TECHNICAL MEMORANDUM

487

Implementation of 1D+4D-Var Assimilation of Precipitation Affected Microwave Radiances at ECMWF, Part I: 1D-Var

Peter Bauer[1], Philippe Lopez[1], Angela Benedetti[1], Deborah Salmond[1] and Emmanuel Moreau [2]

Research Department

[1]European Centre for Medium-Range Weather Forecasts, UK
[2]NOVIMET, France

Accepted for publication in Quart. J. Roy. Meteor. Soc.

February 2006

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

Series: ECMWF Technical Memoranda

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

Abstract

This paper presents the operational implementation of a 1D+4D-Var assimilation system of rain affected satellite observations at ECMWF. The first part describes the methodology and performance analysis of the 1D-Var retrieval scheme in clouds and precipitation that uses SSM/I microwave radiance observations for the estimation of total column water vapor. The second part shows the global and long-term impact of these observations on both model 4D-Var analyses and medium-range forecasts.

The 1D-Var scheme employs a complex observation operator that consists of linearized moist physics parameterization schemes and a multiple scattering radiative transfer model. The observation operator shows a rather linear behavior in most situations except in the presence of very intense precipitation suggesting a possible use even for a direct assimilation of radiances in 4D-Var. A bias correction and observation error estimation method were implemented and indicate stable error behavior. The 1D-Var algorithm quality control shows the largest failure number in areas with mostly frozen precipitation where the SSM/I channels have little sensitivity to changes in hydrometeor contents. From test analyses on a global scale, a small moisture increase was computed that was greatest in dry subtropical areas. Large-scale and convective precipitation were increased similarly but showed a significantly different geographical distribution. The large-scale precipitation scheme has a stronger sensitivity to moisture changes and therefore moisture increments mainly affect stratiform precipitation distributions. While the global mean moisture fields are only weakly affected by the assimilation of rain affected observations, the impact on local systems may be quite large. The forecast of synoptic system development through the 4D-Var analysis can be significant.

1 Introduction

The focus of this two-part paper is the implementation of a rain assimilation framework in the global modeling system at ECMWF. This purpose prescribes already the choice of methodologies to be employed and the definition of some constraints given by the existing model configuration. Firstly, ECMWF operates an incremental formulation of a four-dimensional variational (4D-Var) data assimilation system with four analyses per day, namely two 12-hour assimilation window and two 6-hour assimilation window ones. The former produce the background fields for the latter while the latter produce two 10-day forecasts initialized at 00 and 12 UTC, respectively. A vast set of observations are assimilated of which about 98% originate from satellite data. These are almost exclusively screened for cloud and rain contamination except for those sounder channels that are not sensitive to the mid to lower troposphere. This produces an analysis bias in terms of data coverage towards cloud-free areas that are therefore much more strongly constrained by observations than cloudy and rain affect regions.

The assimilation of rainfall observations in numerical weather prediction (NWP) models has been covered in various research studies over the last 20 years. The most apparent issues when dealing with this new type of observations are (1) the choice of rain-related observation treatment and the estimation of its error characteristics (e.g. Errico et al. 2000), (2) the interaction with the moist physics parameterization schemes in the model, in particular their possibly non-linear response to constraints introduced by the observations (e.g. Fillion and Errico 1997), and (3) the response of the model dynamics that is the trade-off between the desired improvement of moisture distributions through modification of divergence fields and the unwanted effect of, for example, gravity wave excitation by unbalancing local dynamics (e.g. Fillion 2002). Details of these issues that are related to the variational assimilation of rain observations in the present model configuration with respect to other studies will be presented in part II of this paper (Bauer et al. 2006a).

At ECMWF, new satellite observations have been assimilated through one-dimensional variational (1D-Var) retrievals of intermediate parameters before a direct assimilation of electromagnetic radiances was attempted (e.g. Eyre et al. 1993, Phalippou 1996). The advantage of this choice is the better control of a non-linear response of the observation operator to changes in the atmospheric state as well as the additional level of

quality control before data is passed on to the 4D-Var system. For cloud observations, 1D-Var retrieval studies have been carried out (Chevallier et al. 2002, Deblonde and English 2003) that provide a lot of insight into observation operator and minimization performance from flexible sensitivity testing that is based on global model statistics. For rain observations Marécal and Mahfouf (2000, 2002) developed a similar approach using Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) based surface rain rate estimates. Their work was crucial for all further studies at ECMWF because most of the issues listed above were discussed.

Chevallier and Bauer (2003) showed that microwave radiances that are simulated using global model fields are quite realistic. Later, the direct use of microwave radiances instead of rain rate observations was established in the variational retrieval method (Moreau et al. 2002). A direct comparison of radiance with rain rate observations in the 1D-Var (Moreau et al. 2003) reveals that the choice of radiances mainly serves to avoid the dependence on sensor-specific retrieval algorithms and to simplify the rather uncertain error estimation (e.g. Bauer et al. 2002). The developments that lead to the implementation which became operational at ECMWF on June 28, 2005, were (1) the implementation of the 1D-Var radiance assimilation in model grid-point space that is activated at each time step (currently 15 minutes) in the assimilation window, (2) the improvement of the observation operator including its tangent-linear and adjoint versions, (3) extensive testing of its performance with global and long-term data, (4) the implementation of quality control, bias correction and observation error formulation, (5) extensive testing of the impact of the assimilation on global model analyses and forecasts.

This paper consists of two parts, the first being presented here and the second showing the impact of total column water vapour (TCWV) pseudo-observations that were derived from the 1D-Var retrieval on the 4D-Var analyses and forecasts. Section 2 of this paper introduces the 1D-Var set-up and presents the combined moist physics-radiative transfer observation operator. The linearity of the operator is investigated using realistic perturbations that may lead to conclusions regarding a potential direct assimilation of rain affected passive microwave observations in the 4D-Var analysis. Section 3 describes the data processing including the bias correction scheme and evaluates the performance of the 1D-Var algorithm inside the ECMWF analysis system. Problematic cases are identified in which convergence problems occur or unrealistic retrieval results are produced and that have to been screened out. The paper is concluded by a discussion of the developed methodology based on the shown results.

2 Algorithm

2.1 Optimum estimation

The assimilation of radiance data in precipitation represents an inversion problem that is not fully constrained. If, for example, Special Sensor Microwave / Imager (SSM/I) or TMI-type microwave channels are available as observations, they represent only 2-3 statistically independent measurements (e.g. Bauer 2001). Therefore, large weight is put on the 'a priori' constraints that are the background profiles of temperature and moisture, their error structure, and the observation operator that may comprise moist physics parameterizations and a radiative transfer model. However, the feasibility of a variational framework using this set-up has been demonstrated and tested with various radiometer channel combinations (Moreau et al. 2002, 2003).

In the following, the observations always refer to radiances measured by the SSM/I that are expressed as blackbody equivalent brightness temperatures (TB's) in units of degrees Kelvin (K). The SSM/I has seven channels with dual polarization measurements at 19.35, 37.0 and 85.5 GHz, respectively, and vertical polarization measurements at 22.235 GHz. Hereafter, the channels will be referred to as 19v, 19h, 22v, 37v, 37h, 85v, and 85h to identify measurement frequency and polarization in a simple way. The logical flow of the 1D -Var retrieval is shown in Figure 1.

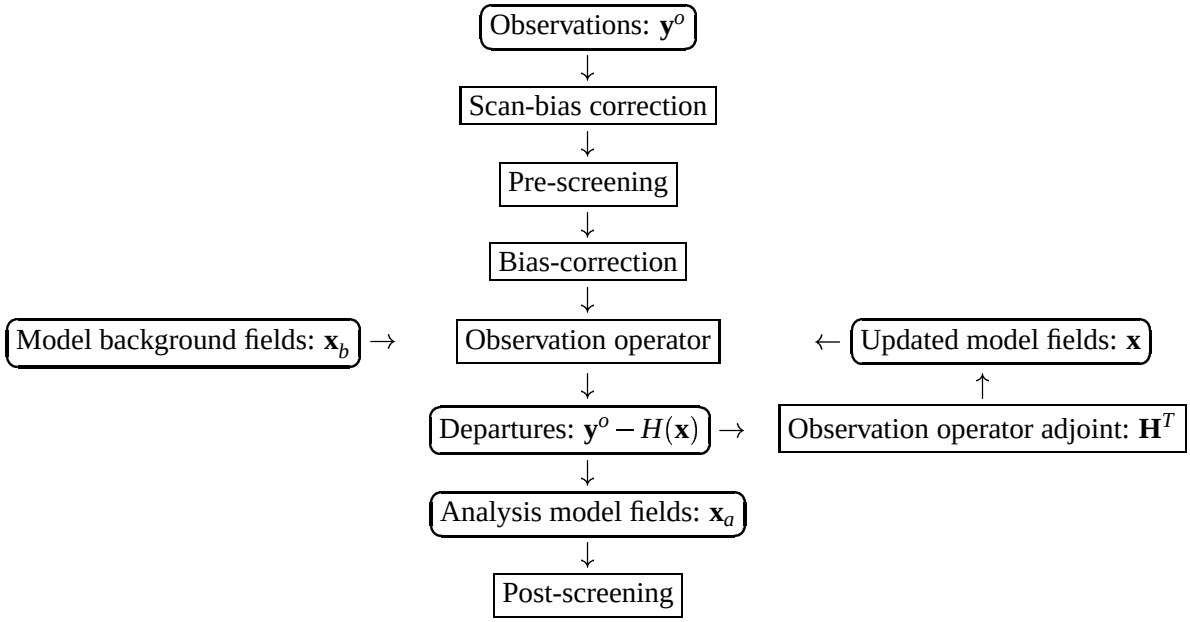

Figure 1: Logical flow of SSM/I radiance processing and 1D-Var retrieval.

In the variational retrieval (e.g. Rodgers 2000), the optimum estimate of a state vector, $\mathbf{x}$, is obtained using an observation vector (TB's), $\mathbf{y}^o$, plus additional a priori information. Due to the uncertainties associated with the background state, observations and the observation operator, H (that maps geophysical space onto observation space), the relation between state and observation space is usually described by probability density functions (pdf's). Applying Bayes' theorem and assuming that the errors of background state, $\mathbf{x}_b$, and observations are uncorrelated and have a Gaussian shape, the inversion problem can be formalized by the minimization of the well-known cost function, J:

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

with background error covariance matrix $\mathbf{B}$ and observation error covariance matrix $\mathbf{R}$. The latter includes the modeling error of operator H that is likely to be larger than the instrument noise of the SSM/I that is typically between 0.5-0.8 K. The issue of the specification of errors and their Gaussian distribution shape will be discussed in Section 3b.

