table 4.1 in terms of bias, standard deviation (stdv), as well as mean absolute error (mae) and root-mean square error (rms). The last two statistical measures are only used for cloud reflectivity which can have both, the positive and the negative values in dBZ. The mean absolute error expresses better how far the model data (either first guess or analysis) are from the observations. In the case of first-guess (FG) departures (differences between observations and the model first guess equivalents), the model is closer to the averaged observations of cloud ice water content than to the matched ones. FG departures for cloud reflectivity are larger for the averaged observations than for the matched ones (in terms of bias, mae and rms). The analysis (AN) departures (differences between observations and analysis) are usually smaller when assimilating observations which have been averaged over the model grid-box. This indicates that the 1D-Var system favours using smoother observations provided by averaging. Table 4.1 also shows that the analysis is closer to the cloud reflectivity observations for all type of assimilation experiments as indicated by all used statistical measures. Of course, the best adjustment is achieved when assimilating directly cloud reflectivity. The analysis obtained from the assimilation of cloud optical depth in combination with cloud reflectivity differs only slightly from the assimilation of cloud reflectivity alone. The bias of the analysis departures for the cloud ice water content is always larger than for the first guess departures, even though the standard deviation is decreased in the analysis. This can be created by a more efficient reduction of the negative departures (i.e. when the observed contents are smaller than the model first guess ones) while not modifying significantly positive ones. Thus the 1D-Var system is more efficient in removing hydrometeors than in increasing their amount.

(b) Analysis increments

Before starting this project and based on our experience with the former operational assimilation system for precipitation-related observations based on the 1D-Var+4D-Var approach (Bauer *et al.*, 2006a,b), it was believed that the model specific humidity field will mainly be modified by the assimilation of cloud-related observations. That would allow us to produce a pseudo-observation of total column water vapour from the 1D-Var retrievals which would be then assimilated in the ECMWF 4D-Var system. Because of this and also to see an impact of the assimilated observations on the control variables of 1D-Var system, analysis increments of temperature and specific humidity have been checked. They are displayed in Fig. 4.11 - 4.3.1. From these figures one can immediately see that the temperature increments cannot be ignored since they

are not negligible compared to the specific humidity increments as it used to be the case for the precipitation assimilation. That means that we will need to include pseudo-observations of temperature and specific humidity profiles from 1D-Var retrievals rather than the total column water vapour into the 4D-Var system.

Overall, analysis increments for both temperature and specific humidity do not seem unrealistic, though there are quite large differences depending on the assimilated type of observations. Similarly as we could see from the comparisons of analyses with respect to cloud reflectivity observations, including cloud optical depth in the assimilation system in combination with cloud reflectivity does not bring too much additional information as it is possible to notice from small differences in both, the specific humidity (Fig. 4.11a, b) and the temperature (Fig. 4.3.1a, b) increments between these two experiments (i.e. **refl** vs. **reflopt**).

Analysis increments from the assimilation of level-2 products differ significantly from those created by the assimilation of cloud radar reflectivity, especially for specific humidity (Fig. 4.3.1a). When combining level-2 products with MODIS cloud optical depth, one can see more similarities with increments obtained from either **refl** or **reflopt** assimilation experiments. Such similarities are obvious for instance around 30°N or from 40°to 60°S in the case of specific humidity (Fig. 4.11) and between 50°and 60°N or from the equator to 10°N for the temperature increments (Fig. 4.13). Since the analysis increments have been significantly modified by adding cloud optical depth to the level-2 products, an experiment has been performed assimilating cloud optical depth only. This experiment clearly shows what contributions are coming to the specific humidity and temperature increments from cloud optical depth. These contributions are actually responsible for more similarities in increments between **liwcopt** and **refl** experiments than between **liwc** and **refl** (even more between **liwcopt** and **reflopt**, of course).

Figure 4.3.1a shows that temperature increments from the **liwc** experiment are getting smaller over 4.5 km and do not go higher than up to around 9.5 km, while those from the **refl** or **reflopt** experiments are well spread up to 15 km. That could indicate a small sensitivity of the 1D-Var system to ice water content due to the small sensitivity of the moist parametrization schemes (used as observation operators) to this quantity or too large observation errors for the cloud ice water content products (i.e. the weight of this observation in the system would be small). Indeed, the observation errors for cloud ice water content (iwc) would very often exceed 75% of the observed value. Therefore an experiment has been performed with iwc observation errors reduced by a factor of 4. Such experimentation clearly led to an increase of temperature increments above 4.5 km. This also indicates how a proper definition of the observation errors is important for the assimilation performance.

(c) Comparisons of the first-guess and analysis against independent observations

For a validation of the performed 1D-Var experiments it is important to use independent observations (i.e. observations which were not assimilated in the system). Such observations may not always be available for the geographical location and/or the period for which experiments have been performed. For the situation on 23 January 2007, we can only evaluate the performance of the **refl** and **liwc** experiments against MODIS cloud optical depth. The results of such evaluation are summarized in Table 4.2 and in Fig. 4.15 where only analyzed values of cloud optical depth from the **refl** and **liwc** experiments (marked with green circles) should be considered. From both, the table and figure, one can see that the analysis gets closer to the independent cloud observations by assimilating either cloud radar reflectivity or cloud liquid and ice water contents. In this comparison, the **liwc** experiment performs better for convective situations between points 240 and 470 (corresponding to the area between 20°N and 25°S), which was already obvious from the analyzed values of cloud radar reflectivity compared to the observations in Fig. 4.3.1. This large area of improvement is also responsible for an overall better performance of the **liwc** experiment over **refl** when compared to independent observations. Experiment **refl** gives better results than **liwc** only between points 490 and 540 (i.e. between 25°and 35°S) and slightly better results between 100 and 150 (between 45°and 35°N).

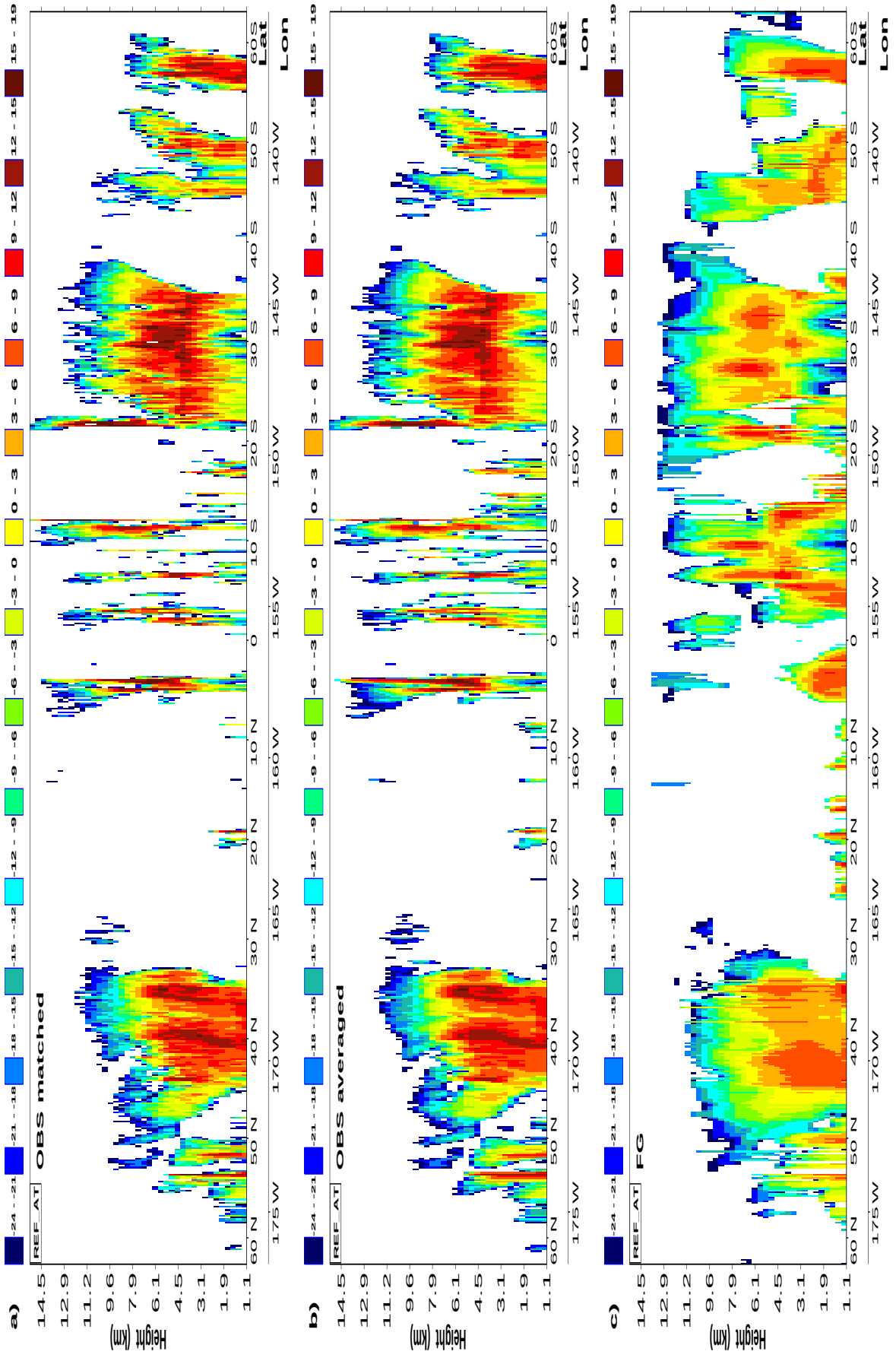


Figure 4.2: *Cloud radar reflectivity (in dBZ) for the situation on 23 January 2007 over the Pacific Ocean (displayed on Fig. 4.1a) - (a) CloudSat observations from 94 GHz radar closest to the model grid-point (matched), (b) observations averaged over the model grid-box, (c) model background (FG).*

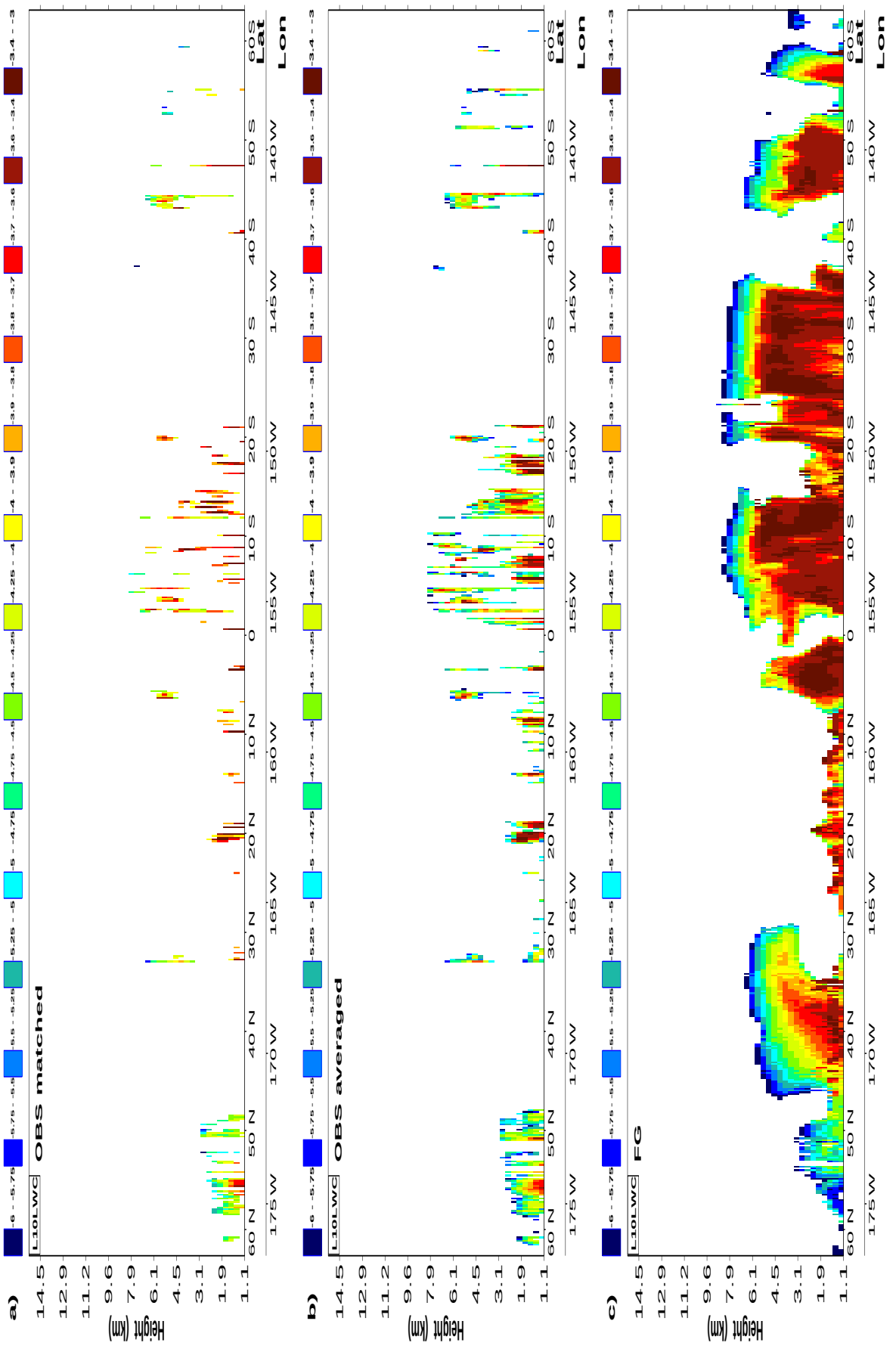


Figure 4.3: *Same as Fig. 4.3.1, but for the cloud liquid water content (in kg.m−3 using logarithmic scale).*

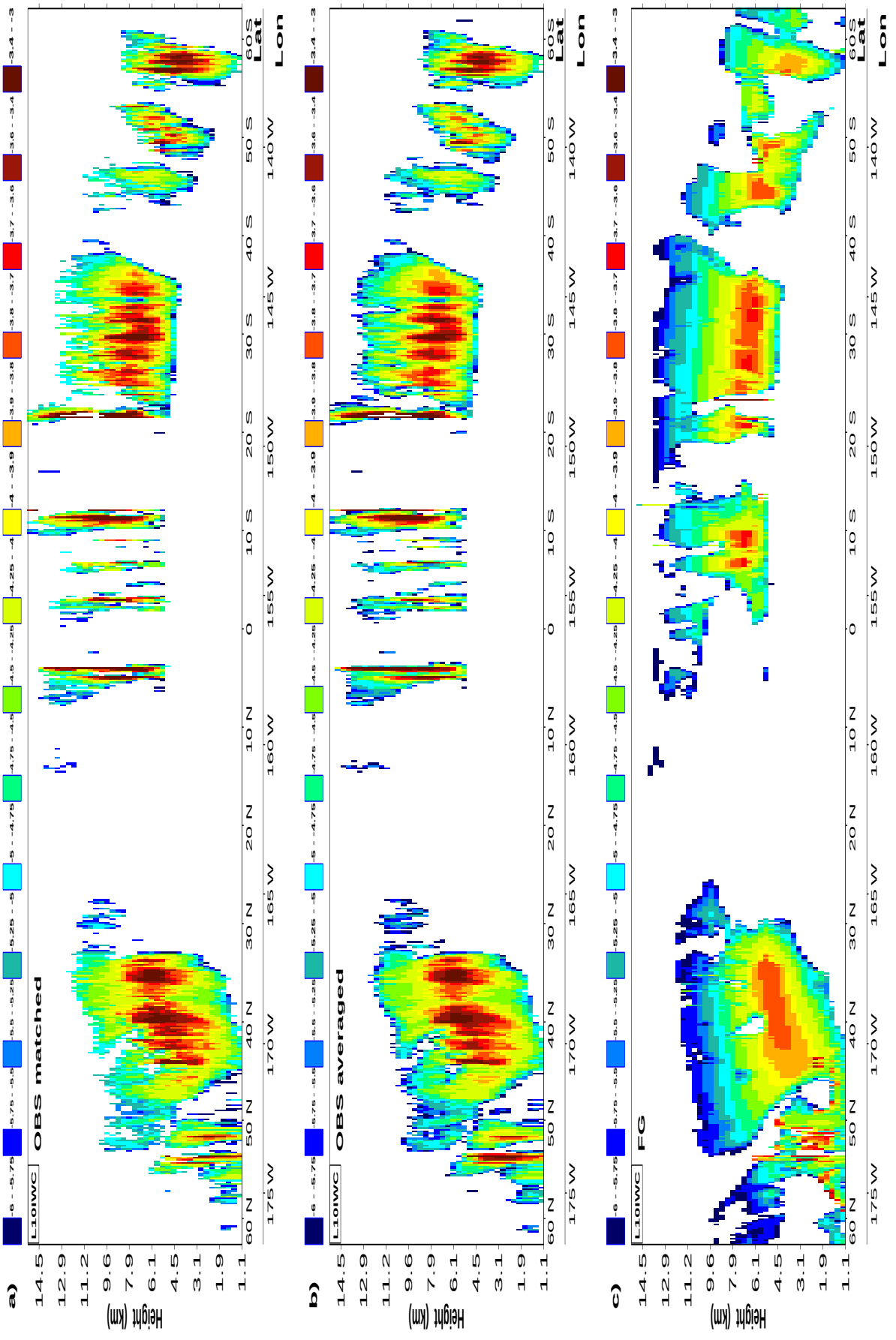


Figure 4.4: *Same as Fig. 4.3, but for the cloud ice water content.*

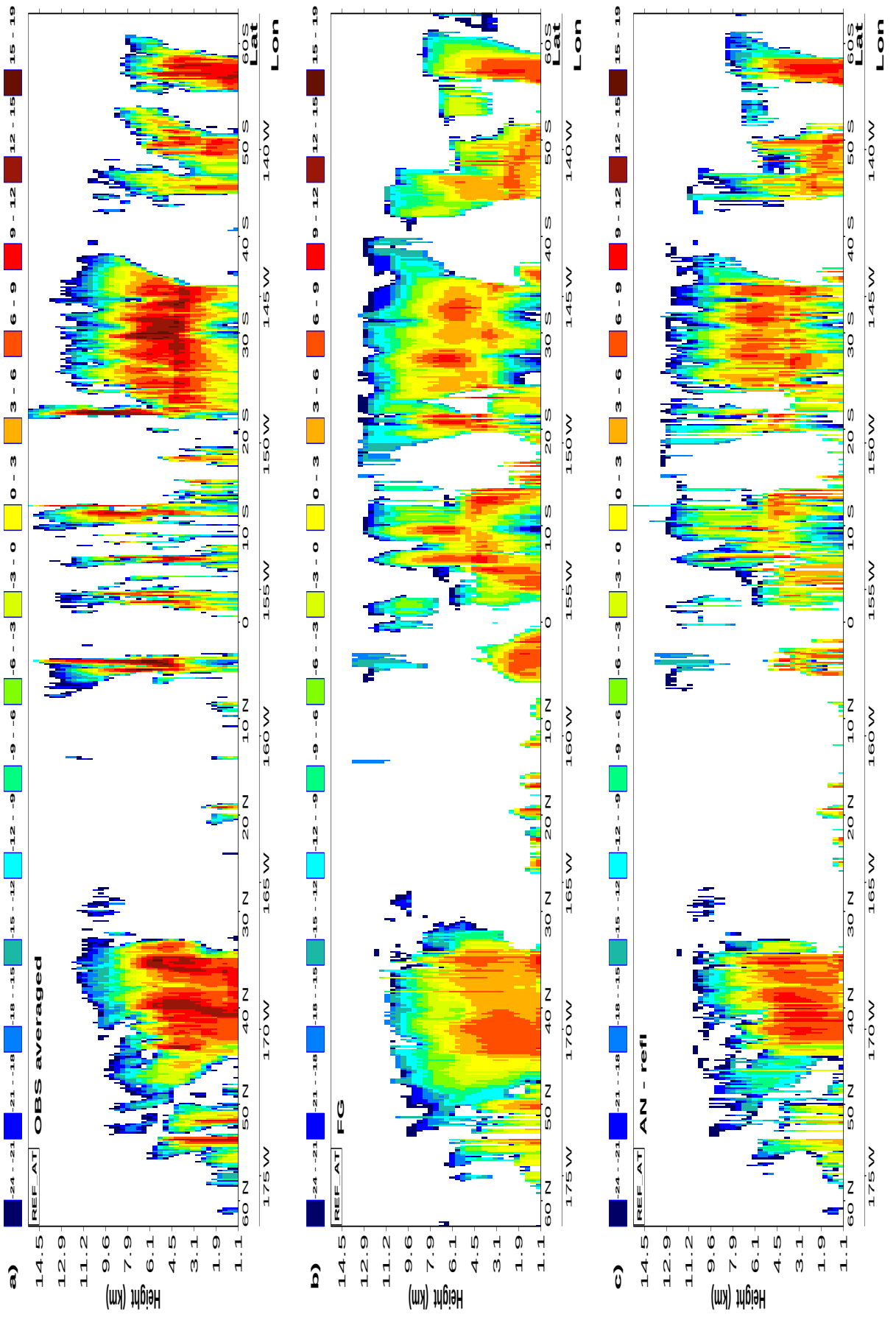


Figure 4.5: *Cloud radar reflectivity (in dBZ) for the situation on 23 January 2007 - (a) CloudSat observations from 94 GHz radar averaged over the model grid-box, (b) model background (FG) and 1D-Var retrieval using averaged cloud reflectivity (AN - refl).*

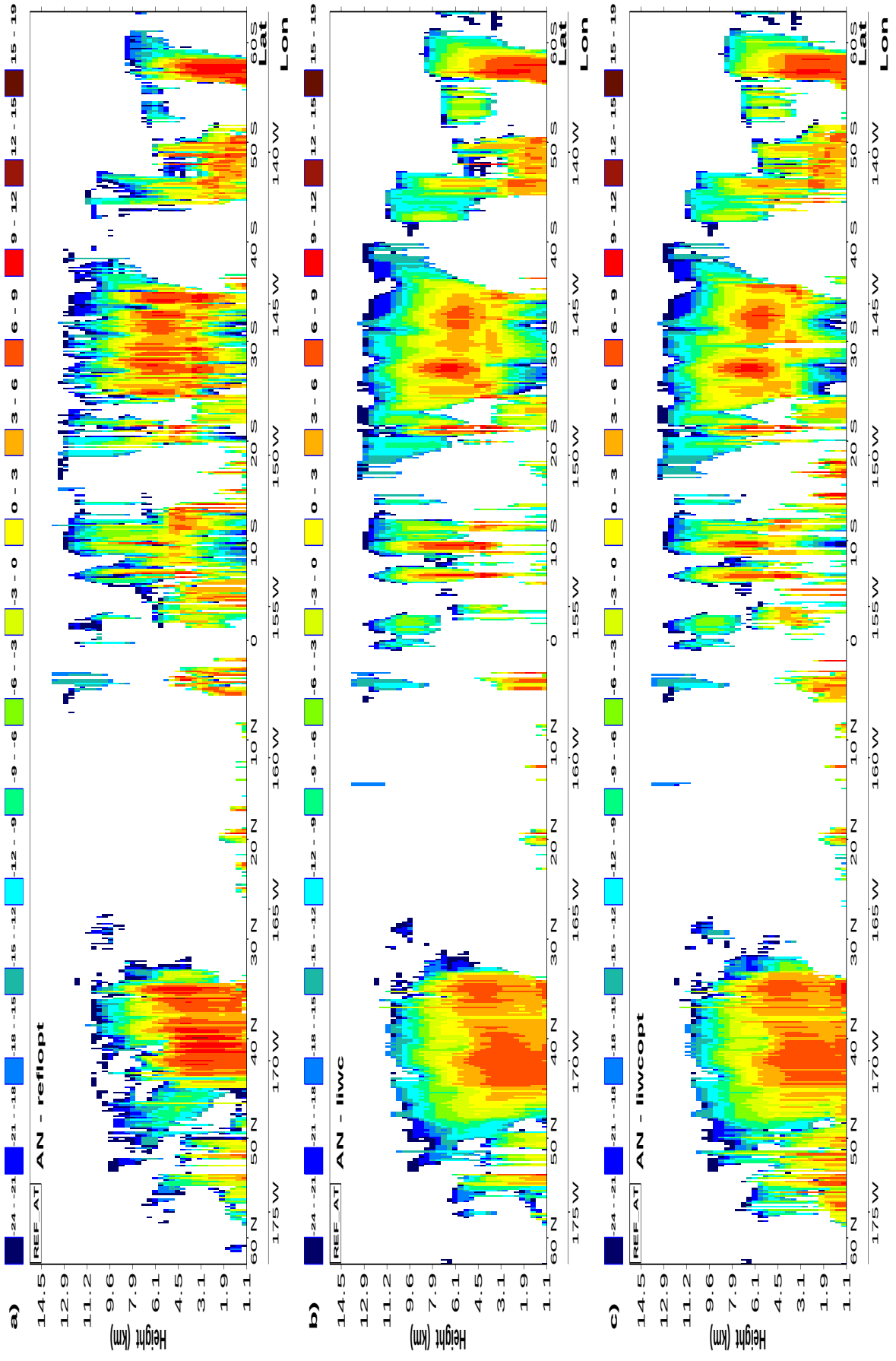


Figure 4.6: *Cloud radar reflectivity (in dBZ) for the situation on 23 January 2007 - 1D-Var retrievals using averaged observations of: (a) cloud reflectivity combined with the MODIS cloud optical depth (AN - reflopt), (b) cloud liquid and ice water contents from CloudSat (AN - liwc), (c) cloud liquid and ice water combined with the cloud optical depth (AN - liwcopt).*

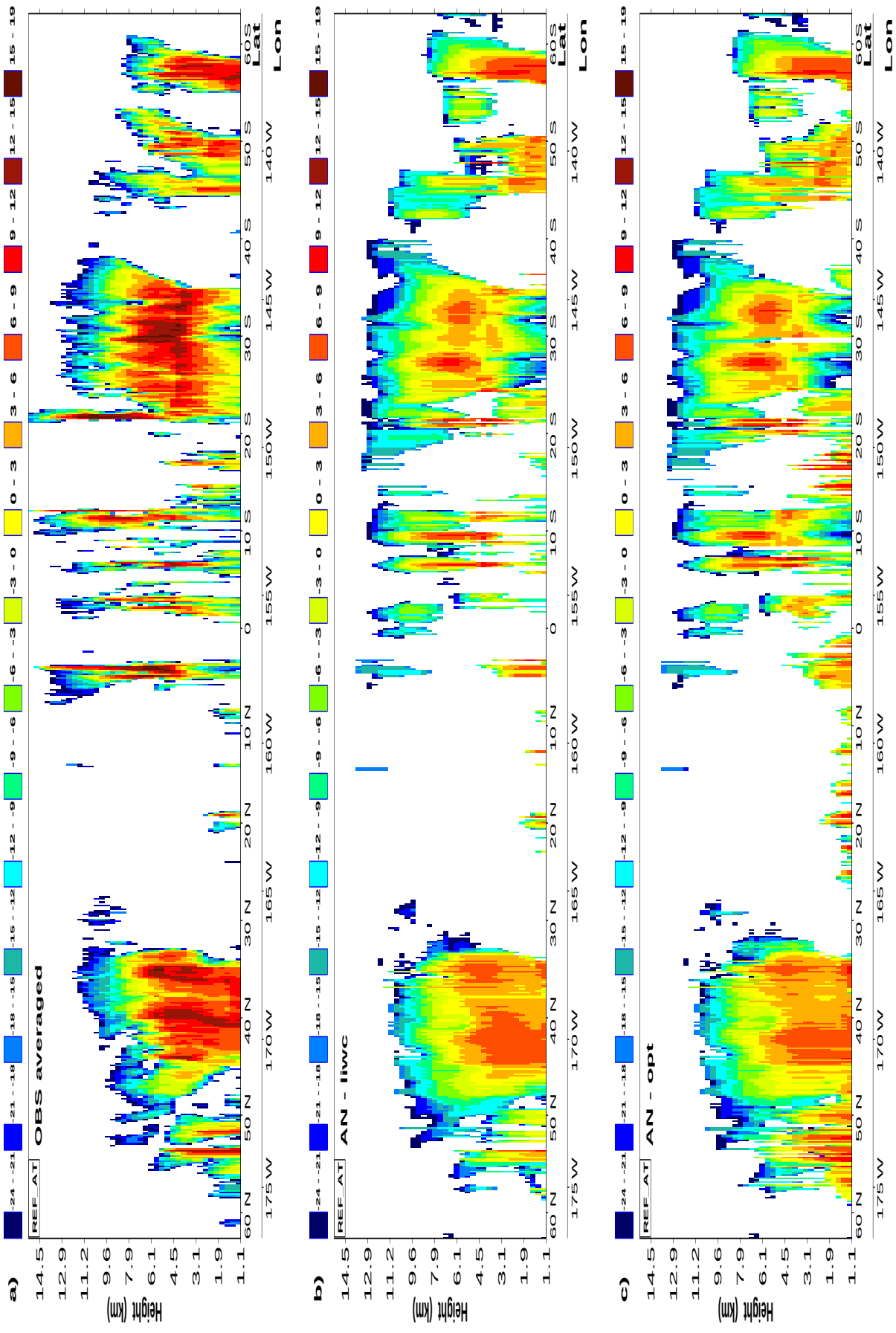


Figure 4.7: *Cloud radar reflectivity (in dBZ) for the situation on 23 January 2007 - (a) CloudSat observations from 94 GHz radar averaged over the model grid-box and 1D-Var retrieval using averaged observations of: (b) cloud liquid and ice water contents from CloudSat (AN - liwc) and (c) MODIS cloud optical depth (AN - opt).*

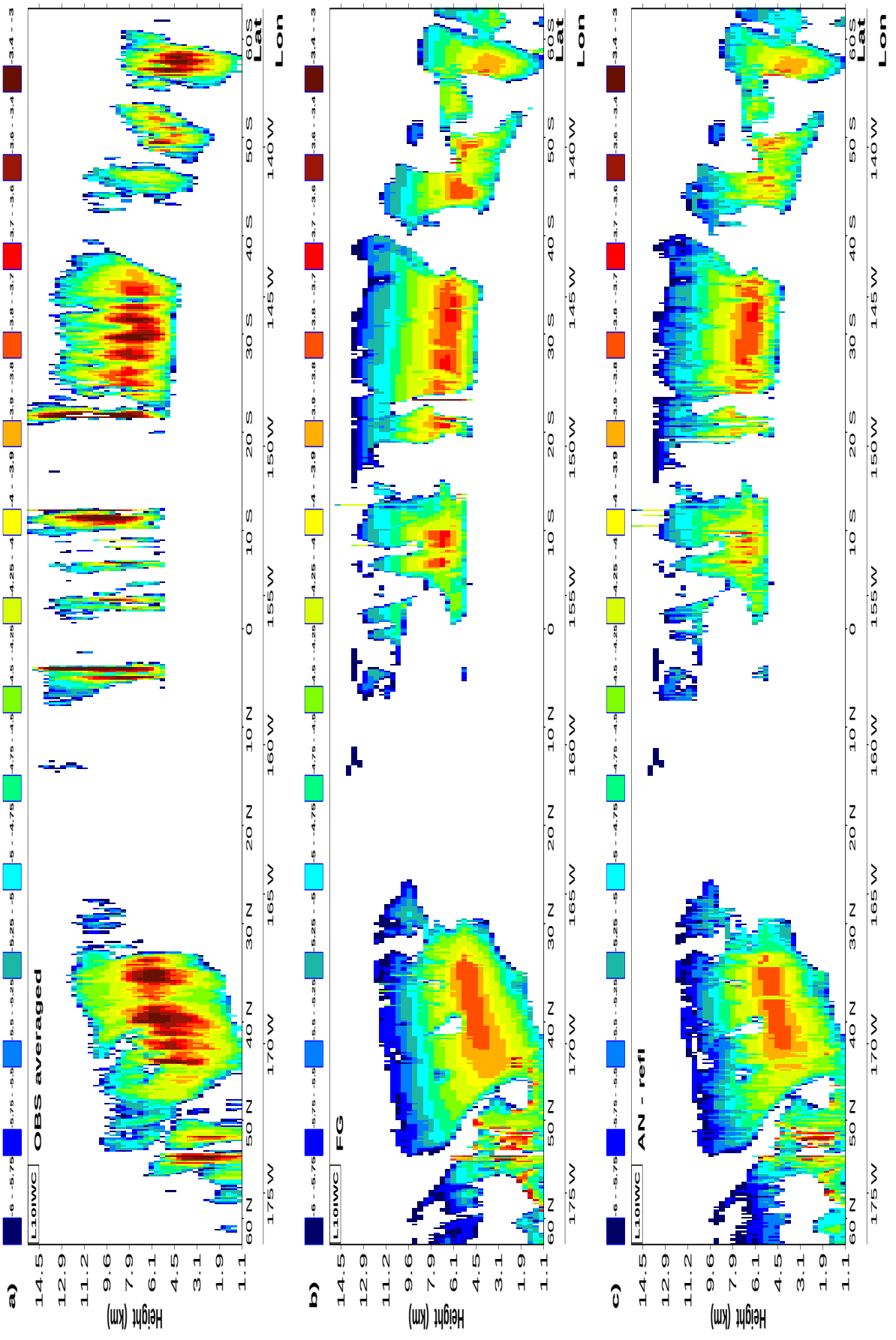


Figure 4.8: *Same as Fig. 4.5, but for CloudSat cloud ice water content (in kg.m−3 using logarithmic scale).*

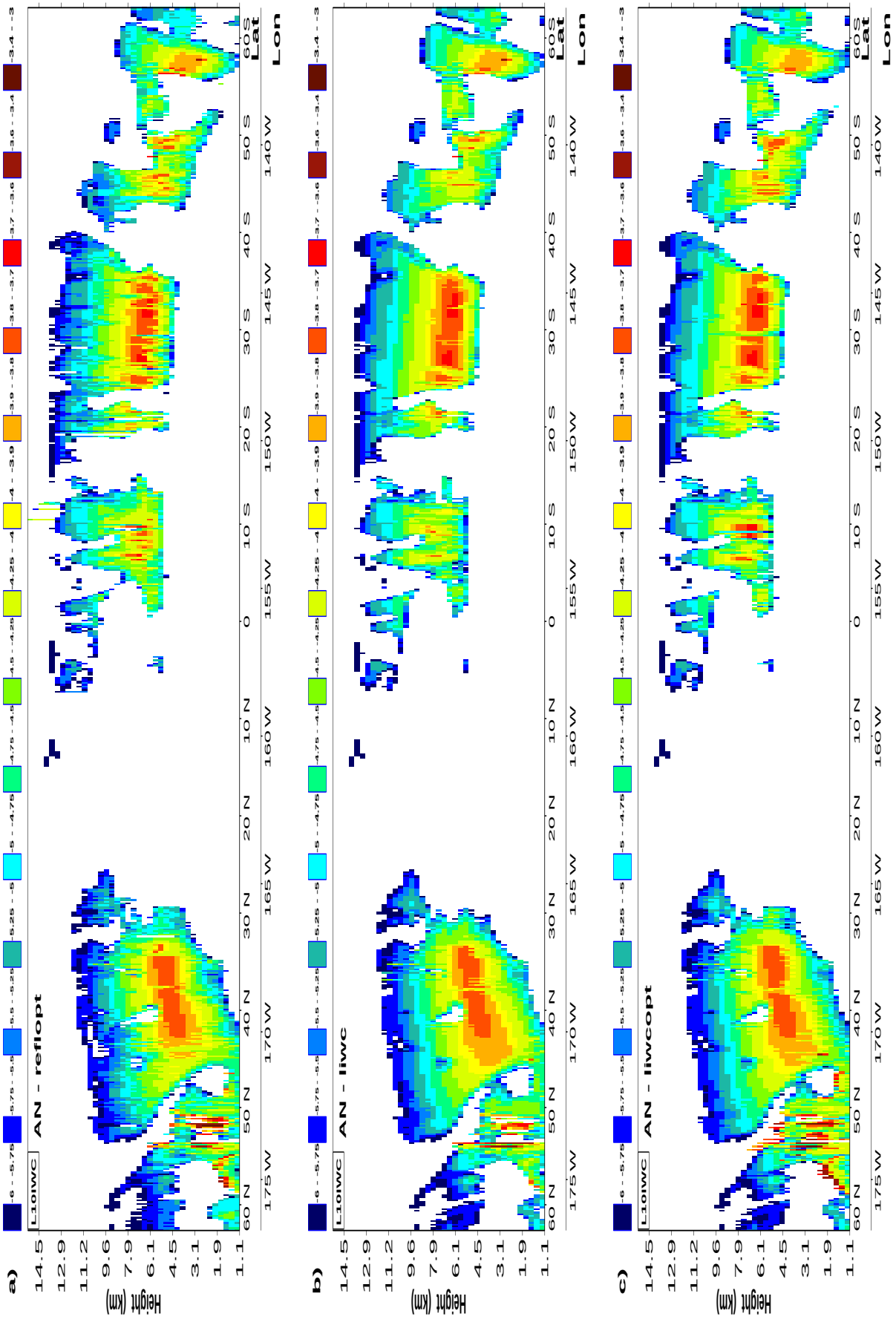


Figure 4.9: *Same as Fig. 4.3.1, but for CloudSat cloud ice water content (in kg.m−3 using logarithmic scale).*

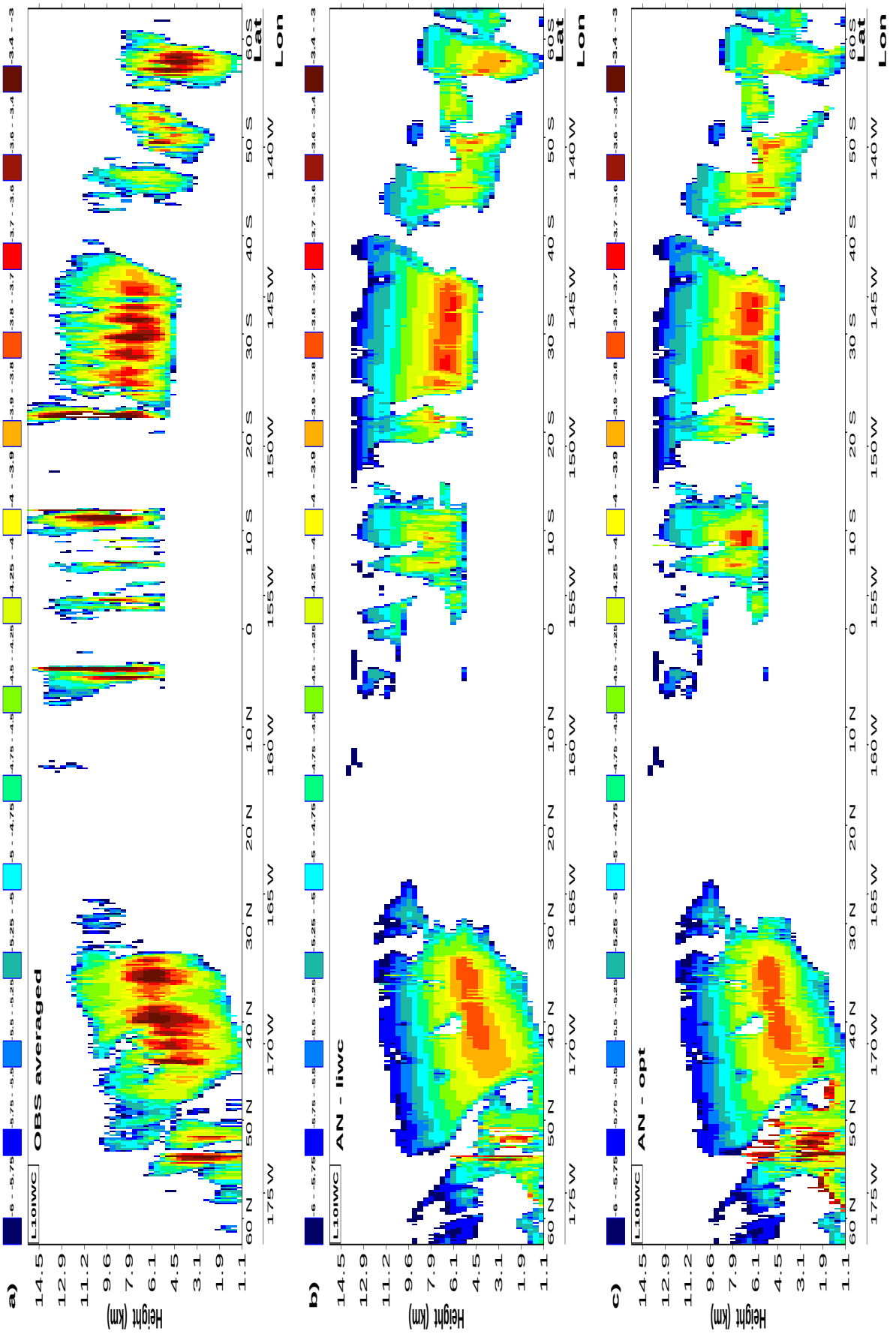


Figure 4.10: *Same as Fig. 4.7, but for CloudSat cloud ice water content (in kg.m−3 using logarithmic scale).*