The non-linearity of H is not necessarily an issue in this implementation because the variational retrieval may employ a non-linear minimization procedure. This represents an advantage for large differences (first-guess departures) between $\mathbf{y}^o$ and $H(\mathbf{x}_b)$. The linearity of the entire observation operator is investigated in Section 2b and the departure statistics are analyzed in Section 3a, respectively.

In our application, the control vector $\mathbf{x}$ contains vertical profiles of temperature and specific humidity on 60 model levels, hence its dimension of 120 in the current ECMWF model version. The minimization of Equation (1) requires the gradient of $J(\mathbf{x})$:

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \mathbf{H}^T\mathbf{R}^{-1}[H(\mathbf{x}) - \mathbf{y}^o] \tag{2}$$

where $\mathbf{H}^T$ is the adjoint of the observation operator. The quasi-Newtonian (M1QN3) software developed by (Gilbert and Lemaréchal 1989) is used to perform the minimization.

The minimization is pre-conditioned for convergence improvement by decomposing $\mathbf{B} = \mathbf{E}\Lambda\mathbf{E}^T$ where $\mathbf{E}$ is the matrix of eigenvectors of $\mathbf{B}$ and Λ is a matrix whose diagonal contains the eigenvalues of $\mathbf{B}$. The state vector $\mathbf{x}$

is transformed to κ with forward and backward transformations:

$$\begin{aligned} \kappa &= \Lambda^{-1/2}\mathbf{E}^T(\mathbf{x}-\mathbf{x}_b) \\ \mathbf{x} &= \mathbf{x}_b + \mathbf{E}\Lambda^{1/2}\kappa \end{aligned} \tag{3}$$

so that the cost function gradient becomes:

$$\nabla J(\kappa) = \kappa + \Lambda^{1/2}\mathbf{E}^T\mathbf{H}^T\mathbf{R}^{-1}[H(\mathbf{x}) - \mathbf{y}^o] \tag{4}$$

The above transformation is applied before and after the minimization, respectively, and is justified by the rather different magnitudes of departures and background errors of these parameters.

2.2 Observation operator

The observation operator contains moist physics parameterization schemes and a multiple-scattering radiative transfer model. The moist physics parameterizations consist of a large-scale condensation scheme (Tompkins and Janisková 2004) and a convection scheme (Lopez and Moreau 2005). Both represent models that aim at a similar performance of the forward models as the non-linear moist physics parameterizations employed at ECMWF (Tiedtke 1989, 1993). However, they represent model version whose sensitivity to perturbations of the input parameters is more linear than that of the Tiedtke-parameterizations. This produces a more controlled behavior in the minimization and avoids excessive increments that may cause convergence problems. The multiple-scattering radiative transfer model is part of the RTTOV software package (e.g. Saunders et al. 2005) that has been extensively tested for data assimilation purposes by Bauer et al. (2006b).

The convection scheme represents subgrid-scale processes and treats several convection types, namely shallow, mid-level and deep convection, in a unified way. In contrast to previous models employed at ECMWF, the tangent-linear and adjoint models account for perturbations of all convective quantities such as convective mass flux, updraught characteristics and precipitation flux. The large-scale condensation scheme uses the convective detrainment prescribed by the convection model with a similar precipitation generation formulation. The cloud scheme applies a statistical method for the description of subgrid-scale cloud fluctuations affecting cloud cover and cloud water. Subgrid-scale variability of humidity is used to provide an improved modeling of precipitation evaporation. The radiative transfer model applies the Delta-Eddington approximation to radiative transfer that is widely considered sufficiently accurate at microwave frequencies (Smith et al. 2002). The output from the moist physics parameterizations, i.e., cloud cover and precipitation fluxes are used to compute cloud layer optical properties based on pre-calculated look-up tables.

The observation operator requires more input variables than those contained in the control vector $\mathbf{x}$. The full set of parameters is referred to as the state vector (see Table 1). It contains the profiles of specific humidity and temperature (which represent the control vector) as well as their tendencies plus surface fields. These are the latent and sensible heat flux, the zonal and meridional components of wind stress and 10-meter windspeed, skin temperature and pressure, 2-meter temperature, 2-meter specific humidity and 2-meter dew-point temperature. The moist physics parameterizations produce profiles of fractional cloud cover, cloud liquid and ice water as well as liquid and frozen precipitation. The hydrometeor profiles represent the input to the radiative transfer model together with the 10-meter windspeed, the skin temperature and both temperature and moisture profiles. Note that only the control vector is contained in the cost-function in Equation (1) while state vector components may be affected by changes in the control vector during minimization but do not affect the cost-function directly.

The accuracy of the observation operator can not be directly determined because the required in situ observations of all parameters are not available. The relative accuracy of the moist physics parameterizations is

Table 1: Input (I) and output (O) parameters of observation operator components.

	Convection	Large-scale condensation	Radiative Transfer
Profiles:			
Temperature	I	I	I
Specific humidity	I	I	I
Temperature tendency	I	I	
Specific humidity tendency	I	I	
Detrained convective cloud water	O	I	
Cloud water content	O	O	I
Cloud ice content	O	O	I
Cloud cover fraction		O	I
Liquid precipitation flux (rain)	O	O	I
Frozen precipitation flux (snow)	O	O	I
Surface fields:			
Latent heat flux	I		
Sensible heat flux	I		
Wind stress (zonal, meridional)	I		
10-meter wind speed (zonal, meridional)			I
2-meter temperature			I
2-meter dew-point temperature			I
2-meter specific humidity			I
Pressure	I	I	I
Skin temperature			I

evaluated from comparison to the non-linear ECMWF cloud and convection schemes. Those have been evaluated in the framework of global model intercomparison studies under the auspices of, for example, the WMO Working Group on Numerical Experimentation (WGNE) or the Global Energy and Water Cyle Experiment (GEWEX) Cloud System Study (GCSS). The radiative transfer model's main sources of error are the approximation of radiation propagation through broken clouds and the assumptions made for calculating particle single scattering properties (e.g. for particle size distributions; Bauer et al. 2006b). As a generally applicable error estimate for data assimilation, an indirect method was applied (see Section 3b).

The assimilation of rain-affected microwave radiances is new and therefore little work has been published on the performance of the combined moist physics - radiative transfer observation operator. Apart from computational efficiency, the most fundamental question is how linearly the observation operator behaves. This is particularly important if the operator is applied within incremental variational assimilation systems. The test of the adjoint code is straight forward because it only relies on a numerical check using both tangent-linear and adjoint code versions. The involved adjoint tests have been carried out successfully and will not be reproduced here.

The linearity tests are based on comparing the output of the tangent-linear model with those from finite-difference calculations using the forward model. For example the ratio:

$$F = \frac{H(\mathbf{x} + \lambda\,\delta\mathbf{x}) - H(\mathbf{x})}{\lambda \mathbf{H}(\delta\mathbf{x})} \tag{5}$$

may be used as a linearity measure in which $\delta\mathbf{x}$ represents the initial perturbation of the control vector and λ is a scaling factor. In the linear case, scaling of the output of the tangent-linear model should produce the same result as the scaling of the input to the forward model.

The scaling factor, λ, usually spans several orders of magnitude to sample all possible perturbation sizes. The

$\delta\mathbf{x}$ should be realistic because, theoretically, $\delta\mathbf{x}$ could be chosen too small so that even rather non-linear models show a nearly linear performance. Errico and Raeder (1999) investigated the limitations of the tangent-linear approximation to moist physics parameterizations and pointed out the importance of adequate testing with realistic perturbations vs. infinitesimal ones. In variational analyses, the perturbations are of the order of the model uncertainties which are similar to the first-guess departures.

In our case, the analysis minus first-guess departures from the 1D-Var analysis serve this purpose because they define the improvement of the first-guess state that was achieved by constraining the retrieval with observations given the defined model and observation errors, respectively. Therefore:

$$\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b \tag{6}$$

which is scaled with λ ranging from 10^{-10} to 1 with increments of 10^2. The linearity of the moist physics parameterizations component and the total observation operator can be tested separately. Since $\mathbf{x}$ is the control vector it contains profiles of temperature and specific humidity. If only the moist physics parameterizations are tested the output of $H(\mathbf{x})$ consists of profiles of hydrometeors and cloud cover. F can then be displayed as a function of the respective output, i.e., hydrometeor contents or SSM/I channel radiances.

Figures 2-4 show histograms of $log_{10}|1-F|$ from 8,290 profiles distributed over global oceans from two model analyses at 00 and 12 UTC on December 31, 2004. Ideally, F should approach zero and the histograms should show the majority of the data at very low values. If $log_{10}|1-F| < 0$, a decently linear regime is achieved, and for example, at $log_{10}|1-F| = -1$, F would lie between 0.9 and 1.1 that is a 10% disagreement between the finite-difference and the tangent-linear calculations.

For rain (Figure 2), the distribution does not show much sensitivity to altitude which is because precipitation is computed with a diagnostic scheme that immediately rains out the hydrometeors once they have formed. For very small λ, the results deteriorate due to numerical limitations. For large λ, the results deteriorate as well because the larger perturbations produce stronger non-linearities. An optimum state can be identified for $\lambda \approx 10^{-6}$ that produces the highest concentration of cases near -7. There are clearly two regimes visible that mark nearly linear and rather non-linear situations which correspond to deeper layers and heavier precipitation (rain flux greater than 10^{-3}kg m^{-2}s^{-1}). This regime is almost independent of λ because for intense systems, large sensitivities to small perturbations may exist. For frozen precipitation (Figure 3) there is a strong altitude dependence. For higher altitudes the linearity improves due to lower snow contents. Above 550 hPa (model level 40) snow mainly occurs in the more intense systems and shows a lower degree of linearity.

Cloud liquid water (not shown here) produces a distribution that is very similar to that of liquid precipitation. Cloud cover, however, (Figure 4) represents a convolution of the rain, snow and cloud water distributions with a good linearity where snow exists at high altitudes and for the weaker systems in the presence of rain. The secondary maximum for intense systems is evident but $log_{10}|1-F|$ mainly remains below 10^{-6}.