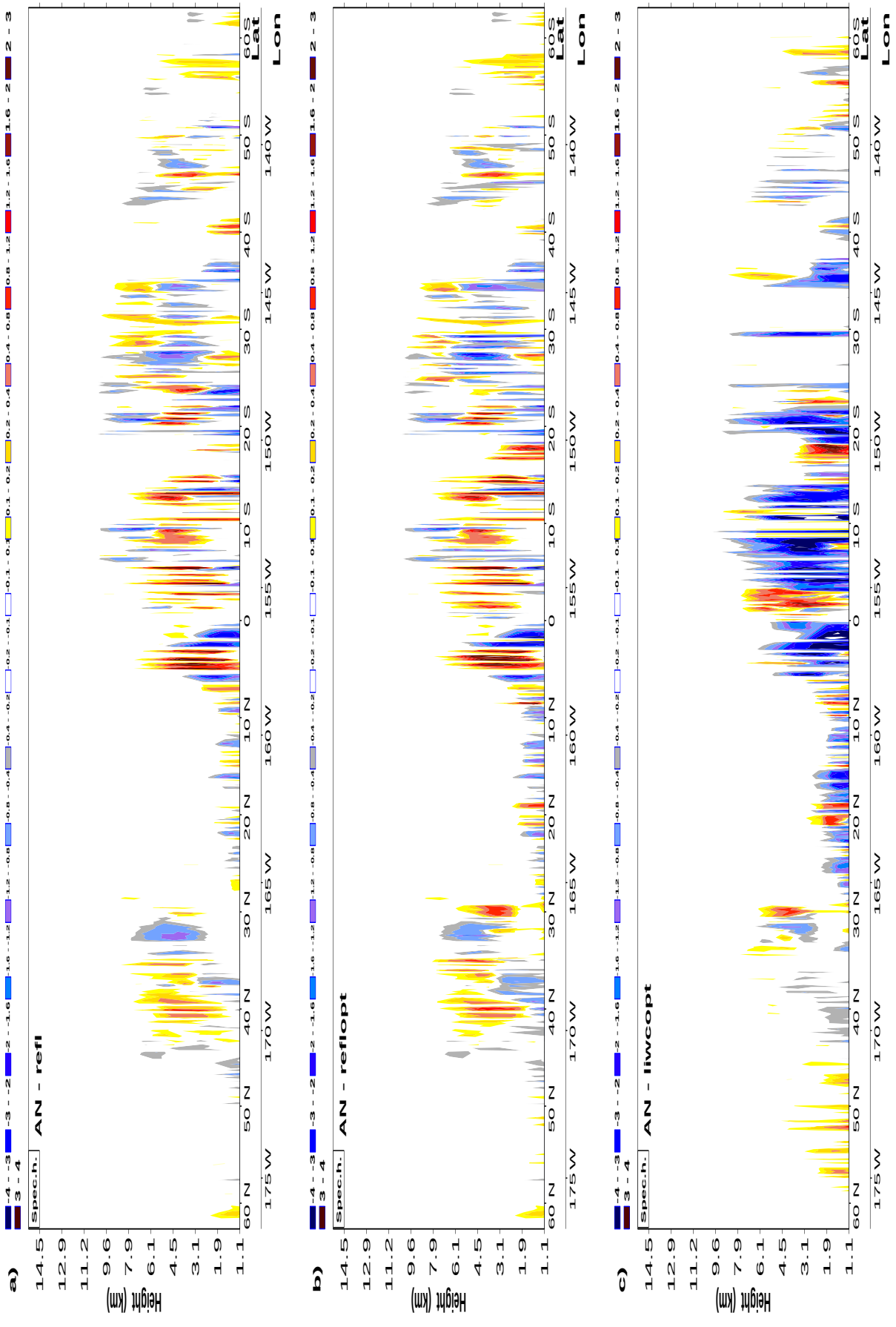


Figure 4.11: *Analysis increments for specific humidity (in g/kg) from 1D-Var retrievals using: (a) cloud reflectivity, (b) cloud reflectivity combined with cloud optical depth and (c) cloud liquid and ice water contents combined with cloud optical depth. Situation of 23 January 2007.*

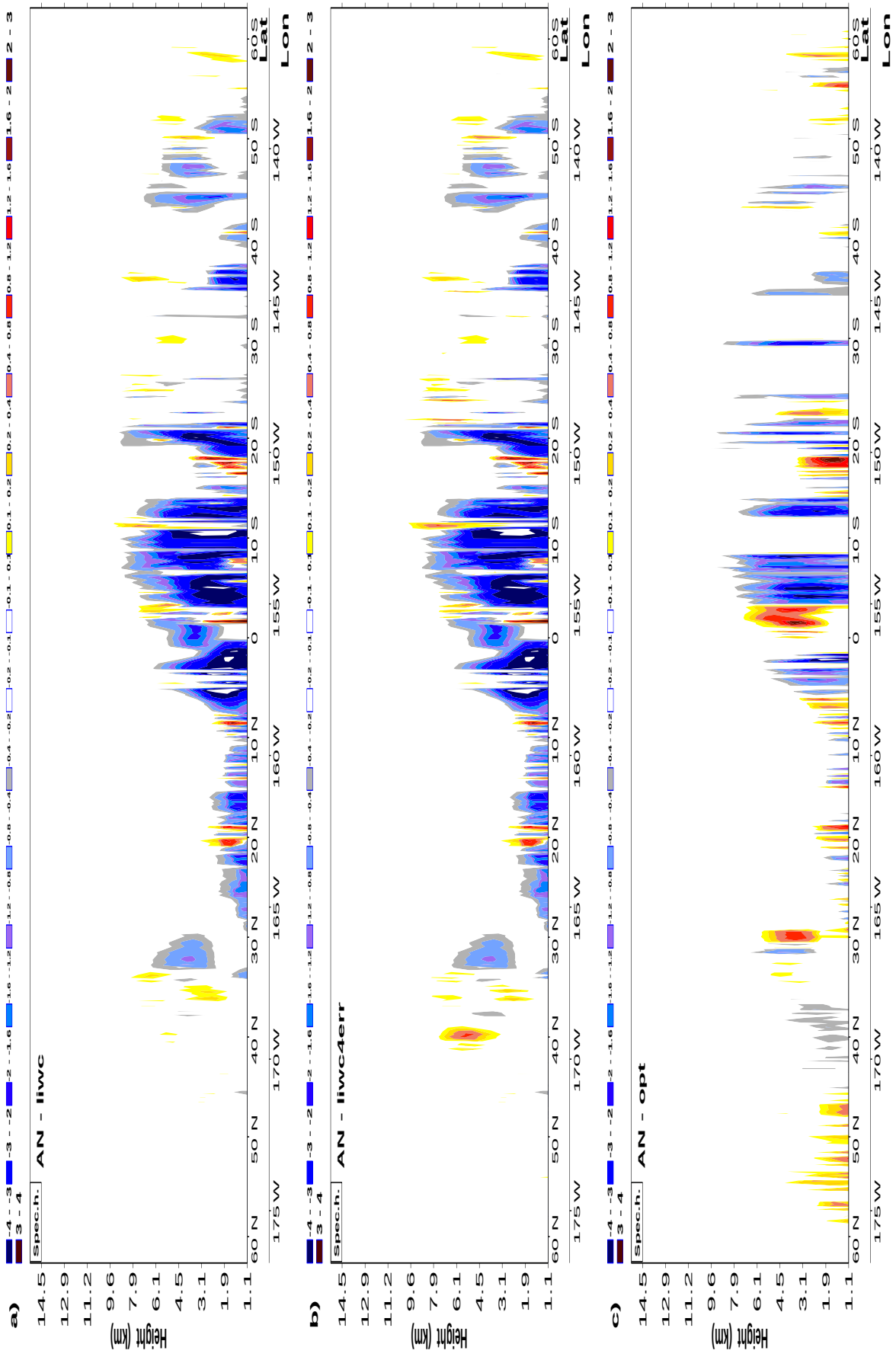


Figure 4.12: *Analysis increments for specific humidity (in g/kg) from 1D-Var retrievals using: (a) cloud liquid and ice water contents, (b) cloud liquid and ice water contents with ice water errors reduced by a factor of 4, (c) cloud optical depth only. Situation of 23 January 2007.*

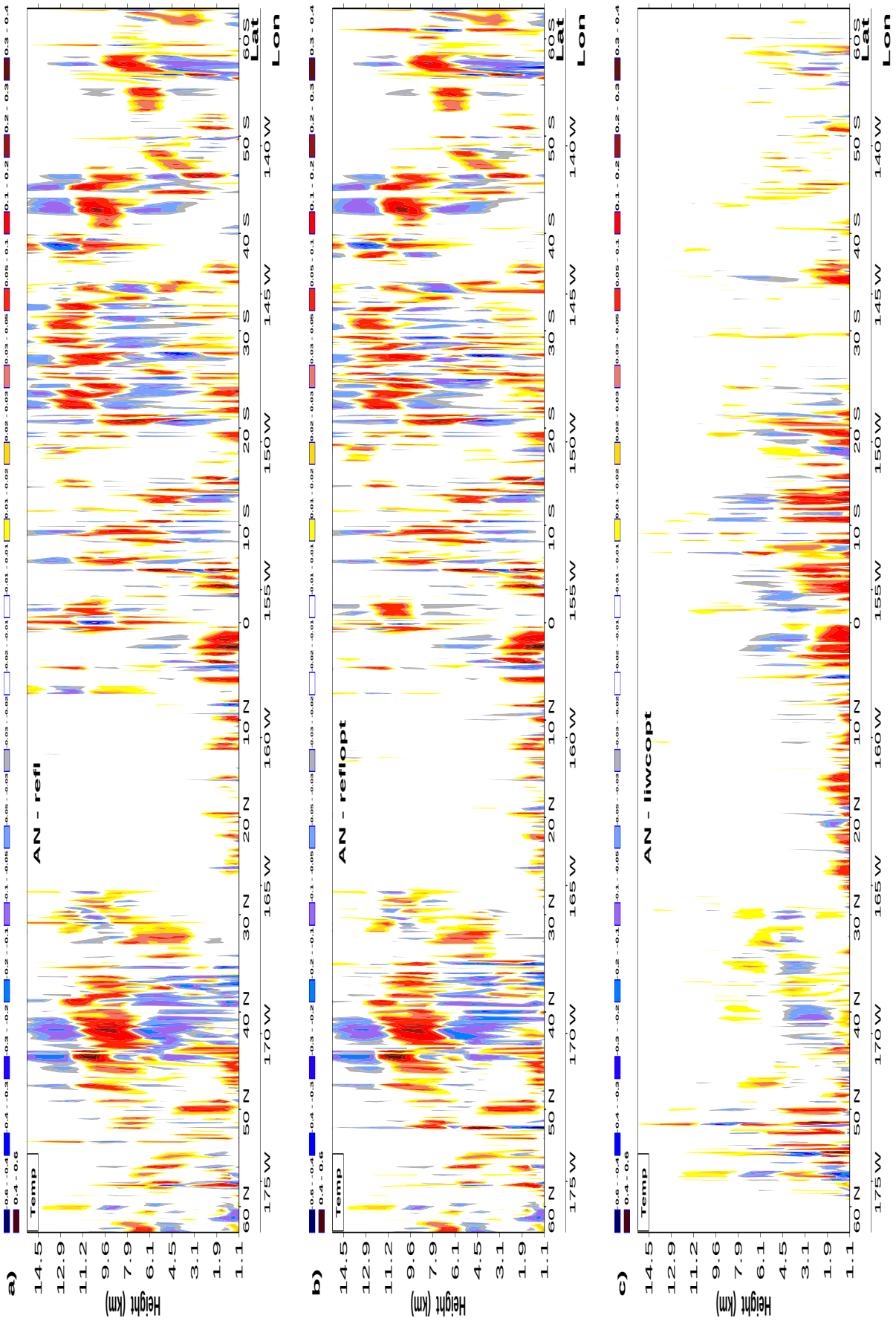


Figure 4.13: *Same as Fig. 4.11, but for analysis increments of temperature (in K).*

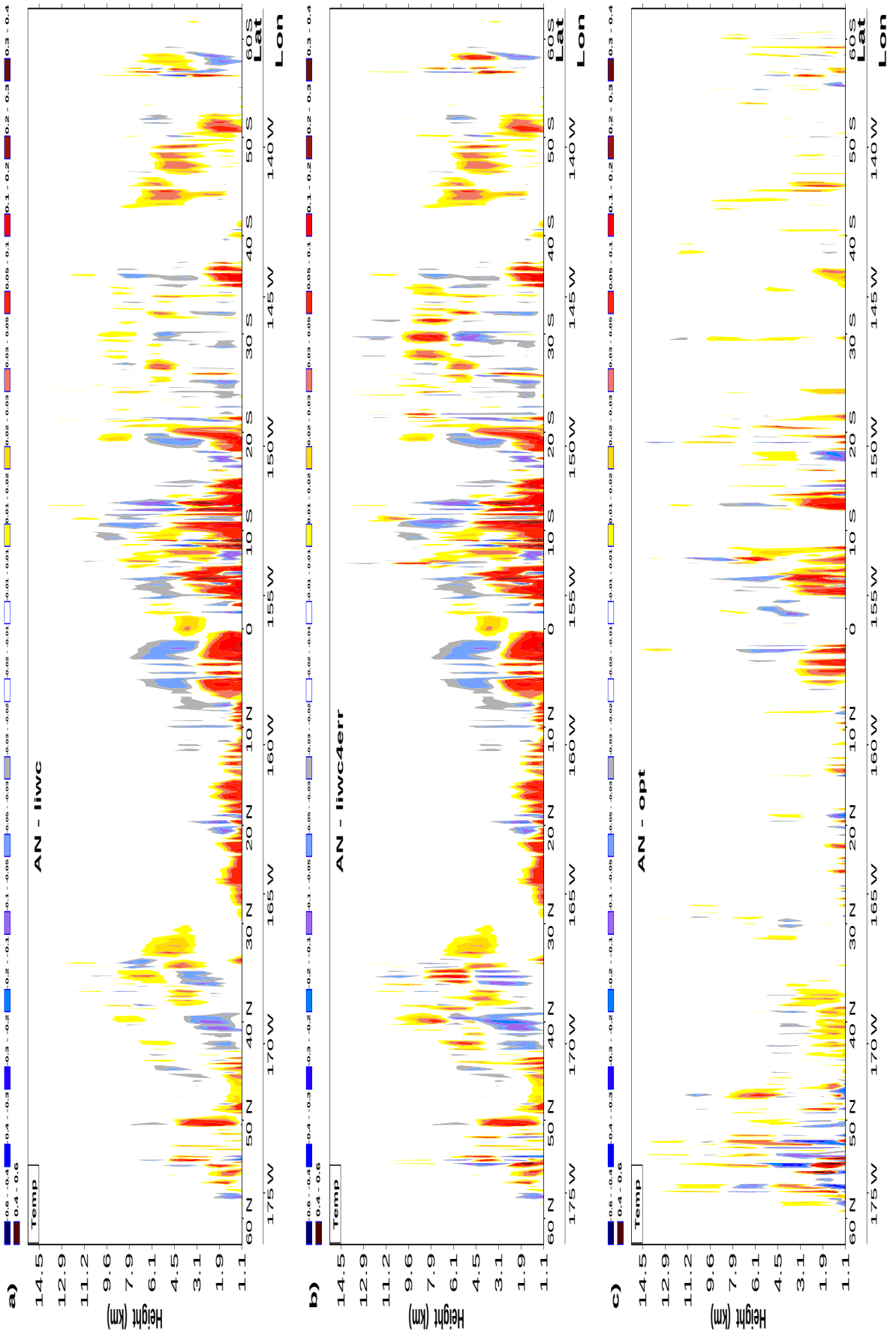


Figure 4.14: *Same as Fig. 4.3.1, but for analysis increments of temperature (in K).*

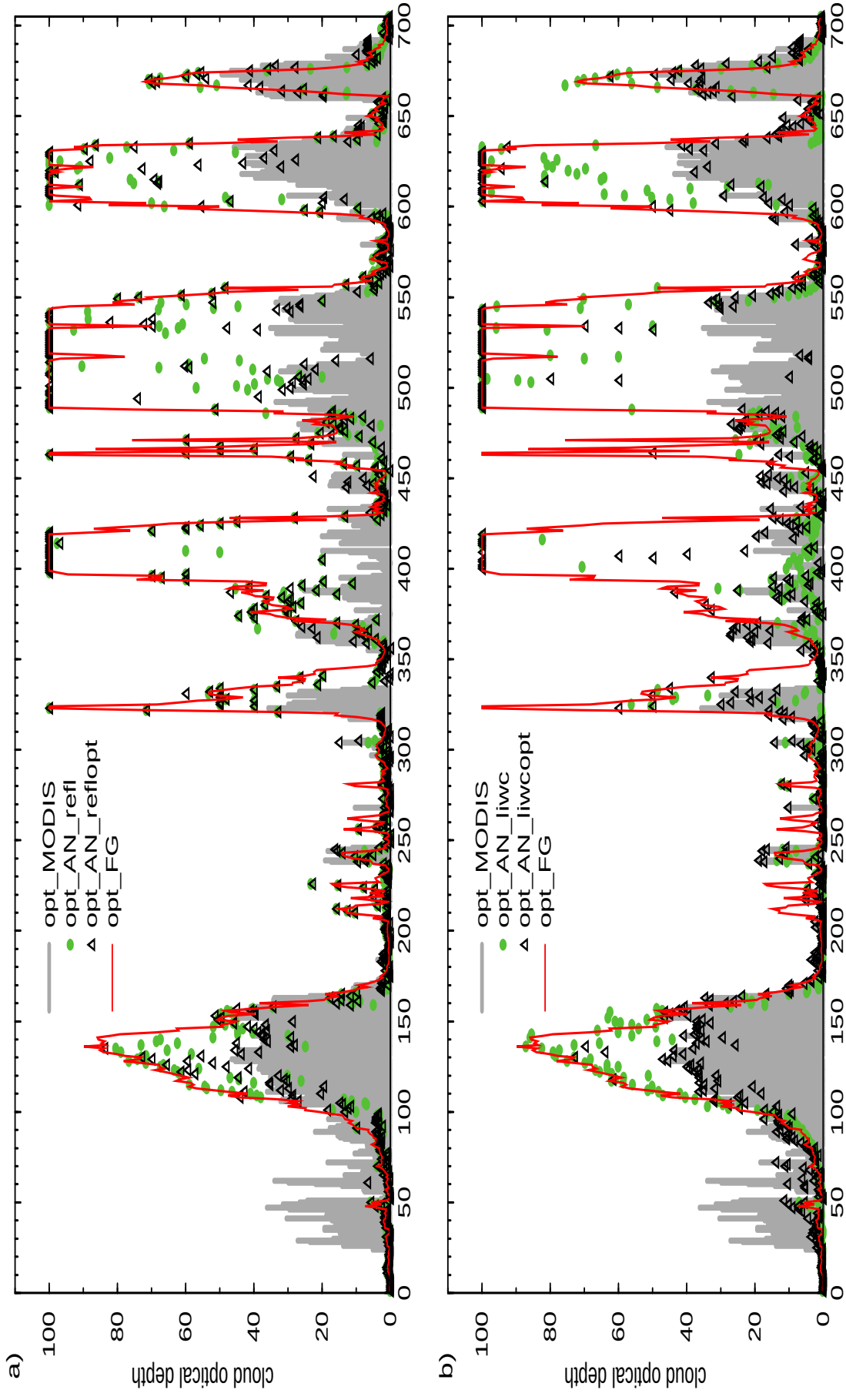


Figure 4.15: *Comparison of the model first guess (red line) and the analyzed values from 1D-Var retrievals using averaged observations of (a) cloud reflectivity (green circle) alone and combined with cloud optical depth (black triangle), (b) cloud liquid and ice water contents (green circle) alone and combined with cloud optical depth (black triangle) against MODIS cloud optical depth observations averaged over the model grid-box (grey shading). Situation of 23 January 2007.*

	MATCHED OBS						AVERAGED OBS					
	reflectivity				cloud iwc		reflectivity				cloud iwc	
	bias	stdv	mae	rms	bias	stdv	bias	stdv	mae	rms	bias	stdv
FG	-4.46	15.06	11.58	15.71	5.69E-3	7.51E-2	-5.27	15.09	11.86	15.98	4.77E-3	5.82E-2
AN-refl	0.36	12.25	7.89	12.25	7.49E-3	7.26E-2	-0.47	12.21	7.79	12.22	6.57E-3	5.56E-2
AN-reflopt	0.38	12.37	8.16	12.38	7.52E-3	7.25E-2	-0.69	12.13	7.86	12.15	6.50E-3	5.59E-2
AN-liwc	-1.91	12.82	9.55	12.96	7.07E-3	7.21E-2	-1.85	12.90	9.55	13.03	6.11E-2	5.44E-2
AN-liwcopt	-2.75	13.40	10.13	13.68	6.63E-3	7.31E-2	-3.74	13.78	10.63	14.28	5.21E-2	5.69E-2

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 CloudSat cloud reflectivity (in dBZ) and ice water content (in g m^{-3}) observations closest to the model grid-box (matched) or averaged over the model grid-box. 705 profiles were included in the statistics for the case of 23 January 2007.*

	MATCHED OBS		AVERAGED OBS	
	opt.depth		opt.depth	
	bias	stdv	bias	stdv
FG	-25.134	37.332	-20.815	34.316
AN-refl	-18.228	36.137	-14.684	31.749
AN-reflopt	-17.498	35.665	-12.937	30.566
AN-liwc	-15.631	34.777	-9.815	29.232
AN-liwcopt	-13.562	31.424	-10.809	28.632

Table 4.2: *Same as Table 4.1, but for bias and standard deviation of the departures from MODIS cloud optical depth.*

4.3.2 Situation on 24 April 2008 over USA

The satellite track for the situation on 24 April 2008 between 19:13 and 19:23 UTC is displayed in Fig. 4.1 and also in Fig. 4.16d. It crosses over a cloud system above USA which is covered by 193 model grid points. This situation has been selected since there is some precipitation in the centre of the system (Fig. 4.17d) and therefore allows us to use the NEXRAD precipitation data as additional independent observations for the validation of our 1D-Var experiments.