If the linearity test is applied to the entire observation operator that is moist physics parameterizations and multiple scattering radiative transfer model, the sensitivity of radiance observables to realistic perturbations in the control vector can be tested. This can help to select channels which are less affected by non-linearities and to develop screening procedures for the exclusion of situations in which the operator can be expected to behave non-linearly. Figure 5 shows the results based on the same dataset as used for Figure 2-4 for the seven SSM/I channels. For clarity, only values of $\lambda = 10^{-10}, 10^{-6}, 10^{-2}$ are shown. The results for $\lambda = -2$ and $\lambda = -10$ are quite similar for most channels. This indicates the limitation of calculating $log_{10}|1-F|$ by the numerical accuracy for small λ. In most cases, the bulk of the probability distribution of $log_{10}|1-F|$ remains below 0. Even for $\lambda = 1$ (not shown), the maximum of the distributions is near -2 that corresponds to a 1% difference between finite-difference and tangent-linear calculations. For $\lambda = -2$ and $\lambda = -6$ two different regimes can be identified that refer to the two regimes seen in Figures 2-4. Among the SSM/I channels, 37v, 37h, 85v,

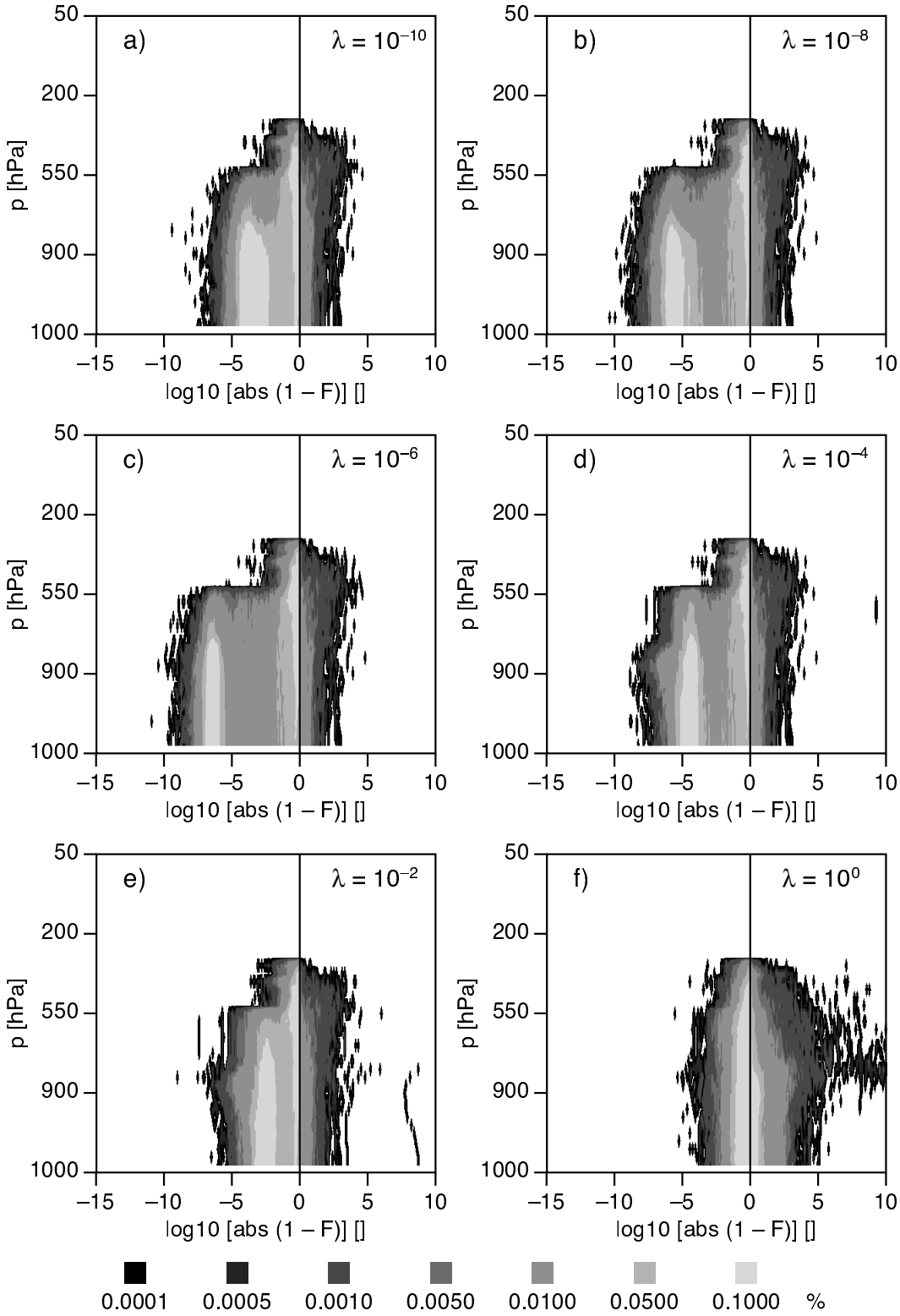

Figure 2: Frequency distribution of linearity test parameter in terms of rain flux for input perturbations of temperature and specific humidity scaled by a factor λ of 10^{-10} (a), 10^{-8} (b), 10^{-6} (c), 10^{-4} (d), 10^{-2} (e), 10^{0} (e), respectively. Ordinate shows model levels and the colour coding unit is %.

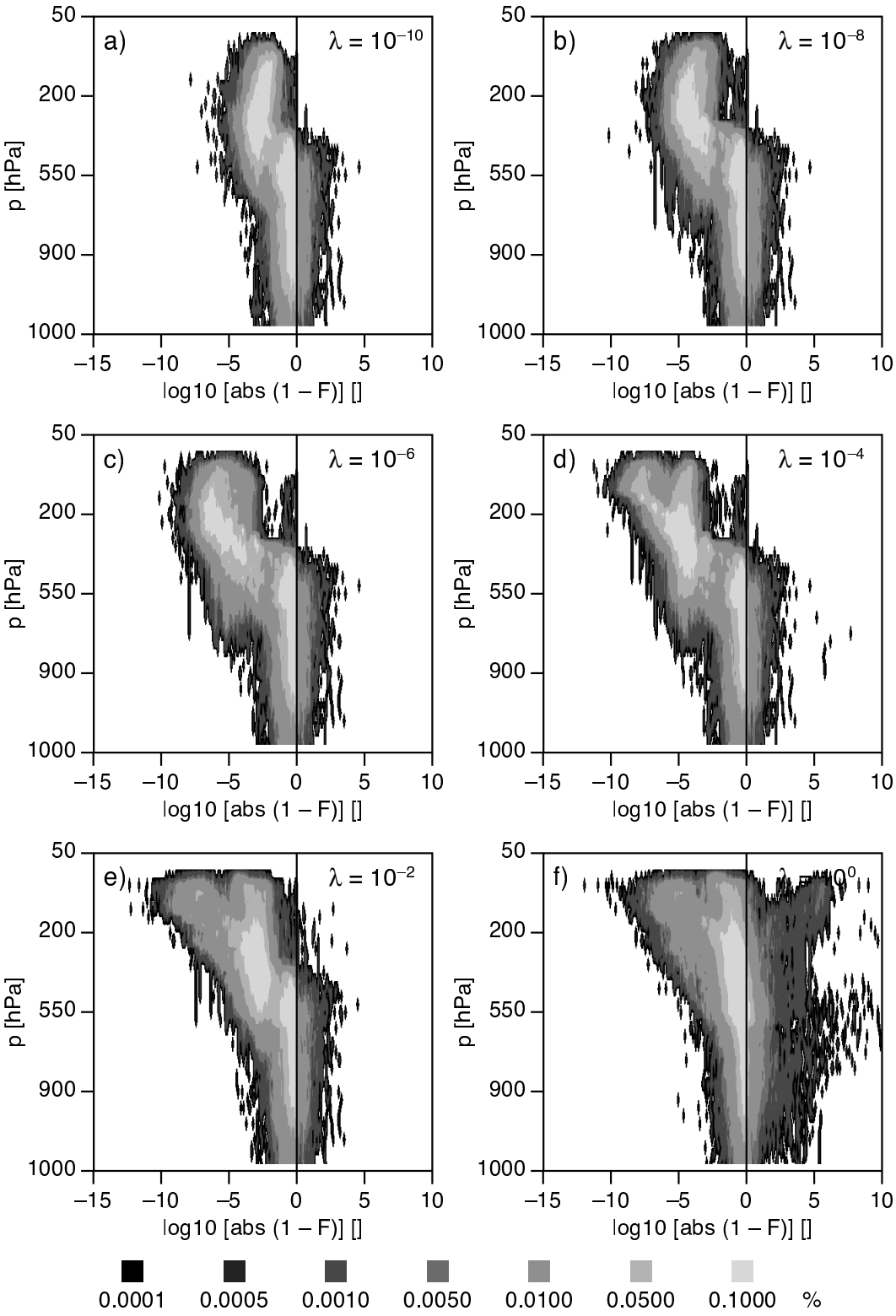

Figure 3: As Figure 2 for snow.