(a) Comparisons of the first-guess and analysis against assimilated observations

The observations of cloud radar reflectivity which are either matched with the model data or averaged over the model grid-box for the situation on 24 April 2008 are shown in Fig. 4.16a, b. Similarly as in the previous case, the differences between the matched and averaged observations are quite small, the averaged observations just being slightly smoother.

The first-guess cloud reflectivity equivalent to the observations is displayed in Fig. 4.16c. There are obvious similarities between the FG reflectivities and the observations. However the FG hydrometeor contents (mainly the ice water content - see also Fig. 4.18) are more wide spread, especially above 6 km, though the values are usually small in the levels above 8 km. On the other hand, the maximum values of model cloud reflectivity and also ice water content are smaller than the observed ones.

Results from the different assimilation experiments, the setup of which has been described in subsection 4.1, are displayed in Fig. 4.16 - 4.20 and in Tables 4.3 - 4.4. The figures only show the results from assimilation experiments using averaged observations.

The analyzed cloud radar reflectivity is closer to the observations by assimilating the different types of

observations (Fig. 4.16). The fit of the analysis to cloud reflectivity observations is the most improved when cloud radar reflectivity is assimilated. Not only the large negative values of the cloud reflectivity (when in dBZ) are removed, but also maximum values are increased bringing the analysis closer to the observations. Adding cloud optical depth to cloud reflectivity in the 1D-Var system does not modify the analyzed cloud reflectivity significantly. In the case of using level-2 products (cloud liquid and ice water contents) in 1D-Var (Fig. 4.16f), the analysed cloud reflectivity is also improved when compared with the FG, but is further away from the observations than in the case of assimilating the reflectivity observation alone or combined with optical depth. Similarly, as in the previous case, adjustments towards the observations are very small above 7 km which could be due to the large observation errors for the cloud ice water content and/or the small sensitivity of the observation operator for moist processes to small values of ice, as mentioned before.

The impact coming from the assimilation of different types of observations on cloud liquid and ice water contents is displayed in Fig. 4.17 and Fig. 4.18, respectively. The modifications of cloud liquid water content seem to be larger than for ice water content, but cannot be directly verified since there are hardly any observations of cloud liquid water content. The assimilation of cloud radar reflectivity leads to larger modifications in the levels above 7 km than for the **liwc** and **liwcopt** cases for the possible reasons already mentioned above.

Table 4.3 summarizes the results from the assimilation experiments using either matched or averaged observations. The FG departures are smaller for the averaged observations of cloud ice water content than for the matched ones, but averaging of cloud reflectivity increases the FG departures. However, similarly as for the case on 23 January 2007, the AN departures are usually smaller when assimilating observations averaged over the model grid-box. The overall analysis performance is also quite similar to the previous case.

(b) Comparisons of the first guess and analysis against independent observations

Two sets of independent observations can be used for a validation of the 1D-Var performance in this case, though only the **refl** and **liwc** experiments can be evaluated using the MODIS cloud optical depth. All performed experiments can be compared against NEXRAD precipitation data representing fully independent observations. Comparisons of the first guess and analysis against independent observations are summarized in Fig. 4.19 - 4.20 (only analyzed results represented by green circles should be considered for independent validation) and Table 4.4 (only values in black). The analyses get closer to MODIS cloud optical depth by assimilating either cloud radar reflectivity or cloud liquid and ice water contents. When assimilating averaged observations, there are very small differences in bias and standard deviation for the **refl** and **liwc** analyses. The validation of the different analyses against the NEXRAD precipitation data (in mm h^{-1}) displayed in Fig. 4.20 and summarized in Table 4.4 shows that there is a better fit to the precipitation for all assimilation experiments. The best results are obtained for the **refl** experiment, followed by **liwc**. One can see that adding MODIS cloud optical depth on top of either level-1 or level-2 CloudSat products drives the analysis further away from the NEXRAD precipitation data, contrary to what one would expect.

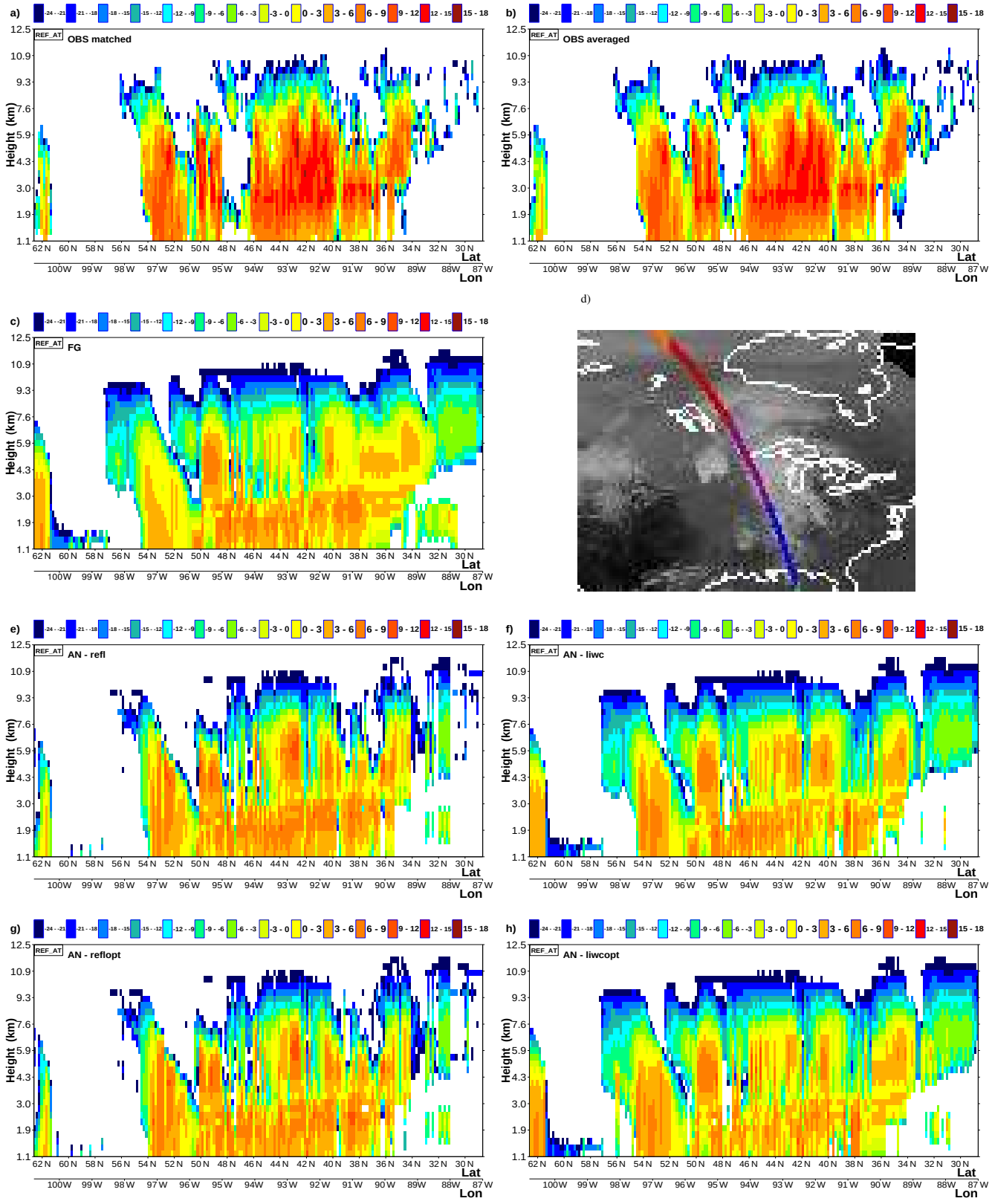


Figure 4.16: *Cloud radar reflectivity (in dBZ) for the situation (d) on 24 April 2008 over USA - (a) CloudSat observations from 94 GHz radar closest to the model grid-point (matched), (b) observations averaged over the model grid-box, (c) model first guess (FG) and 1D-Var retrievals using averaged observations of: (e) cloud reflectivity (AN-refl) alone and combined with MODIS cloud optical depth (AN-reflopt), (f) cloud liquid and ice water contents from CloudSat (AN-liwc) alone and (h) combined with cloud optical depth (AN-liwcopt).*

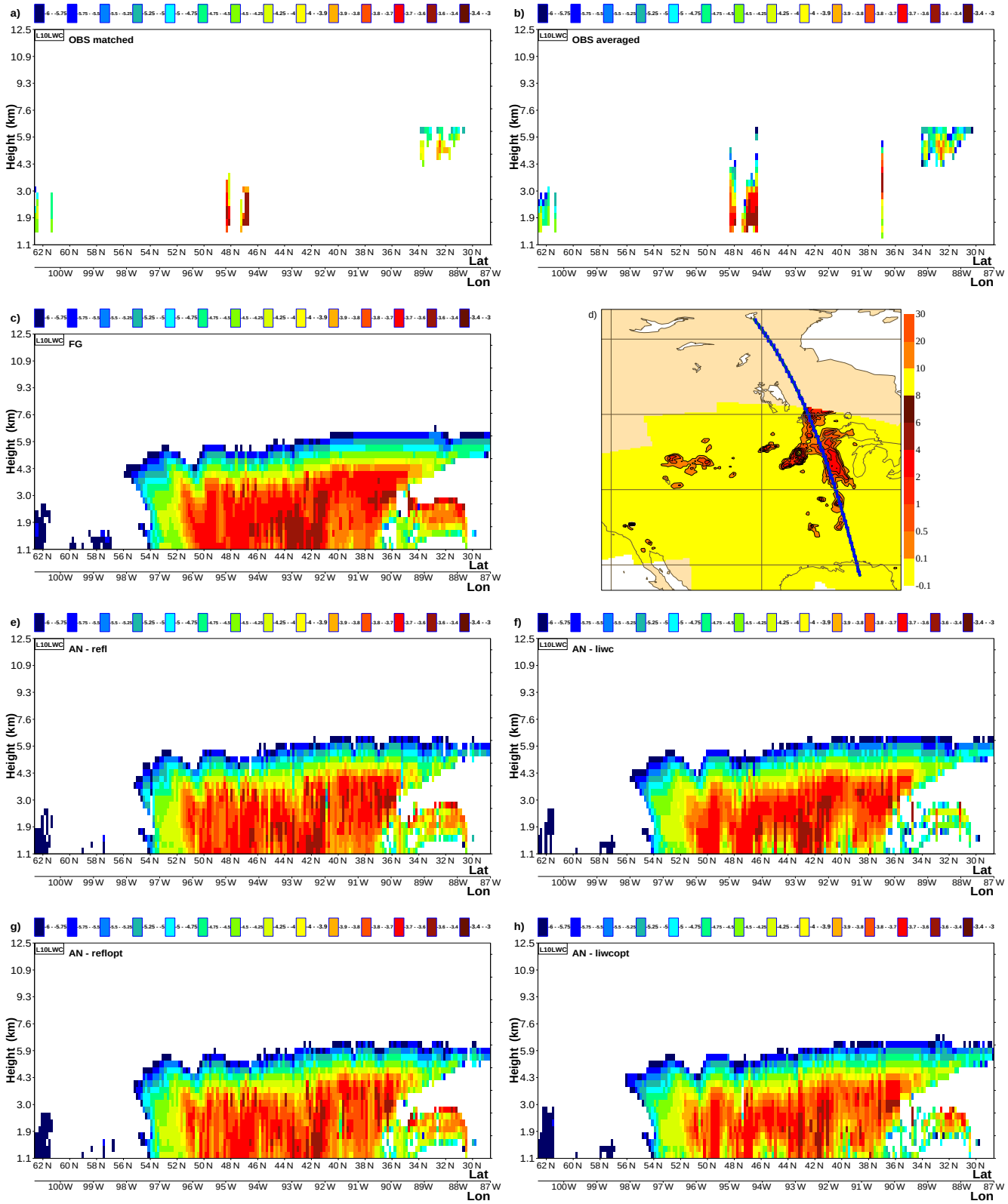


Figure 4.17: *Same as Fig. 4.16, but for cloud liquid water content (in kg m^{-3} using logarithmic scale).(d) NEXRAD precipitation rate in mm h^{-1} (yellow shading indicates observation coverage).*

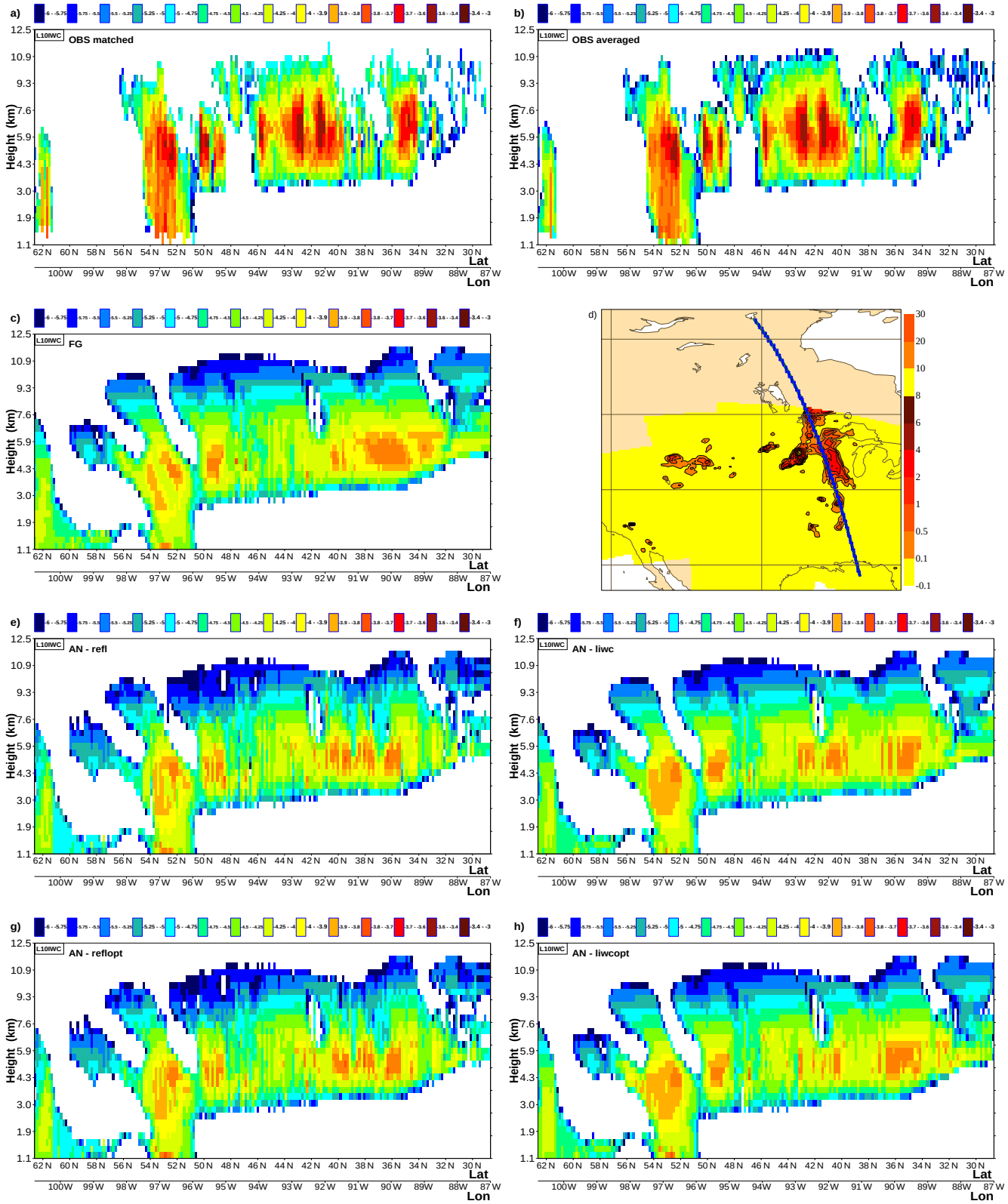


Figure 4.18: *Same as Fig. 4.17, but for cloud ice water content.*

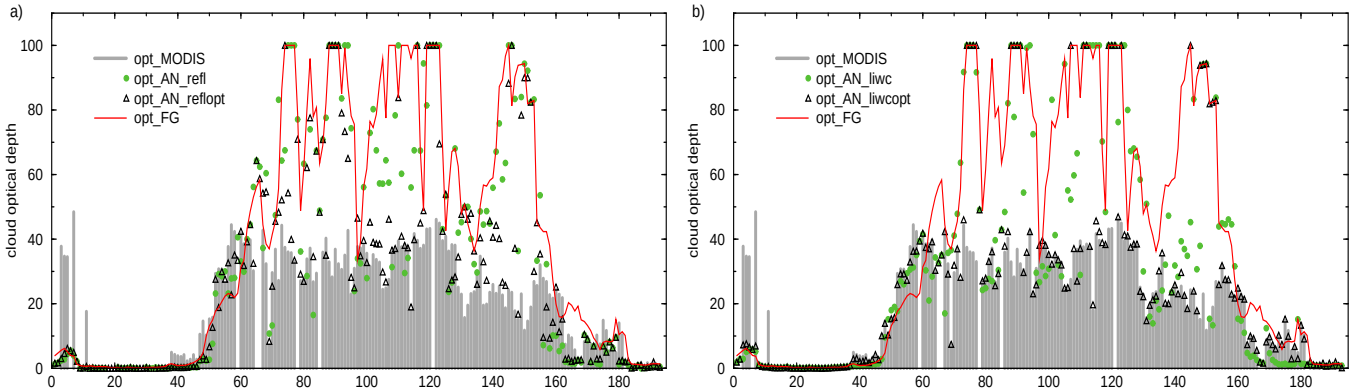


Figure 4.19: *Comparison of the model first guess (red line) and the analyzed values from 1D-Var retrievals using averaged observations of (a) cloud reflectivity (green circle) alone and combined with cloud optical depth (black triangle), (b) cloud liquid and ice water contents (green circle) alone and combined with cloud optical depth (black triangle) against MODIS cloud optical depth observations averaged over the model grid-box (grey shading). Situation of 24 April 2008.*

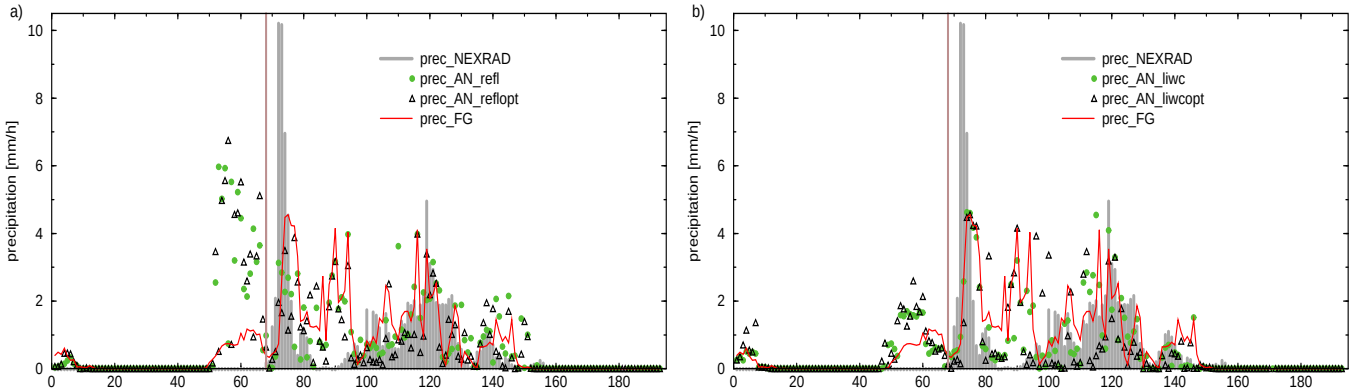


Figure 4.20: *Same as Fig. 4.19, bur for the comparison against the NEXRAD precipitation data. Brown vertical line indicates the beginning of observation availability.*

	MATCHED OBS						AVERAGED OBS					
	reflectivity				cloud iwc		reflectivity				cloud iwc	
	bias	stdv	mae	rms	bias	stdv	bias	stdv	mae	rms	bias	stdv
FG	-2.89	13.17	9.91	13.29	4.99E-3	5.20E-2	-3.59	13.28	10.09	13.44	5.01E-3	5.10E-2
AN-refl	2.31	8.93	5.97	8.83	7.63E-3	4.90E-2	1.79	9.06	5.41	8.53	7.19E-3	4.76E-2
AN-reflopt	2.31	8.98	6.01	8.82	7.49E-3	4.89E-2	1.76	9.50	5.70	8.75	7.07E-3	4.78E-2
AN-liwc	-1.64	12.49	8.80	11.73	5.66E-3	4.54E-2	-1.67	12.81	8.75	11.56	5.65E-3	4.48E-2
AN-liwcopt	-1.91	12.54	8.93	11.98	5.48E-3	4.70E-2	-2.59	13.68	9.63	12.82	5.21E-3	4.73E-2

Table 4.3: *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 CloudSat cloud reflectivity (in dBZ) and ice water content (in g m^{-3}) observations closest to the model grid-box (matched) or averaged over the model grid-box. 193 profiles were included in the statistics for the case of 24 April 2008.*

	MATCHED OBS				AVERAGED OBS			
	opt.depth		precip		opt.depth		precip	
	bias	stdv	bias	stdv	bias	stdv	bias	stdv
FG	-27.04	30.45	-0.117	1.502	-24.41	29.70	-0.117	1.502
AN_refl	-15.97	28.78	-0.033	1.516	-13.91	27.44	-0.024	1.346
AN_refl_opt	-11.92	26.52	-0.080	1.533	-8.94	23.19	0.102	1.461
AN_liwc	-19.24	29.07	-0.058	1.431	-13.56	27.17	0.085	1.371
AN_liwc_opt	-11.81	27.13	0.110	1.404	-8.03	24.21	0.171	1.591

Table 4.4: *Same as Table 4.3, but for bias and standard deviation of the departures from MODIS cloud optical depth and from NEXRAD precipitation data (in mm h^{-1}).*

4.3.3 Situation on 16 September 2007 - tropical cyclone Wipha

Typhoon Wipha used in our 1D-Var experimentation originated from a tropical disturbance on 15 September 2007 and it rapidly intensified into a typhoon in the following date. The peak intensity was reached on 18 September 2007 with winds of 185 km h^{-1} and with a minimum pressure of 925 hPa. On the date of our assimilation experiments, 16 September 2007 00 UTC, Wipha was a tropical storm with a central pressure of 992 hPa. This is a more challenging situation for any assimilation system as the first-guess departures are more significant in this case. Figure 4.21c shows that the model produces a storm system approximately in the right location. However, the model system is more wide spread and has more continuous cloud cover on the northern and especially on the southern edges of the storm system compared to observations.