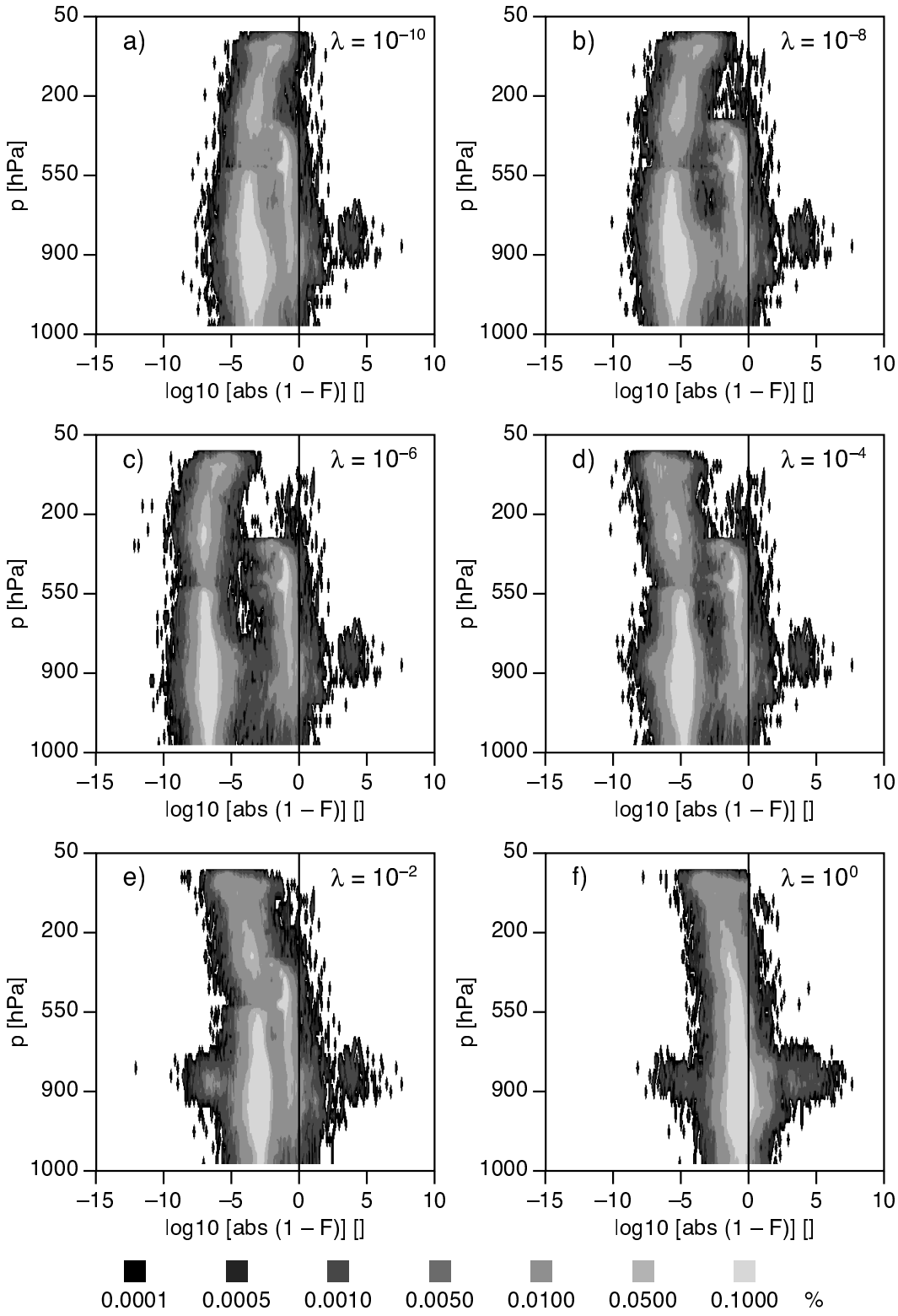

Figure 4: As Figure 2 for fractional cloud cover.

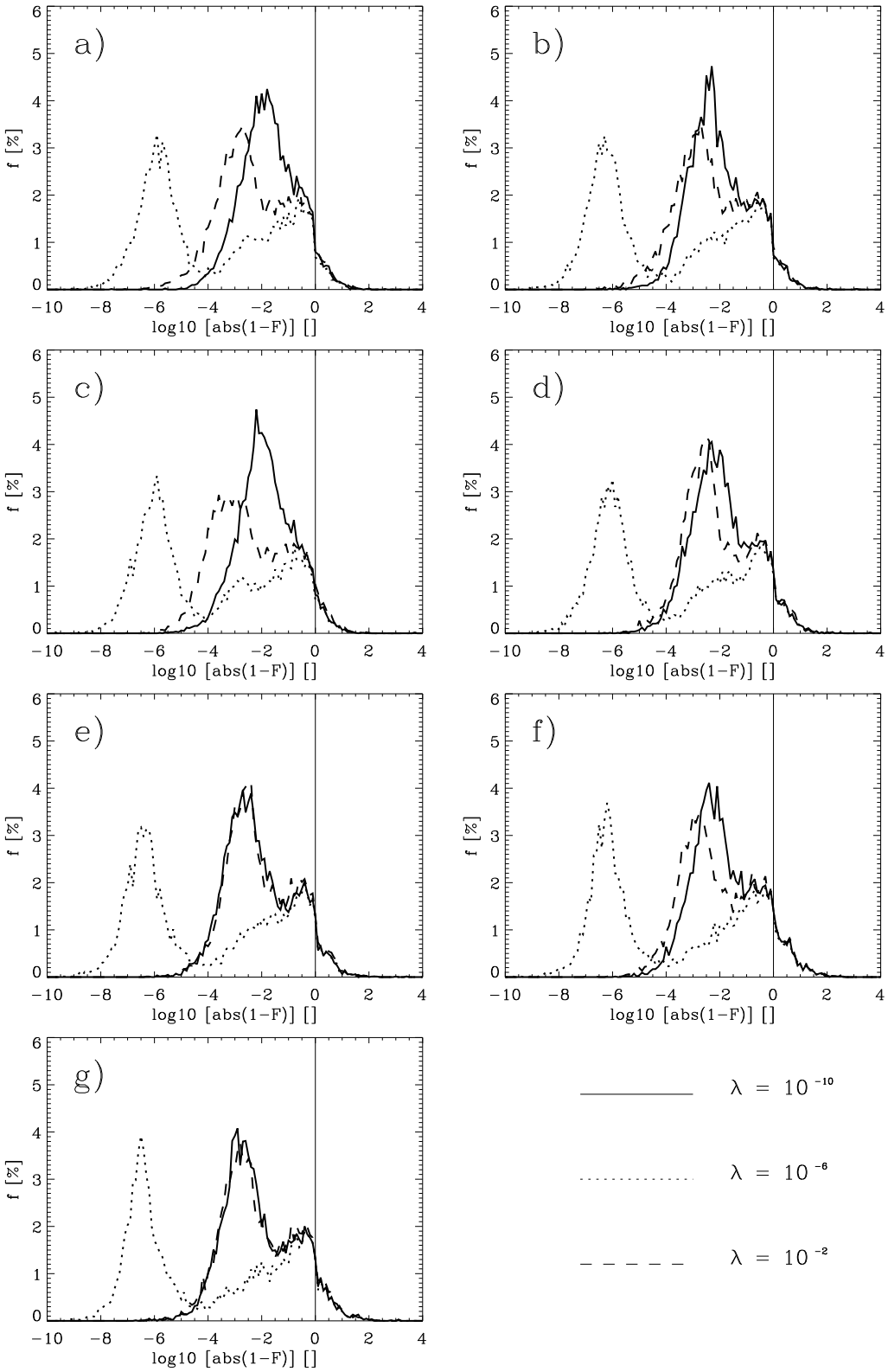

Figure 5: Frequency distribution of linearity test parameter for different scaling factors λ and channels 19v (a), 19h (b), 22v (c), 37v (d), 37h (e), 85v, (f), and 85h (g), respectively.

and 85h show slightly larger $log_{10}|1-F|$-values than the others. This was already shown in an independent investigations of microwave radiative transfer modeling linearity (Bauer et al. 2006b).

3 Results

3.1 Processing

The processing of rain affected SSM/I TB's involves several steps that are also illustrated in Figure 1. These include a data screening prior to the 1D-Var retrieval, a correction of systematic differences between observed and simulated TB's ('bias correction'), a post-retrieval data screening, and a data thinning. Bias correction is a sensitive issue because it represents an area of trade-off between true instrument biases and model biases that can be corrected in the analysis. However, the variational retrieval framework is defined for unbiased departures and, secondly, if systematic increments are introduced into the analysis unwanted effects such as excessive precipitation release (and subsequent spin-down) may occur.

The SSM/I observations suffer from a scan-position dependent bias that is caused by the partial intrusion by the satellite structure mainly near the end of the scan (Colton and Poe 1999, their Figure 3). These biases amount to several degrees K and they are static, which means that constant bias corrections per scan position suffice. These biases have been calculated from 3 months of SSM/I data (F-13, F-14, F-15) over oceans. A cloud check was applied and the differences of clear-sky TBs between scan-center and the other beam positions were averaged over the entire data set. These biases were derived from observations alone to avoid a contamination of the bias calculation by scan-angle dependent model biases produced by, for example, sea surface emissivity model biases as a function of relative azimuth between scan angle and wind direction.

The next processing stage is the data pre-screening that includes a land surface and sea-ice check, the exclusion of cases with high near-surface windspeeds (see Section 3d), a check for valid TB observations and the screening of clear-sky observations not to be treated in the retrieval. Ideally, the 1D-Var retrieval would be applied in all cases except where neither observations nor model first-guess fields show clouds or precipitation. However, the ECMWF data assimilation system employsa screening that is purely based on the observations so that cloud and precipitation screening for clear-sky observations does not rely on model fields. This may change in the future when cloud-affected observations will enter the analysis at least as passive variables inside the observation operator to avoid aliasing of the sensitivity to cloud variables into the temperature and moisture analysis. In our case, a check for cloud liquid water and precipitation presence is applied that is based on TB-thresholds.

Once clouds and/or rain have been identified, the observation operator is applied to the model first-guess profiles and SSM/I TB's, i.e. $H(\mathbf{x}_b)$, are calculated. Note that the entire screening and retrieval procedure is activated along the first model trajectory at full resolution. This is because the 1D-Var observation operator requires a large number of fields that are most optimally accessed immediately after the calculation of the model physics. This implementation requires that the observations are mapped to the model grid (40 km sampling) and to the actual time step (15 minute window). Despite the channel dependent spatial resolution of SSM/I footprints, no adjustment is made and model resolution and observation resolution are considered identical. The implications of this assumption are presented by Bauer et al. (2006b). After the calculation of the first-guess model equivalent TB's, a first-guess departure check is applied.

The bias correction uses a single predictor from the model first-guess fields in a linear regression, namely TCWV. This choice is reasonable because a parameter related to humidity is required that is available in both clear and cloudy situations and for which the TB's show sufficient sensitivity. Coefficients, (slope *a*, offset *b*), of the linear fit were calculated over increasing data volumes in September 2004 to ensure that the fit