(a) Comparisons of the first-guess and analysis against assimilated observations

Figure 4.21 displays the cloud radar reflectivity observations either matched with the model data (a) or averaged over the model grid-box together with the model first-guess equivalents (c) and the analyzed values (e - h) obtained from the different assimilation experiments. The cloud liquid and ice water contents are shown in Fig. 4.22 and Fig. 4.23, respectively. For this tropical storm, there are obviously more differences between the matched and averaged observations. On top of that, similarly to the 23 January 2007 case, averaging also brings more information about the liquid water content which is profitable for assimilation system as shown by the **liwc** and **liwcopt** experiments. The analyzed amount of liquid water content gets significantly closer to the observations between 20°and 28°N as well as between 6°and 8°N (Fig. 4.22f, h). At the same time, the analyzed cloud reflectivity is also much closer to the observations for these assimilation experiments in those areas, actually even closer than in the case of the assimilation of cloud radar reflectivity directly. In the areas of missing observations of the cloud liquid content observations (between 8°and 20°N), it is the assimilation of cloud reflectivity either alone or combined with cloud optical depth which performs better. Using MODIS cloud optical depth in combination with cloud radar reflectivity or with cloud liquid and ice water contents in the assimilation system has only a small impact on the additional adjustment of the analysis with respect to cloud reflectivity. An improvement is observed between 12°and 14°N for the **reflopt** experiment or between 18°to 22°N for the **liwcopt** experiment.

The results from the assimilation experiments using either matched or averaged observations are also summarized in Table 4.5. For this situation, an assimilation of cloud liquid and ice water contents seems to perform remarkably better than the reflectivity assimilation when comparing the analyzed values not only to the observations of cloud liquid and ice water contents, but also to cloud reflectivity data.

(b) Comparisons of the first-guess and analysis against independent observations

The cloud optical depth derived from the MODIS data has been used as an independent observation for the validation of the **refl** and **liwc** experiments. The results of such validation are shown in Fig. 4.24 where only

analyzed values of optical depth represented by green circles should be considered as independent validation. Comparisons are also summarized in Table 4.5 (only values in black represent independent validation). The analysis from the **liwc** experiment is closer to the MODIS cloud optical depth than the analysis from the **refl** experiment, which is consistent with results of the comparison to cloud reflectivity and cloud liquid and ice water contents analyzed above.

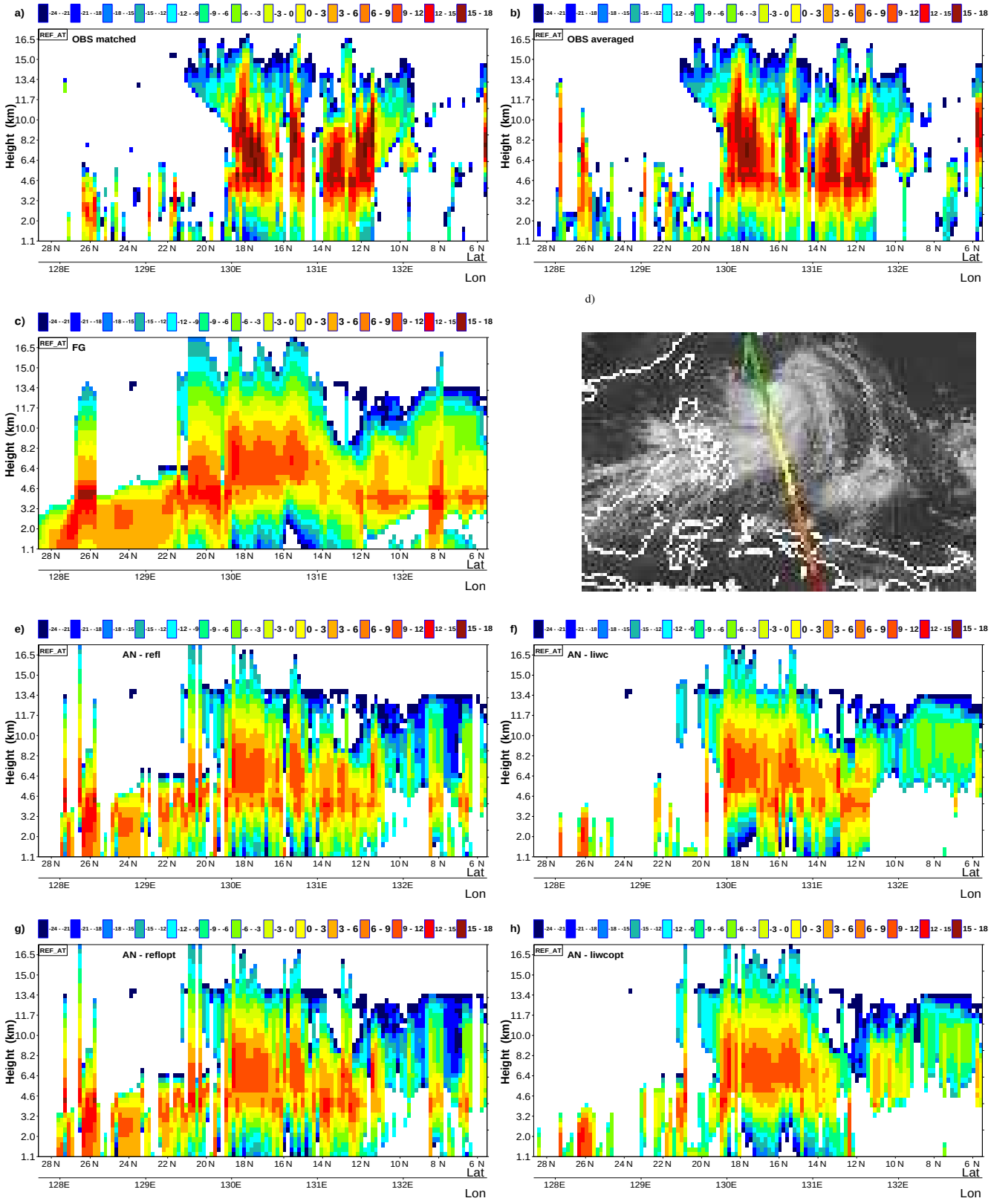


Figure 4.21: *Cloud radar reflectivity (in dBZ) for the situation (d) on 16 September 2007 over (cyclone Wipha) - (a) CloudSat observations from 94 GHz radar closest to the model grid-point (matched), (b) observations averaged over the model grid-box, (c) model first guess (FG) and 1D-Var retrievals using averaged observations of: (e) cloud reflectivity (AN-refl) alone and combined with MODIS cloud optical depth (AN-reflopt), (f) cloud liquid and ice water contents from CloudSat (AN-liwc) alone and (h) combined with cloud optical depth (AN-liwcopt).*

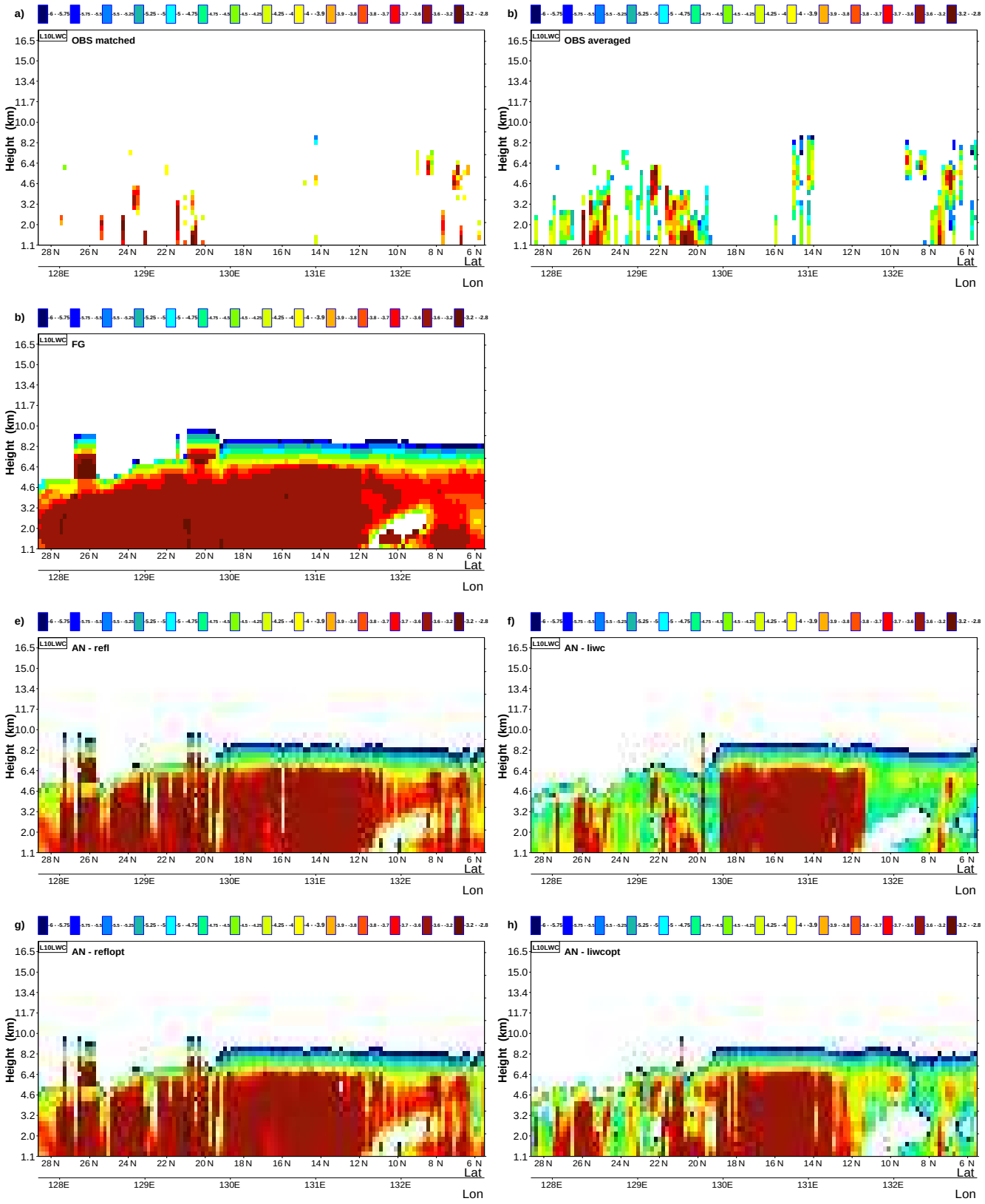


Figure 4.22: *Same as Fig. 4.21, but for cloud liquid water content (in kg m^{-3} using logarithmic scale).*

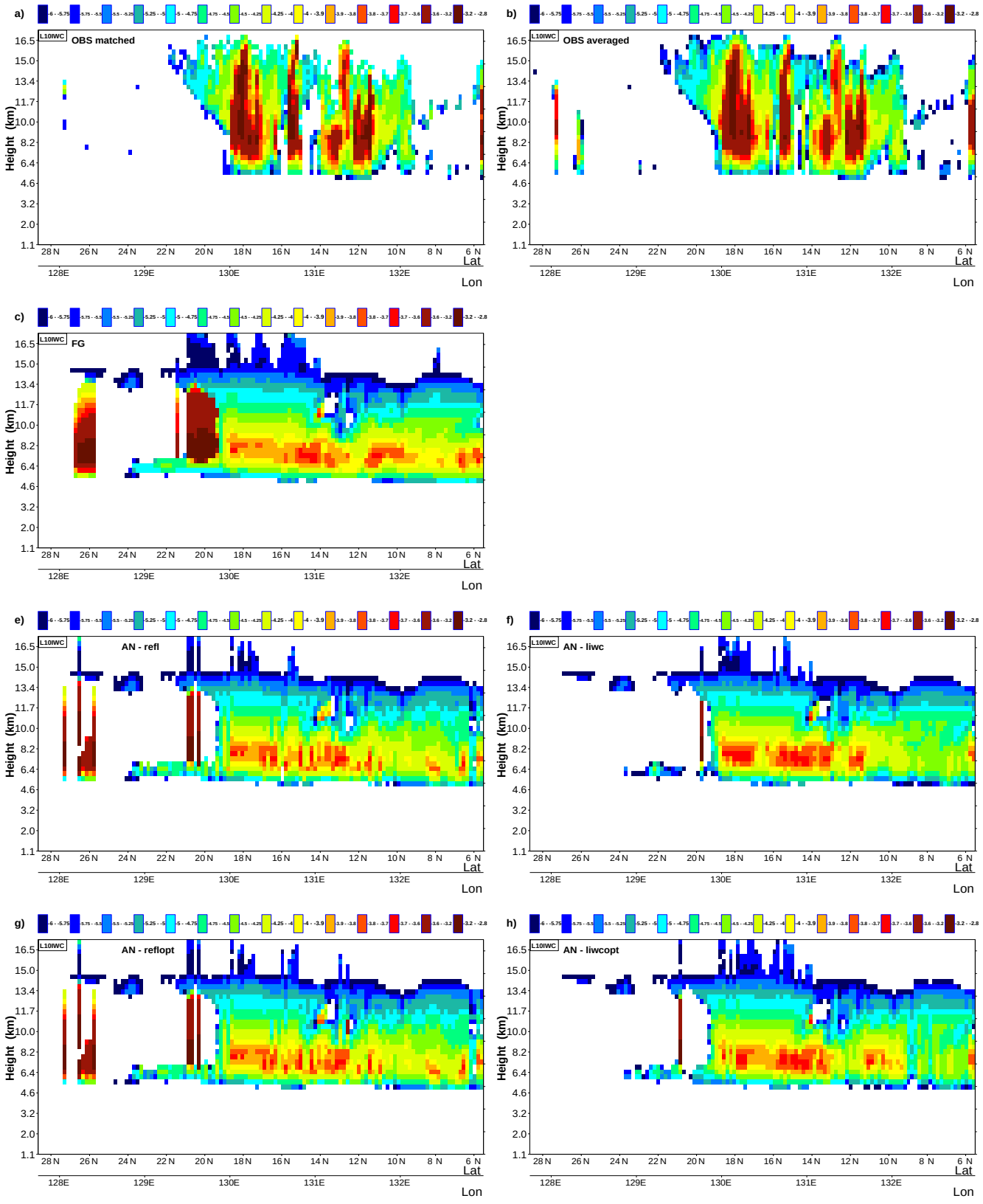


Figure 4.23: *Same as Fig. 4.22, but for cloud ice water content.*

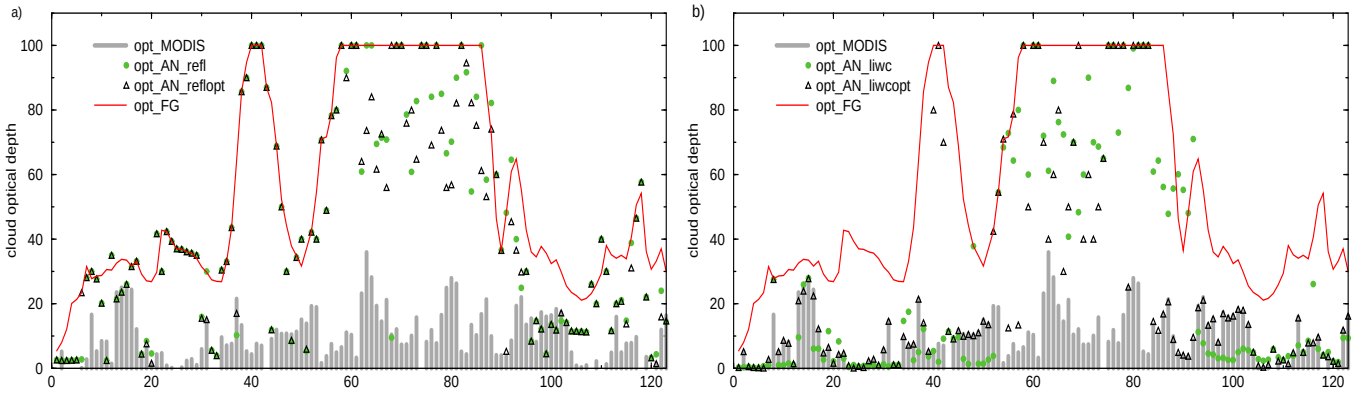


Figure 4.24: *Comparison of the model first guess (red line) and the analyzed values from 1D-Var retrievals using averaged observations of (a) cloud reflectivity (green circle) alone and combined with cloud optical depth (black triangle), (b) cloud liquid and ice water contents (green circle) alone and combined with cloud optical depth (black triangle) against MODIS cloud optical depth observations averaged over the model grid-box (grey shading). Situation of 16 September 2007.*

	MATCHED OBS						AVERAGED OBS					
	reflectivity				cloud iwc		reflectivity				cloud iwc	
	bias	stdv	mae	rms	bias	stdv	bias	stdv	mae	rms	bias	stdv
FG	-8.51	17.76	15.20	19.69	-0.97E-3	0.142	-8.84	17.41	15.26	19.53	-0.45E-3	0.134
AN-refl	-3.27	15.55	11.38	15.88	-5.94E-3	0.129	-3.05	14.56	10.51	14.88	7.75E-3	0.120
AN-reflopt	-3.46	15.65	11.44	16.03	6.07E-3	0.129	-3.48	14.61	10.64	15.02	7.51E-3	0.120
AN-liwc	-1.63	14.38	11.16	14.47	1.54E-2	0.101	-0.07	13.90	10.54	13.90	1.58E-2	0.092
AN-liwcopt	-1.65	15.44	11.94	15.53	1.49E-2	0.106	-0.57	15.59	11.92	15.60	1.61E-2	0.093

Table 4.5: *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 CloudSat cloud reflectivity (in dBZ) and ice water content (in g m^{-3}) observations closest to the model grid-box (matched) or averaged over the model grid-box. 123 profiles included in the statistics for the case of 16 September 2007.*

	MATCHED OBS		AVERAGED OBS	
	opt.depth		opt.depth	
	bias	stdv	bias	stdv
FG	-49.43	30.38	-47.16	29.31
AN-refl	-42.22	32.22	-34.96	31.90
AN-reflopt	-40.12	31.70	-33.77	31.30
AN-liwc	-29.13	36.18	-19.69	33.83
AN-liwcopt	-24.10	35.53	-17.70	31.46

Table 4.6: *Same as Table 4.5, but for bias and standard deviation of the departures from the MODIS cloud optical depth.*

4.3.4 Summary of all cases

The statistical evaluation of the 1D-Var performance has been done for all performed experiments together and the results from this evaluation are summarized in Fig. 4.25 - 4.29 and in Tables 4.7 - 4.8.

Figure 4.25 shows probability distribution functions (PDF) of the first-guess departures for matched and averaged observations of cloud radar reflectivity (a) in $mm^6\ m^{-3}$ (the unit, in which it was assimilated) and cloud ice water content (b). The PDF of the FG departures is slightly better when computed with respect to the observations averaged than matched. The differences are more significant for cloud ice water content, in the case of observation matched with the model, the peak of distribution is not in the middle. PDFs for both, cloud reflectivity and cloud ice water content do not have a perfect Gaussian shape since there is a tail for positive departures. This should be eliminated in the future by building quality control and a suitable bias correction.

As shown in Fig. 4.26 (for the first-guess and analysis departures of cloud radar reflectivity) and in Fig. 4.27 (for the departures of cloud ice water product), the PDFs of analysis departures become more narrow compared to the FG departures, indicating that the analyses are getting closer to the observations, though the tail present already in the FG departures remains nearly unchanged.

Figures 4.28 - 4.29 and Table 4.7 summarize the results from the different assimilation experiments using either matched or averaged observations obtained for all situations used in our experimentation with a total number of 1787 profiles included in the statistics. The outcomes from these statistics are similar to those for the individual cases presented above, i.e.:

- the AN departures are usually smaller when assimilating observations averaged over the model grid-box compared to matched observations;
- analyses obtained from the assimilation of cloud optical depth in combination with cloud reflectivity differs only slightly from the assimilation of cloud reflectivity alone (see also Figures 4.28 - 4.29 b, e);
- the impact of cloud optical depth on the analysis is larger when combined with the cloud liquid and ice water content, which could be a consequence of frequently missing retrievals of liquid water content (Fig. 4.28 - 4.29 c, f);
- when assimilating cloud liquid and ice water contents, the impact of these observations is getting rapidly smaller between 4.5 km and 8 km, and usually does not appear above 9.5 km (even in the tropics) which could indicate a small sensitivity of the 1D-Var system to ice water content due to the small sensitivity of the observation operator (moist parametrization schemes) to this quantity (especially to small amounts of ice water content) or due to too large observation errors associated with the cloud ice water product;
- the bias of the analysis departures for cloud ice water content is always larger than that of first guess departures (though the standard deviation is decreased), indicating that the 1D-Var system is probably more efficient in removing hydrometeors than in increasing their amount.

For the validation of the 1D-Var performance, two sets of independent observations (i.e. observations not assimilated in the system) have been used. Using MODIS cloud optical depth as an independent data type for evaluation, the **refl** and **liwc** experiments could be evaluated for all situations. All different assimilation experiments could be compared against NEXRAD precipitation data, but only for two situations over USA (320 cases included in the statistics) where precipitation observations are available. Comparisons of the analyzed values of cloud optical depth and hourly precipitation rates against independent observations are summarized in Table 4.8. Only values printed in black should be considered for a fully independent validation. For the comparison against cloud optical depth, the statistics have also been run for the cases

when the model values (first-guess or analyzed) were smaller or equal to 50, that is the maximum range of this observational product. The analyses get closer to the MODIS cloud optical depth by assimilating either cloud radar reflectivity or cloud liquid and ice water contents. The **liwc** experiments seem to perform better than the **refl** ones, even though, when comparing only the cloud optical depth ≤ 50, the bias of **refl** is smaller, while the standard deviation remains larger than in the **liwc** experiments. The validation of the different analyses against NEXRAD precipitation data (in mm h^{-1}) reveals that there are small differences in the performance of the **refl** and **liwc** experiments when using matched observations and the best results are obtained when cloud liquid and ice water contents are combined with cloud optical depth (**liwcopt**), while adding cloud optical depth to reflectivity observations does not improve the results significantly. When averaged observations are used, the differences between **refl** and **liwc** are larger and more significantly in favour of **liwc** experiment. Generally, the analyzed values of hourly precipitation rates for all experiments are closer to the observations than the first-guess precipitation.

The performed assimilation experiments show that the analyses obtained by assimilating either level-1 (cloud reflectivity) or level-2 (liquid and ice water contents) data get closer not only to the assimilated observations but, what is important, also to the independent observations.

When compared to independent observations, the **liwc** assimilation seems to perform better than the **refl** assimilation. Analysing the performance of the different assimilation experiments, it was observed that the **refl** assimilation generally modifies more profiles, but has a tendency to "over-shoot" adjustment to the observations sometimes, which results in a degradation of the assimilation performance. This indicates that a better screening of the reflectivity observations should be performed and/or that the reflectivity observation operator should undergo more testing to identify and, when present, to remove possible non-linearities or discontinuities in the operator. The moist parametrization schemes (used as the only operators in the **liwc** experiments) have been widely tested for such effects since they have been used in our operational 4D-Var system for many years.

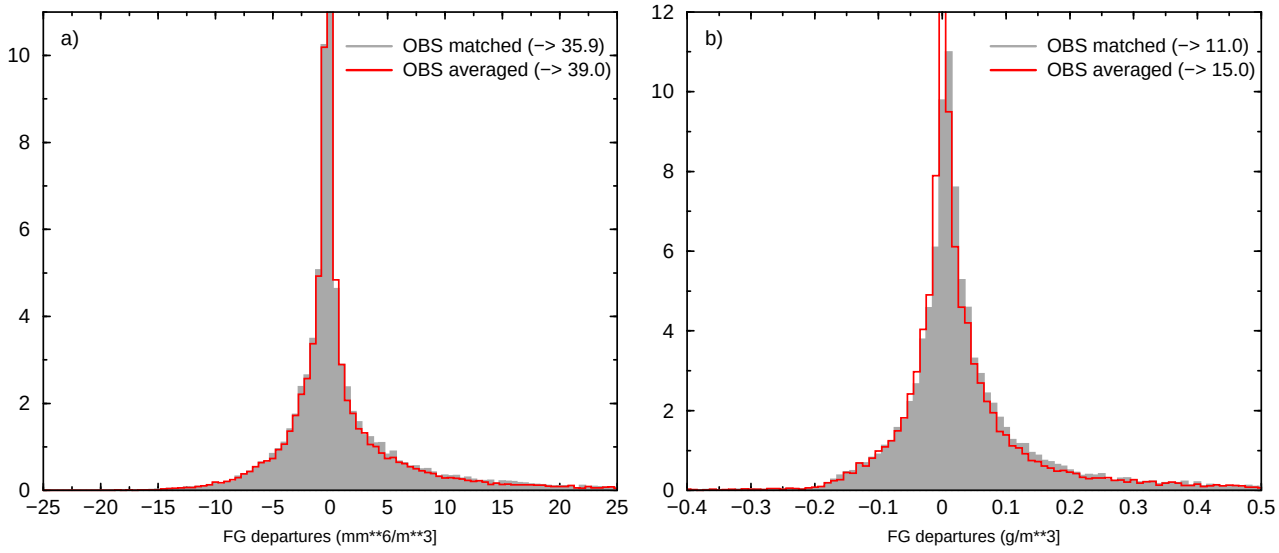


Figure 4.25: *Probability distribution function (PDF) of first-guess departures for matched (grey shading) and averaged (red line) observation of: (a) cloud radar reflectivity (in mm^{-6} m^{-3}) and (b) ice water content (in g m^{-3}). Values in the bracket indicate the peak of distribution. Summary for all experimental cases.*

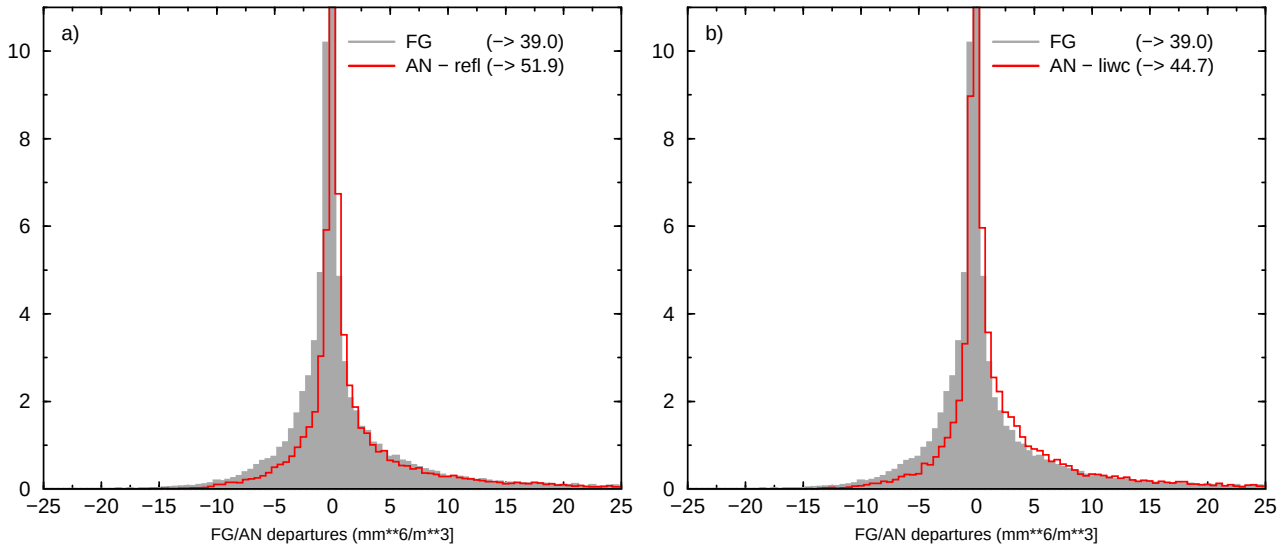


Figure 4.26: *Probability distribution functions of first-guess (grey shading) and analysis (red line) departures for cloud radar reflectivity (in $mm^{-6}\ m^{-3}$) coming from 1D-Var retrievals using averaged observations of (a) CloudSat radar reflectivity and (b) cloud liquid and ice water contents. Summary for all experimental cases.*

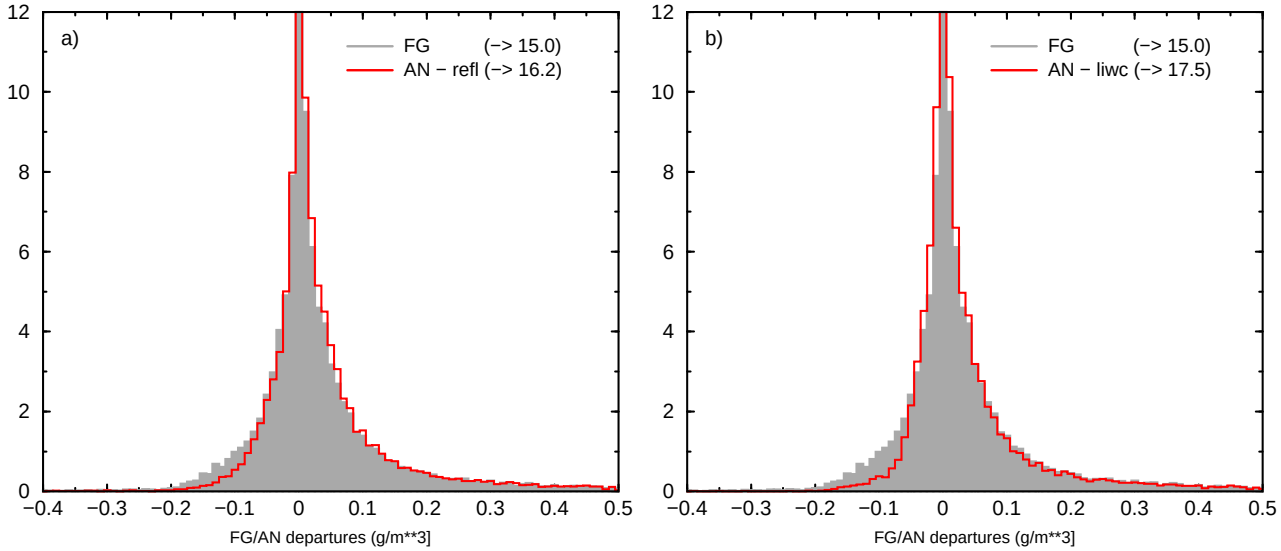


Figure 4.27: *Same as 4.26, but for the ice water content (in g m^{-3}).*

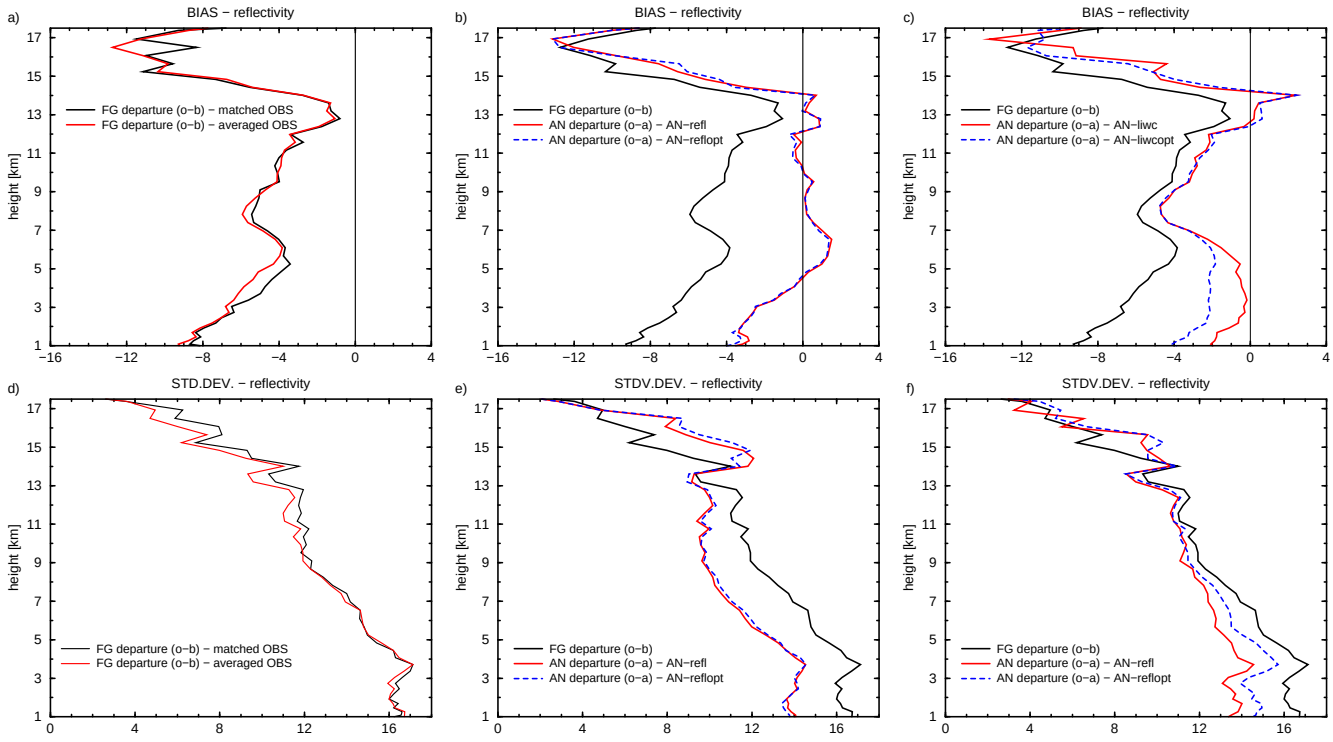


Figure 4.28: *Profiles of radar reflectivity (dBZ) bias and standard deviation for (a, d) the first-guess departures from the matched (black line) or averaged (red line) observations and for analysis departures from 1D-Var retrievals using (b, e) averaged observations of cloud radar reflectivity (red line) alone or combined with cloud optical depth (blue dashed line) and (c, f) averaged observations of cloud liquid and ice water content (red line) alone or combined with cloud optical depth (blue dashed line). Summary for all experimental cases.*

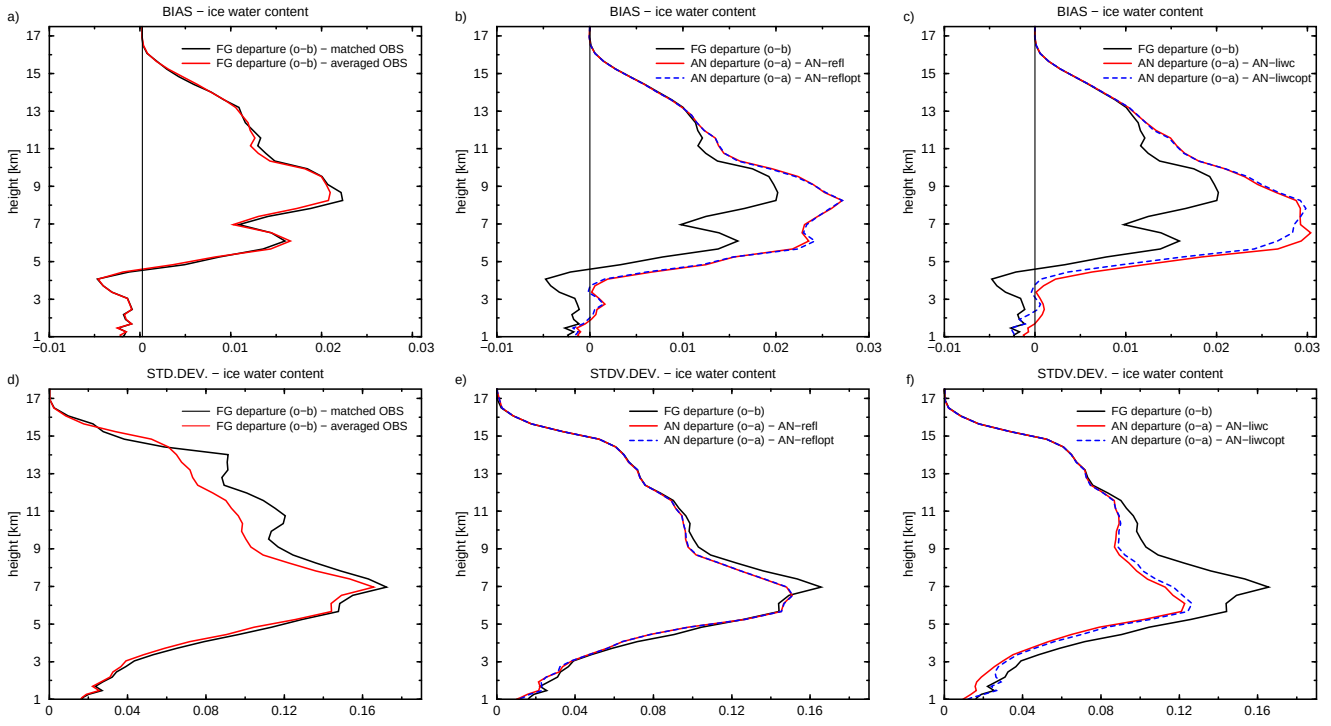


Figure 4.29: *Same as Fig. 4.28, but for cloud ice water content (in g m^{-3}).*

	MATCHED OBS						AVERAGED OBS					
	reflectivity				cloud iwc		reflectivity				cloud iwc	
	bias	stdv	mae	rms	bias	stdv	bias	stdv	mae	rms	bias	stdv
FG	-5.35	14.95	11.55	15.88	5.13E-3	8.30E-2	-5.77	14.87	11.66	15.95	4.89E-3	7.47E-2
AN-refl	-0.31	12.73	8.42	12.74	8.08E-3	7.85E-2	-0.86	12.45	8.09	12.48	7.42E-3	7.19E-2
AN-reflopt	-0.33	12.78	8.52	12.79	8.11E-3	7.85E-2	-0.98	12.46	8.17	12.50	7.36E-3	7.20E-2
AN-liwc	-2.12	12.73	9.41	12.90	8.71E-3	7.11E-2	-1.88	12.74	9.28	12.87	8.31E-2	6.22E-2
AN-liwcopt	-2.50	13.32	9.87	13.55	8.47E-3	7.26E-2	-2.79	13.49	10.03	13.78	7.83E-2	6.39E-2

Table 4.7: *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 CloudSat cloud reflectivity (in dBZ) and ice water content (in g m^{-3}) observations closest to the model grid-box (matched) or averaged over the model grid-box. 1787 profiles were included in the statistics for all experiments.*

	MATCHED OBS						AVERAGED OBS					
	opt.depth		opt.depth≤50		precip		opt.depth		opt.depth≤50		precip	
	bias	stdv	bias	stdv	bias	stdv	bias	stdv	bias	stdv	bias	stdv
FG	-25.15	33.72	-3.09	15.44	-0.420	1.597	-22.55	31.48	-4.54	14.63	-0.420	1.597
AN-refl	-18.37	32.62	1.07	15.08	-0.298	1.452	-16.09	29.17	-1.06	14.15	-0.334	1.415
AN-reflopt	-16.74	31.54	0.38	14.58	-0.273	1.467	-14.09	27.89	-1.30	13.53	-0.288	1.477
AN-liwc	-14.77	31.45	3.01	12.80	-0.250	1.410	-9.67	27.47	3.19	10.38	-0.145	1.363
AN-liwcopt	-11.26	28.29	1.48	6.36	-0.164	1.411	-9.09	25.41	0.73	6.50	-0.092	1.483

Table 4.8: *Same as Table 4.7, but for bias and standard deviation of the departures from MODIS cloud optical depth (see text for explanation) and from NEXRAD precipitation data (in mm h^{-1}). 1787 profiles were included in the statistics for optical depth and 320 cases for precipitation.*

5 1D-Var experiments for CALIPSO aerosol observations

Two test cases were run in order to assess the functioning and the capabilities of the newly designed 1D-Var aerosol assimilation system.