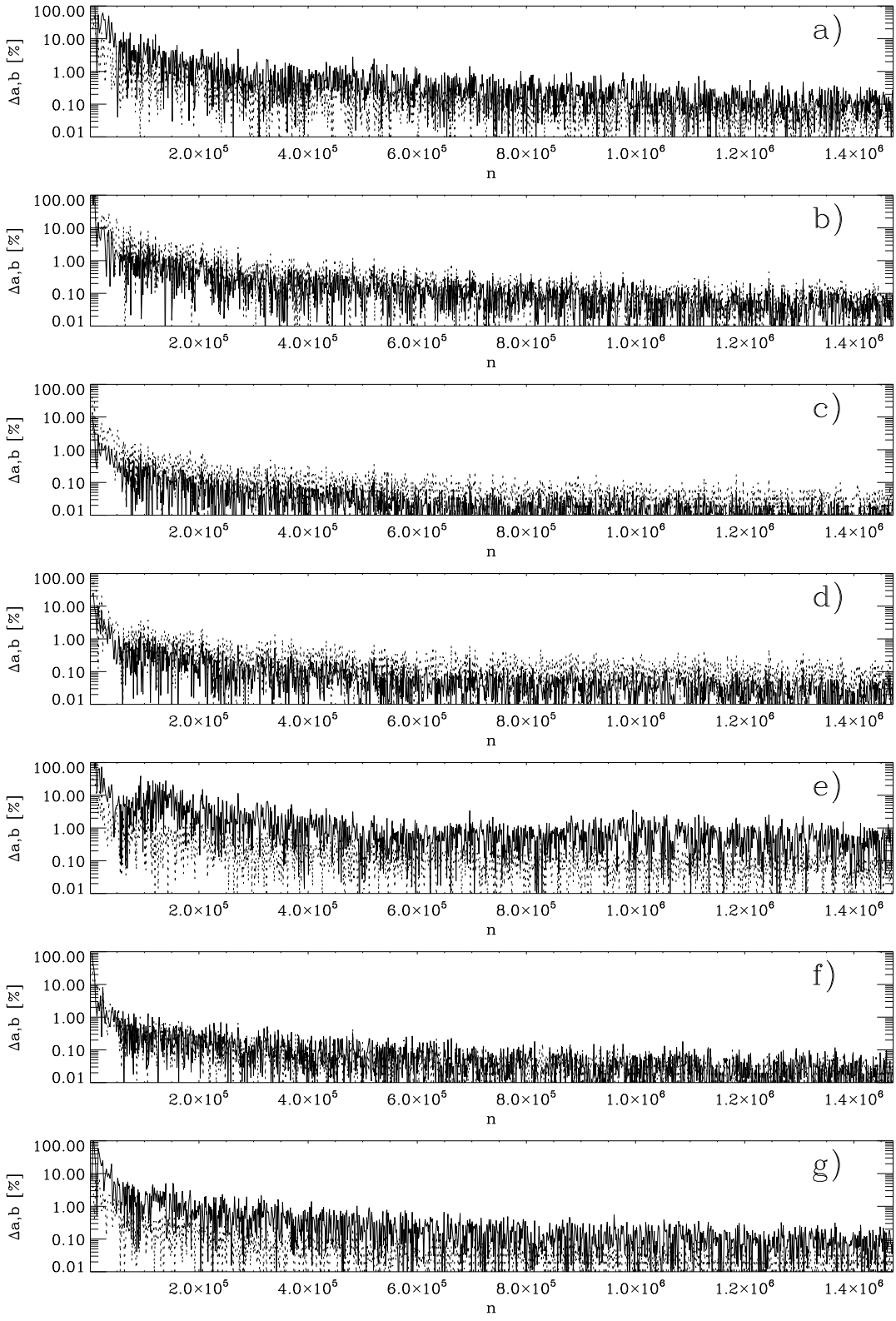

Figure 6: Time-series (expressed as increasing sample size 'n') of gradients for bias-correction coefficient (a, b; September 2004) adjustment in % for channels 19v (a), 19h (b), 22v (c), 37v (d), 37h (e), 85v, (f), and 85h (g), respectively.

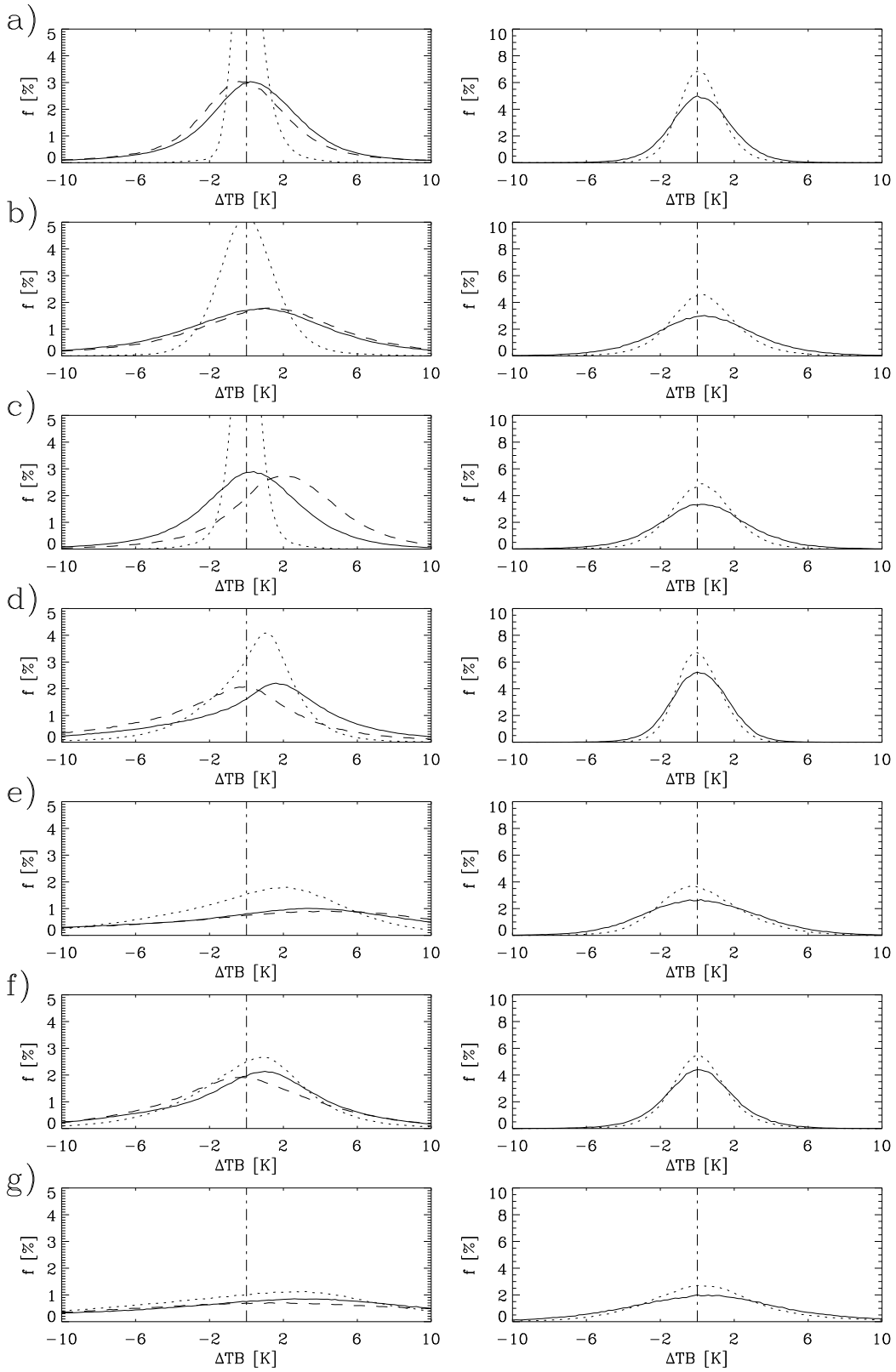

Figure 7: TB-departures between observations and bias-corrected first-guess model values (solid), analysis values (dotted) and uncorrected first-guess model values (dashed). Left panels show distributions for cloud and rain simulations, right panels show clear-sky distributions for channels 19v (a), 19h (b), 22v (c), 37v (d), 37h (e), 85v, (f), and 85h (g), respectively.

Table 2: First-guess and analysis departure statistics (in K) from clear-sky and rain affected SSM/I radiance assimilation (n: sample size, ΔFG: mean first-guess departure, σ_{FG}: first-guess departure standard deviation, ΔAN: mean analysis departure, σ_{AN}: analysis departure standard deviation).

	19v	19h	22v	37v	37h	85v	85h
Clear-sky:							
n	221,202	220,966	220,944	221,358	220,690	220,976	219,895
ΔFG	0.21	0.41	0.33	0.14	0.26	0.13	0.48
σ_{FG}	1.76	3.01	2.67	1.57	3.13	2.06	4.57
ΔAN	0.17	0.37	0.25	0.11	0.26	0.12	0.45
σ_{AN}	1.28	2.03	1.83	1.24	2.39	1.71	3.54
Clouds/rain:							
n	1,472,320	1,472,320	1,472,320	1,472,320	1,472,320	1,472,320	1,472,320
ΔFG	0.30	0.34	0.17	0.30	0.50	-0.04	0.23
σ_{FG}	4.05	7.48	3.46	6.22	13.02	5.78	12.67
ΔAN	0.18	0.10	-0.04	0.16	0.14	-0.04	-0.38
σ_{AN}	0.92	1.81	0.93	2.87	5.74	4.82	8.85

stabilizes. Figure 6 shows the time series of the gradients of these coefficients with increasing data volume for each channel, respectively. The x-axis denotes sample size instead of time and the axis range refers to the entire month. For most channels, the gradient reduces fairly quickly to less than 0.1-1% after 2-3 weeks. The implemented coefficients have been taken therefore from the final calculation at the end of the period.

Figure 7 and Table 2 show the resulting first-guess and analysis departure statistics. For comparison, similar statistics have been generated from the clear-sky SSM/I radiance assimilation in the ECMWF analysis. The mean first-guess departures show that the remaining radiance biases after applying the bias correction are very similar for both clear-sky and rain assimilation data. The first-guess departure standard deviations of rain affected radiances are between 1.3-2.8 times larger than those in clear skies. This result is quite remarkable keeping in mind that the first-guess fields inside precipitation are expected to be less accurate than outside and that the observation operator is more complex than the clear-sky radiative transfer model. The analysis departures are smaller by 50% than the first-guess departures in clear conditions, this reduction inside clouds and precipitation amounts to 30%. This shows the good convergence of the minimization in either case and the impact of the respective observation error definition (see next section).

Figure 7 shows, however, that the bias-correction works successfully only for the lower three SSM/I channels. Obviously, the correction has to be constrained further for the other channels. Adding another degree in the polynominal fit does not fully correct these biases and does not alleviate the non-Gaussian shape of the departure curves for the upper channels. Therefore, only the lower three channels were considered in the further analysis while the other four channels were maintained active in the 1D-Var retrieval for diagnostic purposes. Their observation errors were set to 999 K to eliminate their impact. As shown in Figure 1, the 1D-Var algorithm is then applied and employs both forward and adjoint of the observation operator in an iterative minimization of the cost-function in Equation (1). After minimization, the post-processing is carried out. Details of the 1D-Var performance and the post-processing are described in the next section.

3.2 Error definition

The minimization is constrained by both background and observation plus modeling error covariance matrices **B** and **R**. **B** is specified from the short-range forecast errors of temperature and specific humidity because these parameters also constitute the control vector in the 1D-Var retrieval. The forecast errors are calculated at a rather low spatial resolution that corresponds to a wavenumber truncation of 95, i.e. about 200 km (Rabier et al. 1998) and the vertical error covariance matrix is constant. The spatial variability of **B** is therefore only introduced by the error standard deviation and is supposed to be representative for both clear-sky and cloud scenes.