5.1 Experimental setup

The assimilation of CALIOP backscatter data was restricted to updating only the aerosol fields while other fields were left unchanged by the analysis procedure. Unlike cloud variables, which are strongly linked to the temperature and humidity fields in the model, aerosols are treated as passive tracers by the forecast scheme. They are only altered through some well defined sources and sinks and do not interact with other model variables.

While formally the lidar observation operator shows sensitivities to all background fields mentioned in section 3.1, the comparably large uncertainties of the aerosol fields make the simultaneous assimilation of, e.g., temperature increments unrealistic.

Furthermore, as there are 11 aerosol species (i.e., 11 prognostic variables) represented in the forecast model the problem when assimilating only one observed field is massively underconstrained. We therefore follow Benedetti *et al.* (2009) and increment only the total aerosol mixing ratio (i.e., the sum over the mixing ratios for each individual species) while the fractional contribution of each species is kept constant at the value given by the model background.

5.2 Situations used in experimentation

The situations studied for this report are both taken from the first 12 hours of the 16 April 2008 for which the aerosol forecasts described in section 3.1 had been produced. Figure 5.1 shows the satellite tracks from which the situations have been chosen. The first case (case 1) included a large, mainly cloud free region over the Saharan desert. The second situation (case 2) is located over the east coast of Asia, mostly over the west Pacific intersecting the Japanese islands. This second case exhibited strong concentrations of aerosols in the direct neighbourhood of clouds of different types and altitudes.

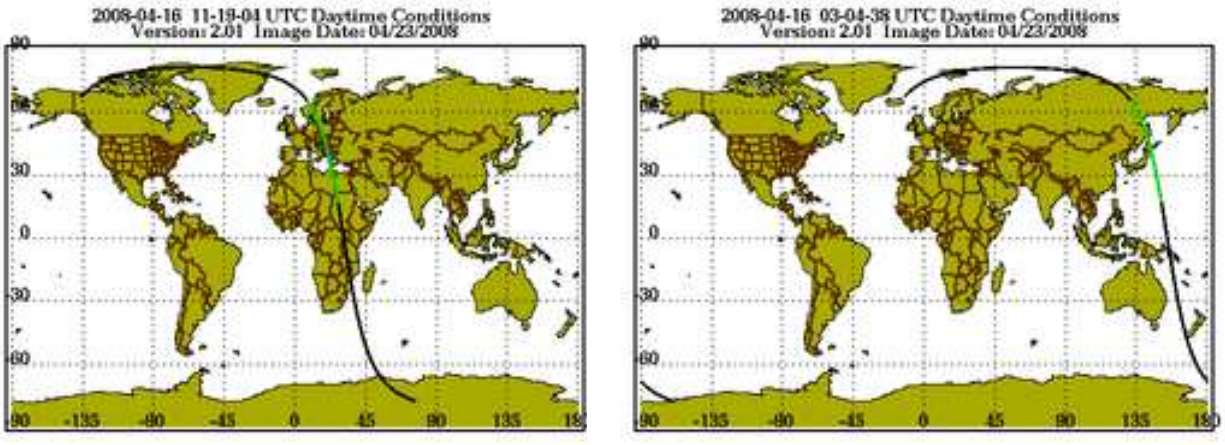


Figure 5.1: *Satellite tracks from which the test cases used for experimentation have been chosen.*

5.3 Results

5.3.1 Case 1

The first case corresponds to the southern portion of the green marked satellite track shown in the left panel of Figure 5.1. The top graph of Fig. 5.2 shows the corresponding CALIPSO feature mask while the two lower graphs give the corresponding 532 nm backscatter coefficient with different types of cloud screening applied in the two graphs, respectively. For the middle graph the cloud top heights from the 1km product have been used as a cloud screening criterion. Comparison with the feature mask indicates that some of the backscatter signal in this graph has been identified as cloud by the feature mask. It has been therefore decided to use a more restrictive cloud screening criterion by taking the cloud top height information from the 5km product, instead. The result is shown in the bottom graph of Fig. 5.1 for which the backscatter signal in the regions where the feature mask reports clouds has been strongly reduced.

The model equivalent of the backscatter coefficient for the first guess is given in the upper graph of Fig. 5.3. While the first guess signal is generally weaker and more confined to the ground, the structure of the main regions where most of the aerosols are present is quite well represented. This is impressive regarding the fact that the modelled aerosols had been spun up without any data assimilation of aerosol data.

The bottom graph of Fig. 5.3 shows the corresponding (model equivalent) backscatter field for the analysis. Generally, the model backscatter has been strongly increased and drawn substantially closer to the observed profile (bottom of Fig. 5.2). As seen in Fig. 5.4, also the first guess model aerosol concentration (top graph) has been, in most places, substantially increased by the analysis (bottom graph).

It should, however be noted that, while typically increasing the intensity of the model backscatter, a more careful comparison between first guess and analysis in Fig. 5.3 also identifies some regions where the analysis actually decreases the strength of the model backscatter. This demonstrates the capability of 1D-Var to adjust the model observations in both directions.

5.3.2 Case 2

The satellite track for the second case can be obtained from the right graph of Figure 5.1. The corresponding CALIPSO feature mask and backscatter fields are shown in Fig. 5.5. This case exhibits a much larger population of clouds of different sizes and altitudes. While, according to the feature mask, no aerosols can be seen in the vicinity of the very high clouds (around 10km altitude) most of the lower clouds seem to have regions in their neighbourhood have been classified as aerosols. Similar to case 1, the cloud screening with the 1 km level-2 product (middle graph of Fig. 5.5) seems to be insufficient and only the screening with the 5km product removes the strong backscatter signals in regions where the CALIPSO feature mask (top graph) shows clouds without any aerosols next to them.

Again, as in case 1, the first guess (model equivalent) backscatter coefficient shown in the top graph of Fig. 5.6 seems to show a large part of the observed features but with smaller intensity while the analysis (bottom graph of Fig. 5.6) increases the intensity and draws the backscatter profiles substantially closer to the observations. Also, the corresponding model aerosol concentrations (Fig. 5.7) are strongly increased by the analysis procedure. This is particularly true for the region around 40^0 latitude which is strongly polluted according to the feature mask but for which the first guess aerosol field shows comparably small numbers. The analysis increases these aerosol concentrations in this region so that the analysed aerosol field seems to agree much better with the feature mask.

5.4 Summary of all cases

The case studies demonstrate that the 1D-Var system is a powerful tool which drives the model equivalent observations substantially towards the CALIPSO backscatter data while incrementing (i.e. updating) the model aerosol fields accordingly. This indicates that the software development for this part of the work package has been successful and that the variational data assimilation system is in principle able to benefit from the new data types.

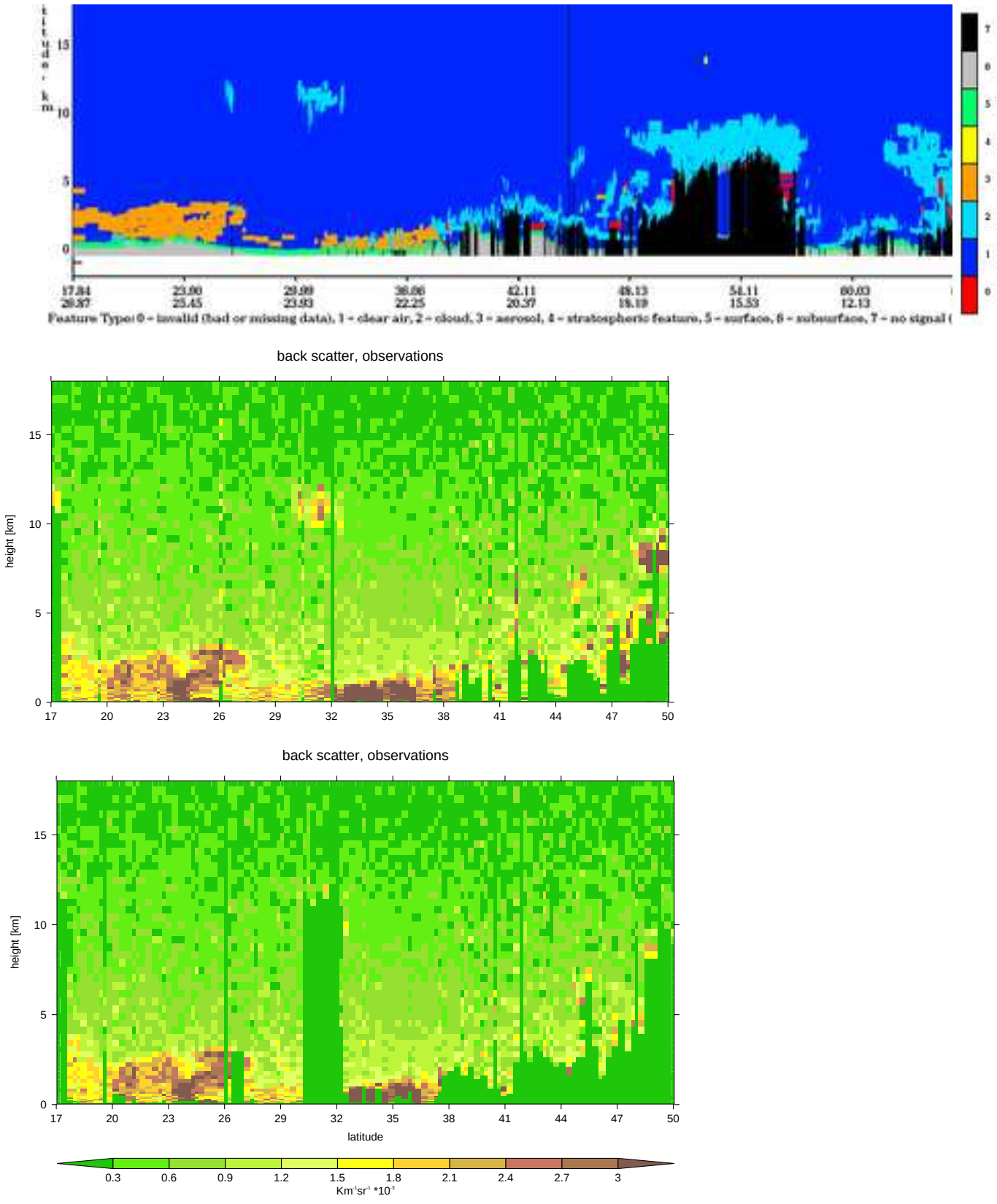


Figure 5.2: *Top: CALIPSO feature mask. Middle and bottom: Aerosol affected CALIPSO backscatter 532 nm. Cloud screening was performed with the CALIPSO level 2 cloud layer descriptor from the 1km product (middle graph) and the 5km product (bottom graph), respectively.*

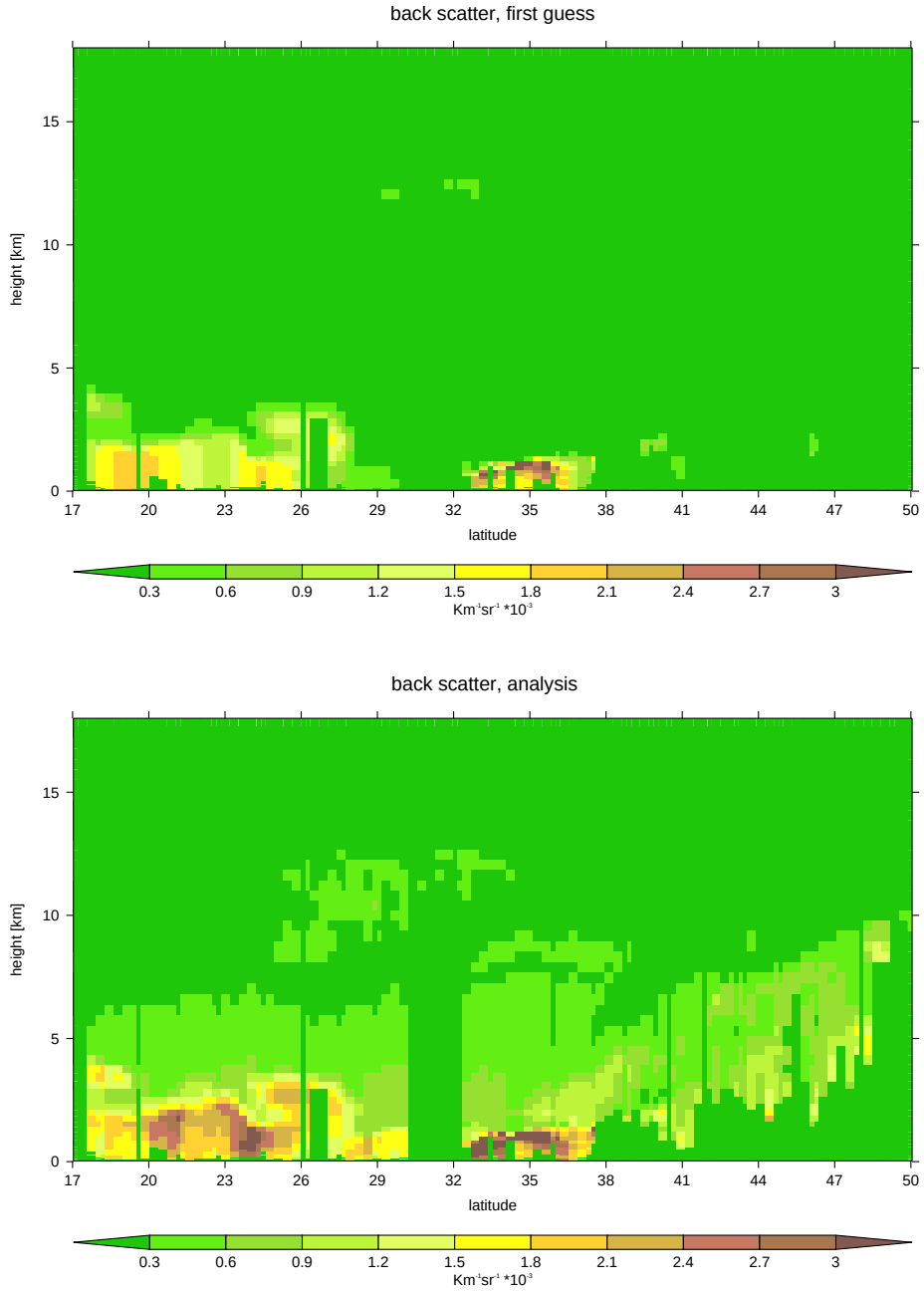


Figure 5.3: *First guess (top) and analysis (bottom) for the model equivalent of 532 nm lidar backscatter.*

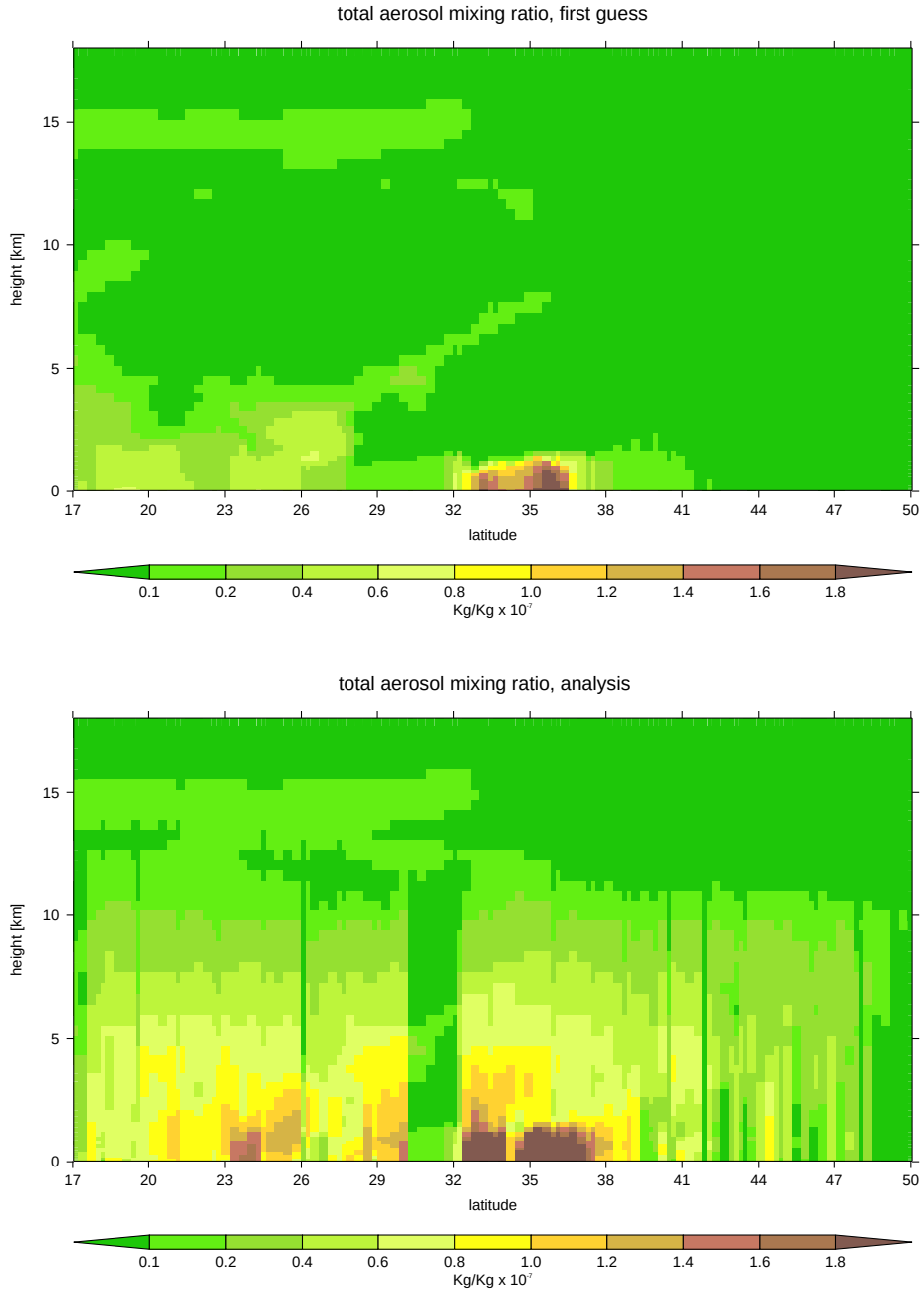


Figure 5.4: *First guess (top) and analysis (bottom) for the total aerosol mixing ratio.*