Part II of this paper (Bauer et al. 2006a) performs a more detailed study of background error covariance inside vs. outside precipitation. For this, the mean difference between 24-hour and 48-hour forecasts of temperature and specific humidity for the same target time were carried out assuming that short-range forecast errors can be represented by differences between forecasts over different periods. The results indicate that the currently available statistics do not allow a precipitation specific error formulation. Only dedicated validation programs will allow to produce error statistics from independent observations.

Hollingsworth and Lönnberg (1986) established a technique for the separation of background errors and observation errors using background departure statistics over areas with densely sampled ground-based observation networks. The method is based on the assumption that observation errors are spatially uncorrelated while background errors are spatially correlated. Therefore, plotting covariances of background departures from different points against their separation distance can be used to distinguish between non-zero covariances for separation distances $d > 0$ (between observations) solely due to the background term and those due to both background and observation terms at $d = 0$. Figure 8 shows the result for all seven SSM/I channels using 116,569 samples. Its is assumed that at $d = 0$ the variance of background departures $\sigma_{FG}^2 = \sigma_B^2 + \sigma_R^2$ with σ_B^2 being expressed as a variance in radiance units. The background term, σ_B^2, can then be obtained from extrapolating the histogram for $d > 0$ to $d = 0$. This seems justified because the shape of the histogram is rather smooth and flattens towards $d = 0$. The resulting error terms are summarized in Table 3.

The observation errors are slightly larger than the background errors and larger for horizontally polarized channels than for vertically polarized channels. The latter is explained by the larger dynamic range of the signal at horizontal polarization. This also means that the observation errors are mainly a result of modeling errors because the radiometric noise is similar for all channels (between 0.5 and 1 K) while the modeling error will depend on the magnitude of the simulated signal. This raises the question if the Hollingsworth and Lönnberg technique can be applied here because the observation errors must be spatially uncorrelated. This must also apply to the modeling errors because the background departures are obtained from $\mathbf{y}^o - H(\mathbf{x})$. While $\mathbf{y}^o$ is most likely uncorrelated, $H(\mathbf{x})$ may be not. This also holds for the inter-channel correlation that may be introduced by errors in H. However, Moreau et al. (2003) have shown that including non-zero off-diagonal elements in σ_R does not significantly affect the 1D-Var performance.

As an alternative solution, the background errors could be calculated from the diagonal terms of $\mathbf{HBH}^T$ that is the translation of the operational model background error covariance matrix into observation space and by applying $\sigma_R^2 = \sigma_{FG}^2 - \sigma_B^2$. However, this would also introduce modeling errors into $\mathbf{HBH}^T$ and neither would produce a clean separation of error contributions. Figure 9 shows the result from this calculation that is the square root of the diagonal terms of the $\mathbf{HBH}^T$ matrix for each SSM/I channel. The left panels denote situations where only the large-scale condensation scheme was active in the 1D-Var minimization while the both large-scale condensation and convection schemes were active in the data reproduced in the right panels. Both median (dotted line) and mode (dashed line) of the distributions are overplotted and their values are given in Table 3. Since the distributions are rather wide, the values were rounded off to integer numbers.

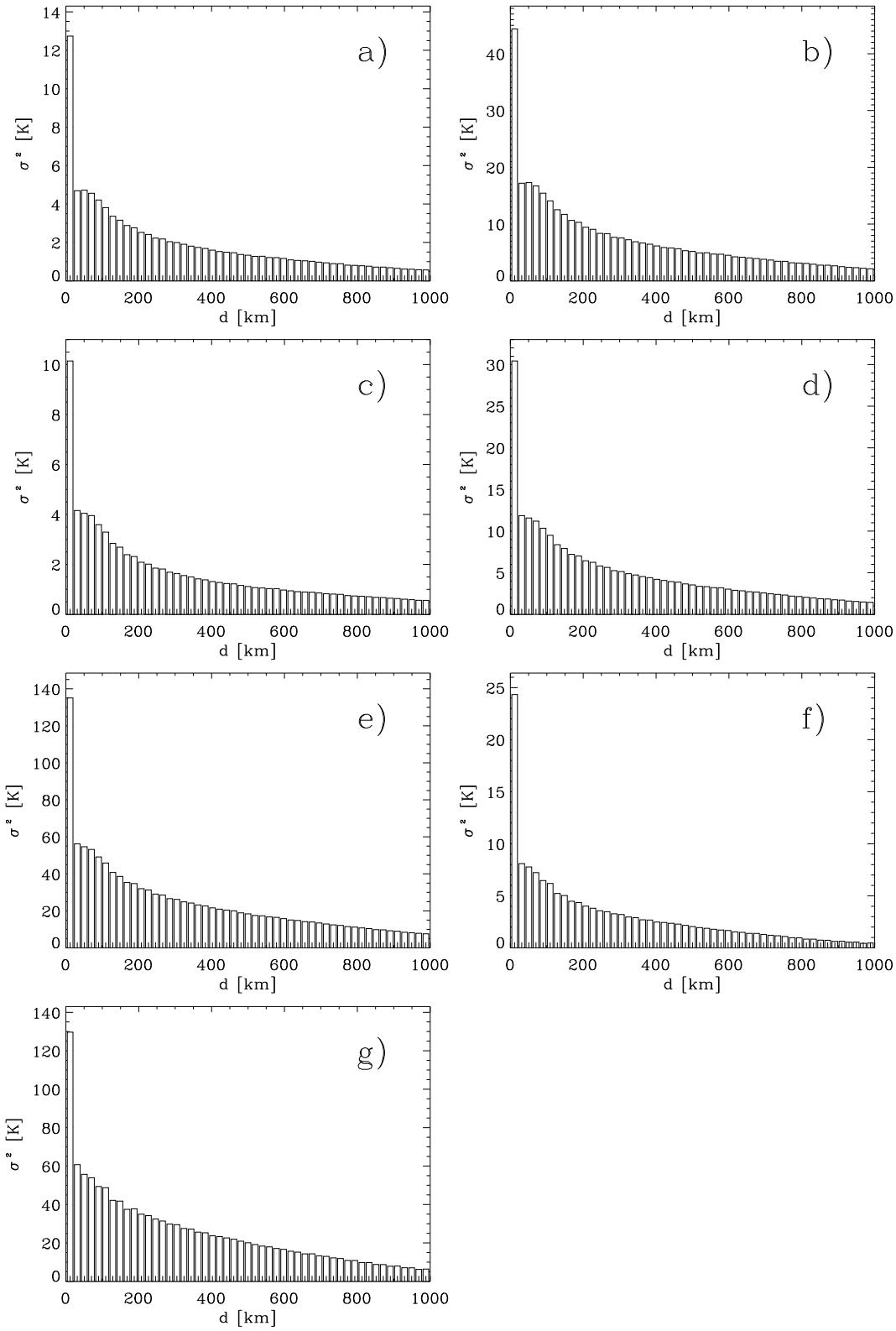

Figure 8: Spatial covariance of SSM/I TB's for channels 19v (a), 19h (b), 22v (c), 37v (d), 37h (e), 85v, (f), and 85h (g), respectively, as a function of separation distance d (in km) from 116,569 observations.

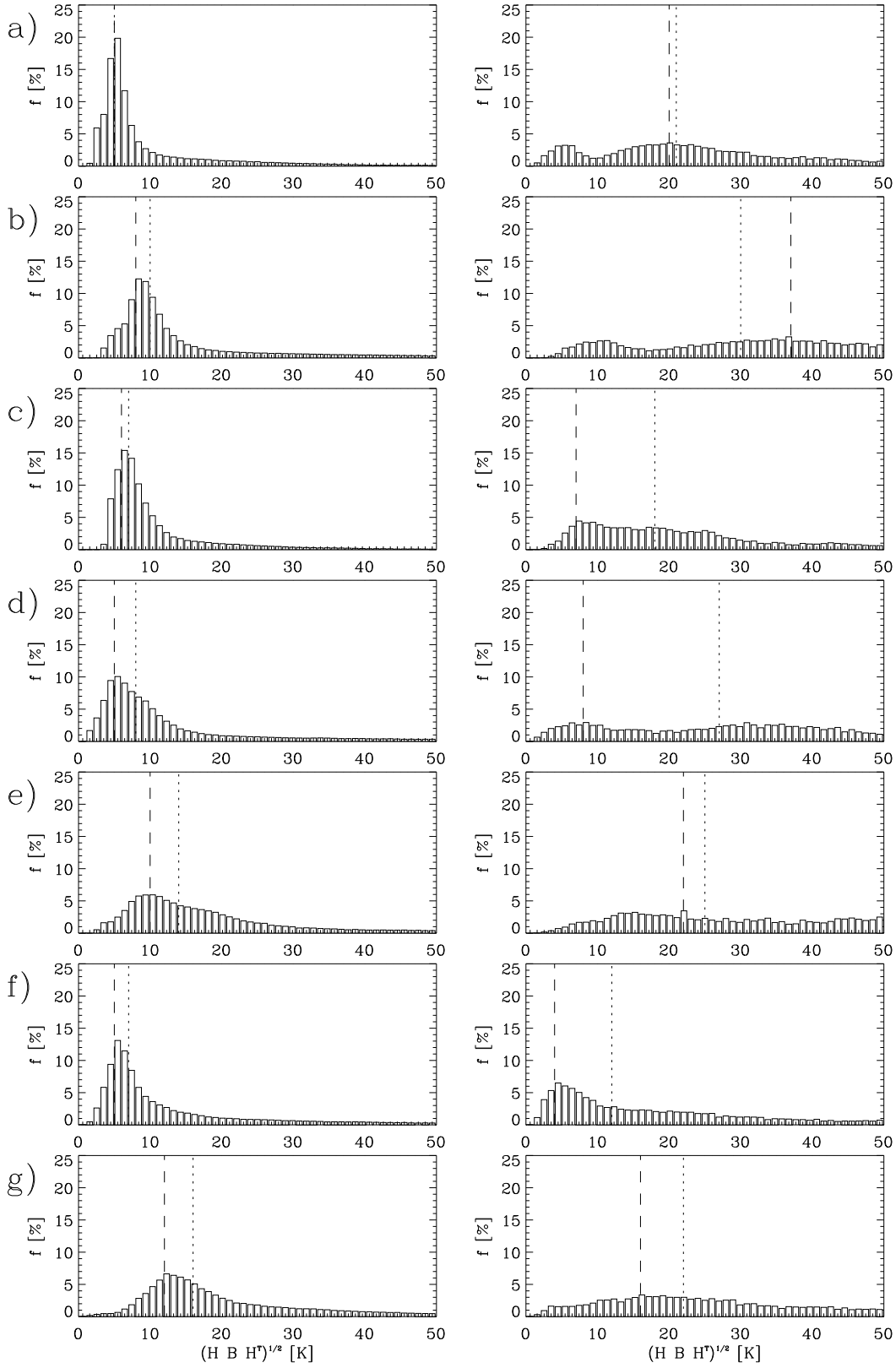

Figure 9: Background error standard deviation expressed in radiance terms for channels 19v (a), 19h (b), 22v (c), 37v (d), 37h (e), 85v, (f), and 85h (g), respectively, from 116,569 observations. Left (right) panels represent cases with large-scale condensation (+ convection) physics parameterization schemes activated.

Table 3: Estimation of background and observation error standard deviations (in K) from the spatial covariance of background departures (Hollingsworth and Lönnberg 1986) based on 116,569 observations. Median and mode of background error distributions in radiance space from $\mathbf{HBH}^T$ *calculations for cases with active large-scale condensation (LS) and convection (CV) schemes.*

	19v	19h	22v	37v	37h	85v	85h
σ_B	2.2	4.1	2.0	3.5	7.4	2.8	7.7
σ_R	2.8	5.2	2.5	4.4	9.0	4.0	8.4
Median: σ_{HBH^T}, LS	5	10	7	8	14	7	16
Median: $\sigma_{HBH^T}, LS+CV$	21	30	18	27	25	12	22
Mode: σ_{HBH^T}, LS	5	8	6	5	10	5	12
Mode: $\sigma_{HBH^T}, LS+CV$	20	37	7	8	22	4	16

The magnitude of background and observation errors obtained from the Hollingsworth-Lönnberg method were rather similar and by a factor of 2-3 smaller than those obtained from the $\mathbf{HBH}^T$-calculations. The main explanation for this is that the forward calculations using background profiles of temperature and moisture produce radiances that are close to the observations while the calculation of background variance properties in radiance space is not close to the observed minus modelled variance. This would suggest that the B-matrix in clouds and precipitation contains too large covariances. The cases where only the large-scale condensation scheme is active show much lower values for the median and mode of the distributions compared to those where also convection is found. When convection is active the atmosphere is likely to be more unstable and thus the combined sensitivity of large-scale condensation and convection scheme increases. The convection scheme produces detrained water that amplifies precipitation generation in the convection scheme. Secondly, convection is generally more active in the Tropics where also **B** for specific humidity is larger.

Both background and observation errors are difficult to specify in cloud and rain affected areas. At present, no operational weather forecasting system has an explicit formulation of model errors for these cases so that indirect methods such as the Hollingsworth and Lönnberg technique represent the only alternative. Moreover, the retrieved errors obtained from this method seem realistic and provide a balance between background and observation errors given the calculated background departures. The discrepancy between the results for the background errors in radiance terms obtained from the Hollingsworth-Lönnberg and the $\mathbf{HBH}^T$-method also suggest that there is a need for a better description of moisture background errors inside clouds and precipitation.

3.3 Convergence

Figures 10a and 10b show an example of the average reduction of both cost function and its gradient for a 12-hour analysis for 70,149 samples. In the current implementation, a maximum number of iterations, *k*, of 19 is allowed, otherwise, no convergence is assumed. Figure 10c shows that most retrievals require between 5 and 15 iterations. The cost-function gradients are reduced by 4 (2) orders of magnitude for the observation (background) components, respectively. This suggest a stable performance.

A zonal cross-sections of the proportion between successful 1D-Var retrievals and screened retrievals is shown in Figure 11a for the first 20 days in January 2005 with 5 degree resolution. The fraction of successful retrievals exhibits lower numbers in the Winter hemisphere. This reduction is mainly due to the increase of rejections due to excessive TCWV increments, i.e., the difference between TCWV retrievals and the model first-guess values. A threshold of 20% with respect to the first-guess TCWV was chosen to avoid excessive moisture changes with possibly negative impact due to the moist physical parameterization performance.

The areas with TCWV increment rejections are mainly located in the regions of frontal overpasses in the Northern oceans. In these regions, sea-surface temperatures are rather low and most of the precipitation is frozen. Generally, the three chosen SSM/I channels show little to no sensitivity to frozen precipitation so that the TCWV retrievals use little information from the observations.

The distribution of retrieval rejections due to non-convergence shows higher rejection rates in dry sub-tropical areas. This rejection is activated during the minimization if increments larger than 10 times the background error standard deviation appear or if negative moistures or unrealistic temperatures are produced. Again, this rejection flag prevents the moist physical parameterization schemes from failing. As expected, the average number of 1D-Var iterations (Figure 11b) follows the distribution of rejections due to convergence problems. The relative TCWV increments, $\Delta TCWV$, in Figure 11b show a zonally symmetric distribution due to the screening of excessive increments described earlier. However, the zonal structure reveals that the average increments are positive (about 1%) and that they are slightly larger in the Subtropics. This is in contrast to the result of Marécal et al. (2002) who identified a negative average TCWV increment from the assimilation of TMI rain rate estimates. Depending on the rainfall retrieval algorithm, Marécal et al. found an average increment of $-0.1 kg\,m^{-2}$. Our result of 1% translate to about $0.3 kg\,m^{-2}$. This mean departure is reduced in the 4D-Var analysis to about $0.04 kg\,m^{-2}$ (Bauer et al. 2006a). There are many potential explanations for this difference, namely the approach of assimilating rain rates vs. TB's and the different moist physical parameterizations schemes used in the 1D-Var algorithms as well the evolution of the non-linear model physical parameterizations. Another reason could be that Marécal et al. did not apply a bias-correction to the retrieved rain rates.

3.4 Impact

Figure 12 shows a mid-latitude front over the North Atlantic on December 31, 2004 from the analysis at 00 UTC. Figure 12a displays the difference between the 19v and 19h SSM/I channels that is indicative of increasing precipitation intensity with decreasing polarization difference. The polarization is produced by surface emission and has a maximum in clear skies. With increasing optical depth due to radiance absorption and multiple scattering from hydrometeors, the polarization signal is obscured. Since the polarization difference only decreases to about 30 K, the total condensed water amount is not very large.

TCWV increments from 1D-Var retrievals are overlayed and show large-scale (up to 30%) positive increments in the post-frontal cloud bands weaker negative increments ahead of the front while the cloud bands themselves are only locally modified. The mean profile increments suggest that positive TCWV increments are mainly produced by an increase of moisture below model levels 50 ($\approx$ 900 hPa) and a temperature decrease between model levels 43 and 58 ($\approx$650 and 1000 hPa). However, the profile increment standard deviations are large indicating a substantial spatial variability. Liquid precipitation only increases near the surface with a strong impact on frozen precipitation since the largest increments are located in the cold sector of the disturbance. Cloud liquid water and cloud cover are nearly unchanged. As previously mentioned, the three active SSM/I channels mainly respond to changes in liquid water and rain but the strongest impact is observed through the moist physical parameterizations on frozen precipitation.

The second example (Figure 13) originates from the same analysis and shows a tropical disturbance ('Chambo') in the Southern Indian Ocean passing by the islands of La Réunion and Mauritius. The polarization differences are still larger than 20 K suggesting that the system is not fully developed. The relative 1D-Var TCWV increments are much smaller than in the mid-latitude cases that is less than 15%. Keeping in mind tropical TCWV amounts, 10% in relative increments represent quite substantial absolute increments. The increments are mainly positive in the system's rain bands and negative in the less intense areas. This indicates an overestimation of rain coverage in the first guess. The profiles of increments show warming and drying (Figure 13b, c) over a deep layer between surface and 750 hPa. Above the freezing level, both liquid and frozen precipitation are

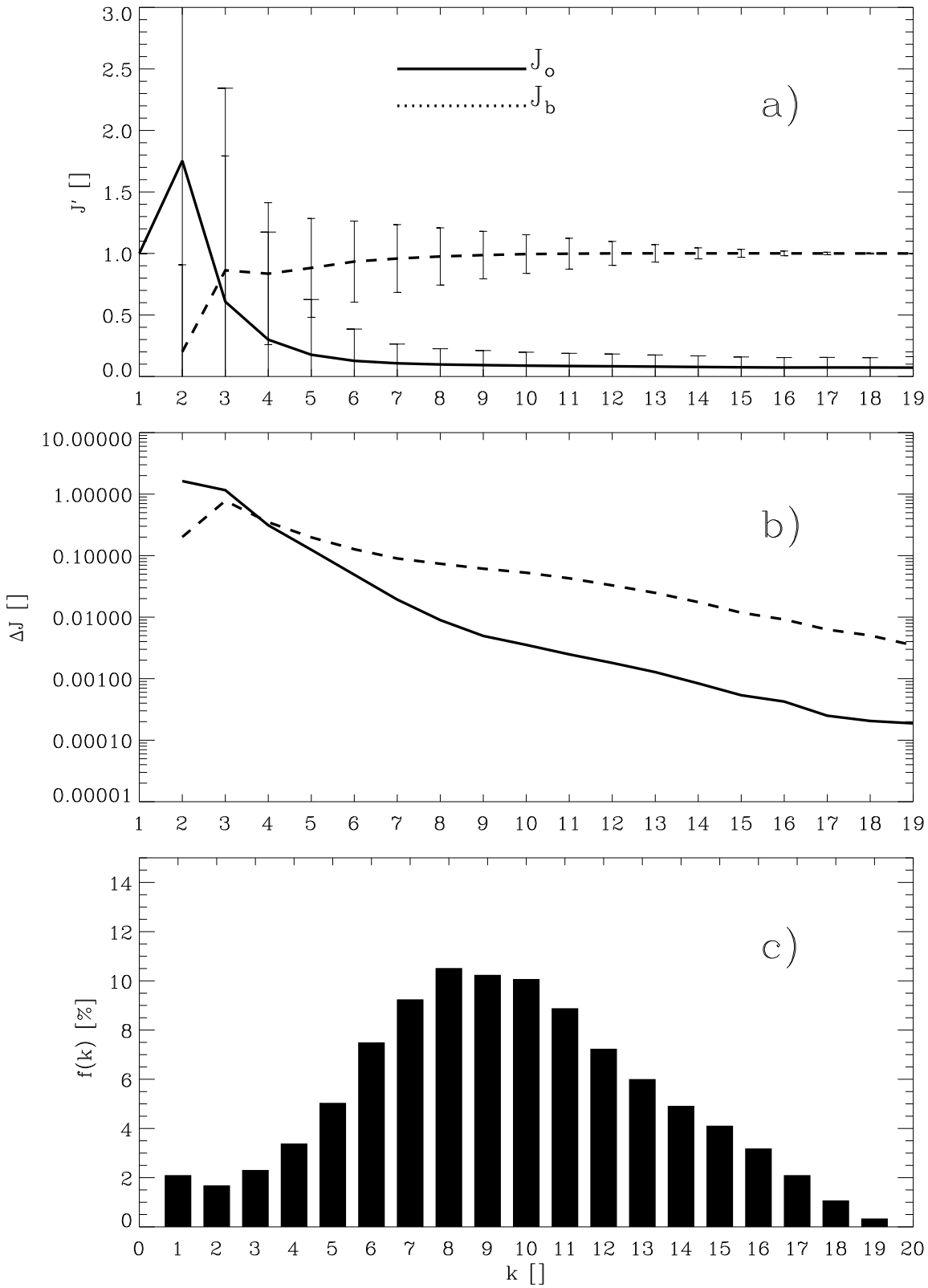

Figure 10: Mean and standard deviations of normalized 1D-Var cost-function, J', for observation (solid) and background (dashed) terms (a); gradient of cost-function (b), respectively, as a function of iteration number, k. Distribution of frequency of occurrence of 1D-Var iterations (c)