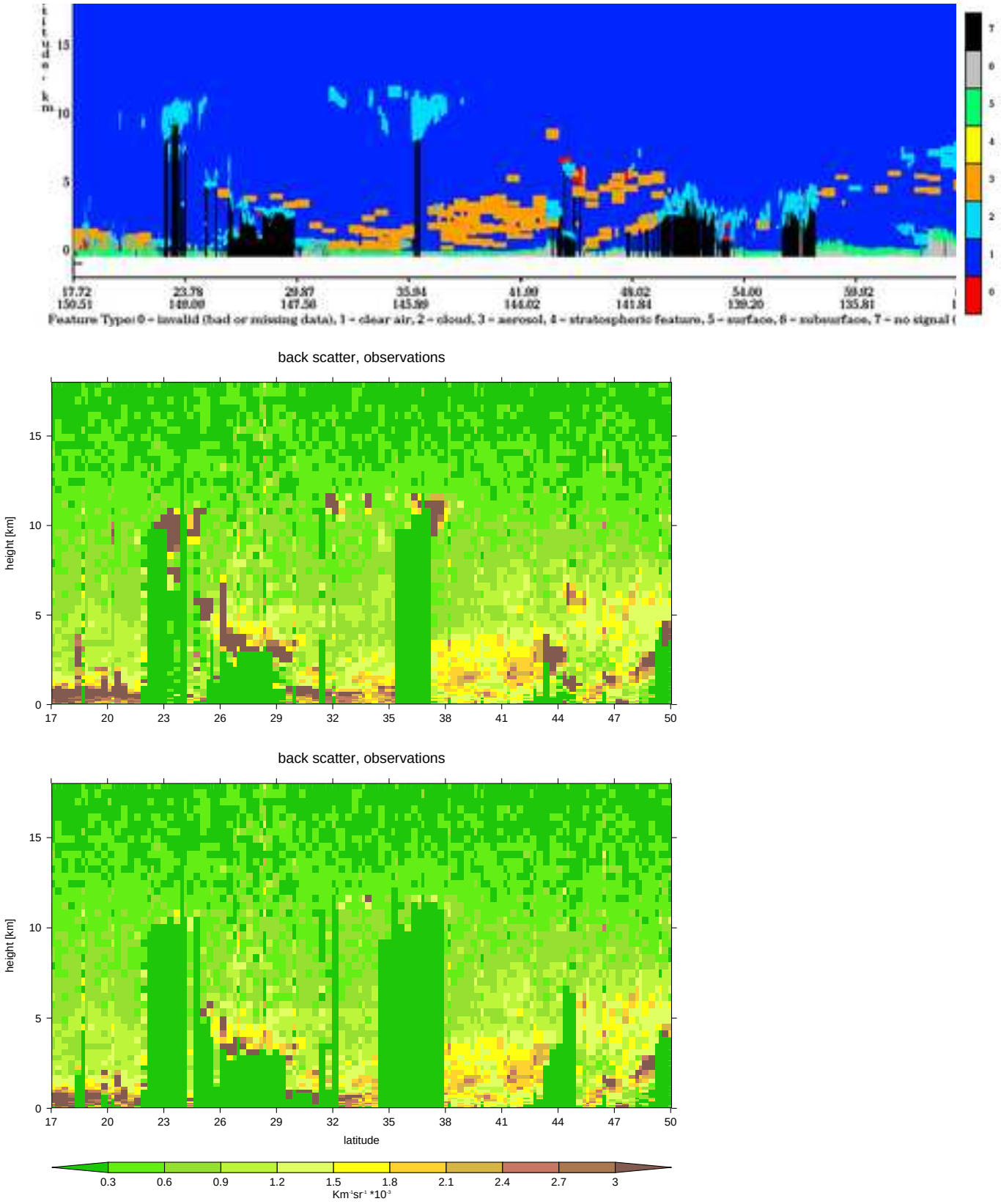


Figure 5.5: *The same as Fig.5.2 but for case 2.*

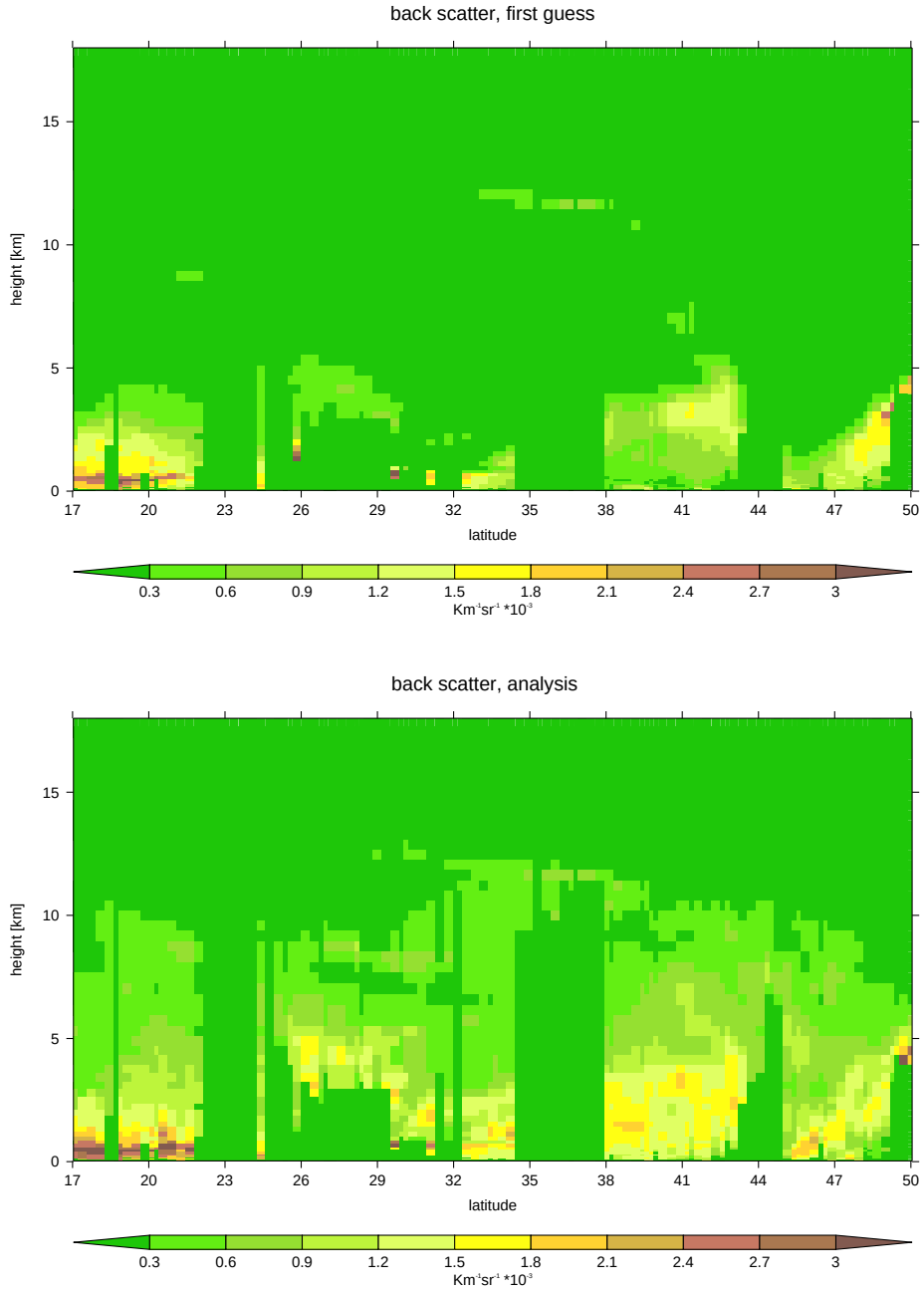


Figure 5.6: *The same as Fig.5.3 but for case 2.*

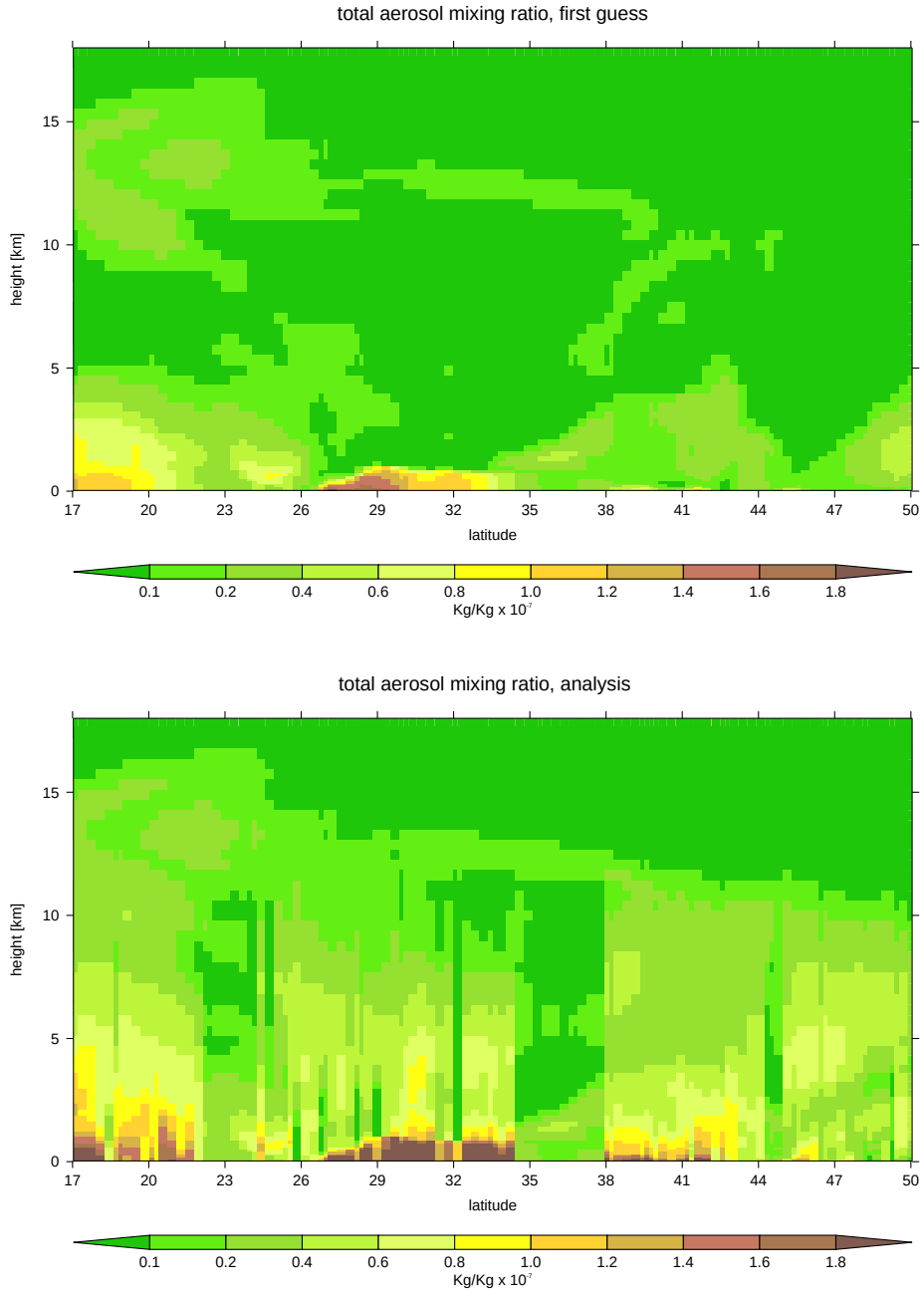


Figure 5.7: *The same as Fig.5.4 but for case 2.*

6 Conclusion and perspectives

In this work package, feasibility studies have been performed in order to develop strategies for radar and lidar assimilation. Since in the past it was proven that the 1D-Var approach can provide very useful experience on how to assimilate new types of observations, 1D-Var systems have been built for the assimilation of CloudSat reflectivities and CALIPSO backscatter cross-sections (level-1 products), as well as for retrieved cloud liquid and ice water contents from CloudSat (level-2 products). Additionally, the CloudSat profile observations have also been combined with MODIS cloud optical depth which is a column integrated measurement with larger spatial coverage. This combination has been tested as a possible way to eliminate a possible representativeness problem when using measurements with very limited spatial coverage as provided by CloudSat.

In order to reduce the computational cost of the 1D-Var experimentation and thus to be able to run more experiments for defining the most suitable approach for the assimilation of cloud/aerosol profile observations, adjoint versions of the cloud reflectivity model and the aerosol backscatter forward operator have been coded and combined with already existing adjoint versions of all other operators used by the 1D-Var system. Using the adjoint technique instead of the finite difference approach for the computation of the transpose of the observation operator required by variational assimilation system reduces the computational cost by approximately 60 times.

1D-Var experiments using either CALIPSO 532 nm backscatter data or different CloudSat observations have been performed for several selected situations and demonstrated that the method is principally able to benefit from the new types of observational data. For CloudSat observations this was shown through a statistical evaluation of the 1D-Var performance which included a study of the impact that different combinations of observations have on the analysis. Particularly, using a CloudSat profiling information in combination with column integrated quantity has revealed that assimilating cloud optical depth on top of cloud reflectivity yields only small improvements (compared to the assimilation of the cloud reflectivity only), while the impact of cloud optical depth on the analysis is larger when combined with cloud liquid and ice water contents. There can be several reasons for such behaviour, such as often missing cloud liquid water content data (the cloud optical depth would then supply missing observational information) and/or larger observation errors for liquid and ice water contents than for cloud optical depth. The errors for cloud reflectivity and optical depth may be more balanced and also observational coverage for these quantities is better.

When comparing the performance of assimilation systems for level-1 and level-2 products, it has been observed that the impact from the assimilation of cloud liquid and ice water contents is limited in the vertical (getting fast smaller between 4.5 and 8 km and usually not exceeding 9.5 km, even in the Tropics). This could indicate a small sensitivity of the 1D-Var system to ice water content due to the small sensitivity of the observation operator (moist parametrization schemes) to this quantity (especially to small amounts of ice water content) or due to too large observation errors for the cloud ice water product. On the contrary, the cloud reflectivity assimilation modifies more profiles and has an impact in larger vertical extent, but sometimes has a tendency to "over-shoot" adjustments towards the observations which, as a consequence, degrades assimilation performance. That suggests that maybe a better screening of reflectivity observations should be done and/or the reflectivity observation operator should undergo more testing to identify and, when present, to remove non-linearities or discontinuities in the operator. The moist parametrization schemes (used as the only operators in the assimilation of cloud liquid and ice water contents) have been widely tested for such effects since they are already used in our operational 4D-Var system for several years. Over-adjustment of analyses with respect to the observations when assimilating the cloud radar reflectivity could be a reason, while the assimilation of cloud liquid and ice water contents perform slightly better when compared to the independent observations (i.e. observations not assimilated in the 1D-Var system).

Analysis increments of temperature and specific humidity have been also evaluated since they can provide information about impact of the assimilated observations on the control variables of the 1D-Var system, i.e. temperature and specific humidity. This evaluation shows that both increments are modified by the assimilation of cloud related observations and therefore pseudo-observations of temperature and specific humidity profiles from 1D-Var retrievals should be included in the 4D-Var system instead of the originally planned pseudo-observations of the total column water vapour (used before operationally in a 1D-Var+4D-Var approach for precipitation observations).

Overall, the results clearly show that the analyses obtained by assimilating either level-1 or level-2 products get closer not only to the assimilated, but also to independent observations. This indicates that analyses could benefit from the assimilation of these types of observations and therefore effort should be made to further explore the possibilities of their use in the assimilation system of NWP models.

Also for aerosols the analysed model observations were drawn substantially towards the analysis by the 1D-Var system. While this demonstrates the technical capability of the newly developed 1D-Var system, tests with independent observations would be desirable for future studies. Also the combined assimilation of lidar profiles with other observations (such as imager data as, e.g. from MODIS) or of lidar data of different channels (e.g. 532 nm and 1064 nm) could be very useful for future applications of the method. Combining different observations might allow to adjust not only the total aerosol mixing ratio but also to improve the adjustment of the fraction of the individual aerosol species Ackermann (1998); Huneeus and Boucher (2007).

While at present the model aerosol field is strongly underconstrained with regard to observations it is remarkable how the first guess data used in this study seems to reproduce the observed aerosol structures. Indeed, the model aerosols employed in this study had been obtained from simulations where the aerosol field was run in a free wheeling mode (i.e., spinning up from zero over a period of one month without any data assimilation). The structure of the model aerosols were therefore completely determined by the physical parametrisations of the sources and sinks (and the wind field of the ECMWF forecasting model) which apparently have remarkable skill. Enhancing this through improved data assimilation yields exciting perspectives for the future of air quality monitoring and forecasting.

Improvements in several areas are still necessary. For example, the probability distribution functions of the first-guess departures for cloud reflectivity and the cloud ice water content do not have a perfect Gaussian shape (requirement for a well-behaved variational assimilation). This should be eliminated in the future by building a suitable bias correction. A better screening of observations and improved observation error definition (including representativeness error) may be required. The observation operator for cloud reflectivity should be re-tested for non-linearities which could have a negative impact on assimilation performance.

Acknowledgements

The NASA CloudSat Project is kindly acknowledged for providing the CloudSat data. CALIPSO data and the CALIPSO cloud mask graphics were obtained from the NASA Langley Research Center Atmospheric Science Data Center. The authors are also grateful to the Goddard Earth Sciences Data and Information Services Center (GES DISC) for providing MODIS data. NCEP for producing and JOSS/UCAR for providing the Stage IV precipitation observations used in this study are also acknowledged. We would further like to thank Peter Bauer, Angela Benedetti and Philippe Lopez for proof reading and their constructive suggestions concerning this manuscript. The aerosol assimilations presented here were greatly supported by Jean-Jaque Morcrette (who provided us with the forward operator, its description and initial conditions for aerosol modeling) and Angela Benedetti (who gave us the aerosol background error statistics)

A List of Acronyms

1D-/4D-Var	One-/Four-Dimensional Variational assimilation
ARM	Atmospheric Radiation Measurement
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization on the CALIPSO satellite
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CAPE	Convective Available Potential Energy
CloudSat	NASA's cloud radar mission
EarthCARE	Earth, Clouds, Aerosols and Radiation Explorer
ECMWF	European Centre for Medium Range Weather Forecasts
GEMS	Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data
GES DISC	Goddard Earth Sciences Data and Information Services Center
HC	hydrometeor content
IFS	Integrated Forecast System
JOSS/UCAR	Joint Office for Science Support of the University Corporation for Atmospheric Research
M1QN3	quai-Newton optimization routine
MACC	Monitoring Atmospheric Composition and Climate
mae	mean absolute error
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NEXRAD	NEXt-generation RADar
NMC	National Meteorological Center
NWP	Numerical Weather Prediction
rms	root mean square error
stdv	standard deviation
TCWV	total column water vapour
ZmVar	Z (reflectivity) Model for Variational assimilation of ECMWF

References

Ackermann, J., 1998: The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study, *Journal of atmospheric and oceanic technology*, **15**(**4**), 1043–1050.

Baldwin, M. E. and K. E. Mitchell, 1996: The NCEP hourly multi-sensor U.S. precipitation analysis, in *Preprints 11th Conf. on Numerical Weather Prediction, Norfolk, VA (USA), 19–23 August 1996*, pp. J95–J96.

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, *Qurt. J. Roy. Meteor. 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, *Qurt. J. Roy. Meteor. Soc.*, **132**, 2307–2332.

Bayler, G. M., R. M. Aune, and W. H. Raymond, 2000: NWP cloud initialization using GOES sounder data and improved modeling of nonprecipitating clouds, *Mon. Wea. Rev.*, **128**, 3911–3920.

Benedetti, A. and M. Fisher, 2007: Background error statistics for aerosols, *Quart. J. Roy. Meteor. Soc.*, **133**(623), 391–406.

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

Benedetti, A. and M. Janisková, 2008: Assimilation of MODIS cloud optical depths in the ECMWF model, *Mon. Wea. Rev.*, **136**, 1727–1746.

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J. W. Kaiser, *et al.*, 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, *J. Geophys. Res.-Atmospheres*, **114**(D13), D13205, doi:10.1020/2008JD011115.

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

Crewell, S., and Co-authors, 2004: The Baltex bridge campaign: An interated approach for a better understanding of clouds, *Bull. Amer. Meteor. Soc.*, **85**, 1565–1584.

Fillion, L. and J.-F. Mahfouf, 2003: Jacobians of an operational prognostic cloud scheme, *Mon. Weather Rev.*, **131**, 2838–2856.

Fu, Q., 1996: An accurate parameterization of the solar radiative properties of cirrus clouds for climate models, *J. Climate*, **9**, 2058–2082.

Fulton, R. A., J. P. Breidenbach, D. J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm, *Weather and Forecasting*, **13**, 377–395.

Gilbert, J.-C. and C. Lemaréchal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms, *Math. Programming*, **45**, 407–435.

Holtslag, A. A. M. and C.-H. Moeng, 1991: Eddy diffusivity and countergradient transport in the convective atmospheric boundary layer, *J. Atmos. Sci.*, **48**, 407–435.

Huneeus, N. and O. Boucher, 2007: One-dimensional variational retrieval of aerosol extinction coefficient from synthetic LIDAR and radiometric measurements, *J. Geophys. Res.*, **112**(D14), D14303, doi:10.1029/2006/JD007625.

Illingworth, A.J., and Co-authors, 2007: CLOUDNET continuous evaluation of cloud profiles in seven operational model using ground-based observations, *Bull. Amer. Meteor. Soc.*, **88**, 883–898.

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

Kessler, E., 1969: On the distribution and continuity of water substance in atmospheric circulation, *Meteorol. Monog.*, **10**, American Meteor. Soc., Boston, Mass.

Klett, J. D., 1985: Lidar inversion with variable backscatter/extinction ratios, *Applied Optics*, **24**(11), 1638–1643.

Lipton, A. E. and G. D. Modica, 1999: Assimilation of visible-band satellite data for mesoscale forecasting in cloudy conditions, *Mon. Wea. Rev.*, **127**, 265–278.

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

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

Macpherson, B., B. J. Wright, W. H. Hand, and A. J. Maycock, 1996: The impact of MOPS moisture data in the U.K. Meteorological Office mesoscale data assimilation scheme, *Mon. Wea. Rev.*, **124**, 1746–1766.

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

Martin, G. M., W. Johnson, and A. Spice, 1994: The measurement and parameterization of of effective radius of droplets in warm stratocumulus, *J. Atmos. Sci.*, pp. 1823–1842.

Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, J. W. Kaiser, M. Razinger, *et al.*, 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, *J. Geophys. Res.-Atmospheres*, **114**(D6), D06206, doi: 10.1029 /2008JD011235.

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

Ou, S.-C. and K.-N. Liou, 1995: Ice microphysics and climate temperature feedback, *Atmos. Res.*, **35**, 127–138.

Rabier, F., A. 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, *Quart. J. Roy. Meteor. Soc.*, **124**, 1809–1829.

Simpson, J. and V. Wiggert, 1969: Models of precipitating cumulus towers, *Mon. Wea. Rev.*, **97**, 471–489.

Slingo, A., 1989: A GCM parameterization for the shortwave radiative properties of water clouds, *J. Atmos. Sci.*, **46**, 1419–1427.

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

Teixeira, J., 2001: Cloud fraction and relative humidity in a prognostic cloud fraction scheme, *Mon. Wea. Rev.*, **126**, 1750–1753.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Wea. Rev.*, **117**, 1779–1800.

Tiedtke, M., 1993: Representation of clouds in large-scale models, *Mon. Wea. Rev.*, **121**, 3040–3061.

Tompkins, A. M., P. Bechtold, A. C. M. Beljaars, A. Benedetti, S. Cheinet, M. Janisková, M. Köhler, P. Lopez, and J.-J. Morcrette, 2004: Moist physical processes in the IFS: Progress and plans, ECMWF Technical Memorandum 452, 91 pp.

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