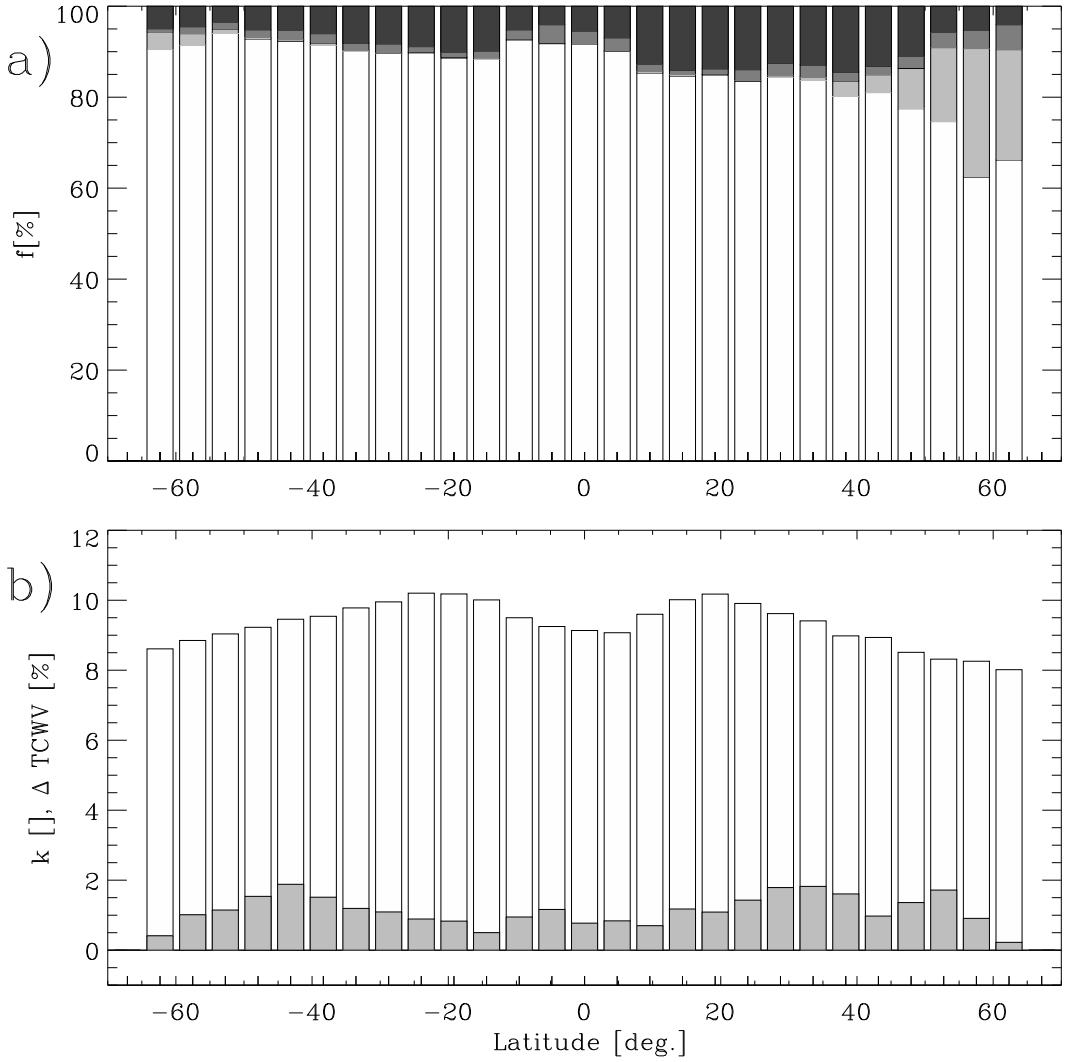

Figure 11: Zonal distribution of percentage of successful 1D-Var retrievals (white), too large TCWV analysis departures (light grey), too large TB analysis departures (dark grey) and convergence failures (black) for January 1-20, 2005 (a). Zonal distribution of mean 1D-Var iteration number (white) and mean relative TCWV increment (in %; light grey) (b).

slightly enhanced but rain is reduced below. Cloud liquid water and cloud coverage respond similarly in both cases with increments of the same sign along the profile.

Global increment statistics from the entire month of September 2004 are summarized in Figure 14 for TCWV (a), large-scale (b) and convective (c) precipitation, respectively. The ≈0.5% increase in TCWV from assimilating SSM/I TB's is reflected in the frequency distribution and propagates into those for both stratiform and convective precipitation. For precipitation, the increments are displayed as dBR that is $10log_{10}\Delta w_R$ with $\Delta w_R = w_{R,AN} - w_{R,FG}$. This is because, globally, precipitation follows a rather log-normal distribution.

The geographical distribution of relative increments is displayed in Figure 15. There are regions with very specific positive and negative TCWV increments (Figure 15a), respectively. Subtropical areas with lower rain intensities receive large positive relative TCWV increments with little effect on precipitation. Very small negative TCWV increments in the Northern oceans produce a significant reduction of large-scale and sometimes convective precipitation. In the Southern Indian Ocean, the strongest positive precipitation increments are produced while the strongest negative ones occur in the Southern Pacific. The areas of large convective rain

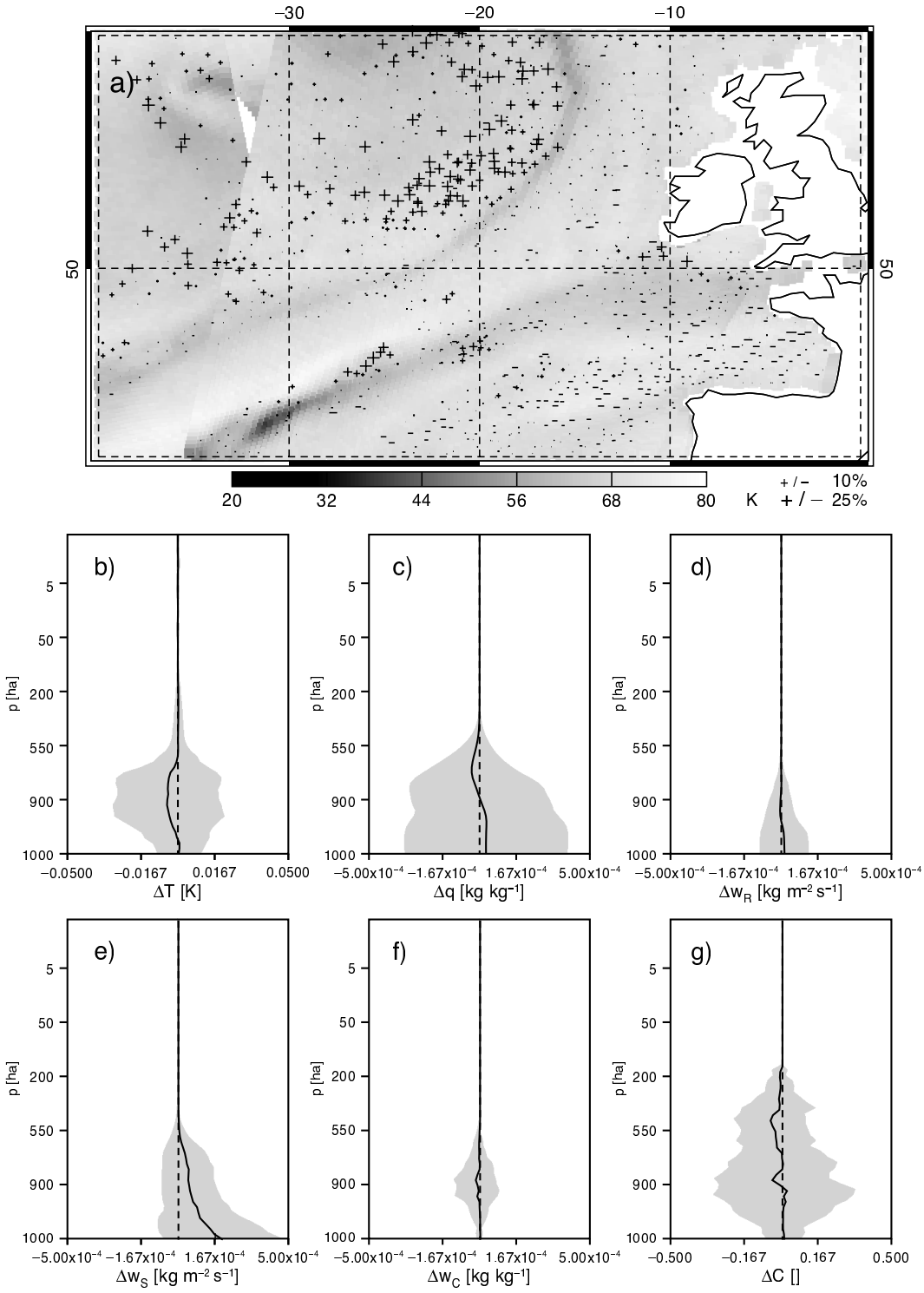

Figure 12: Example of relative TCWV increments superimposed on SSM/I polarization differences at 19.35 GHz (a) (in K). Symbols indicate sign of increments and symbol size denotes increment magnitude (see legend). Panels (b), (c), (d), (e), (f) and (g) denote averaged vertical increment distributions of temperature, specific humidity, rain flux, snow flux, liquid water mixing ratio and fractional cloud cover as a function of model level, respectively. Shaded area refers to ±1 standard deviation. 00 UTC analysis on December 31, 2004.

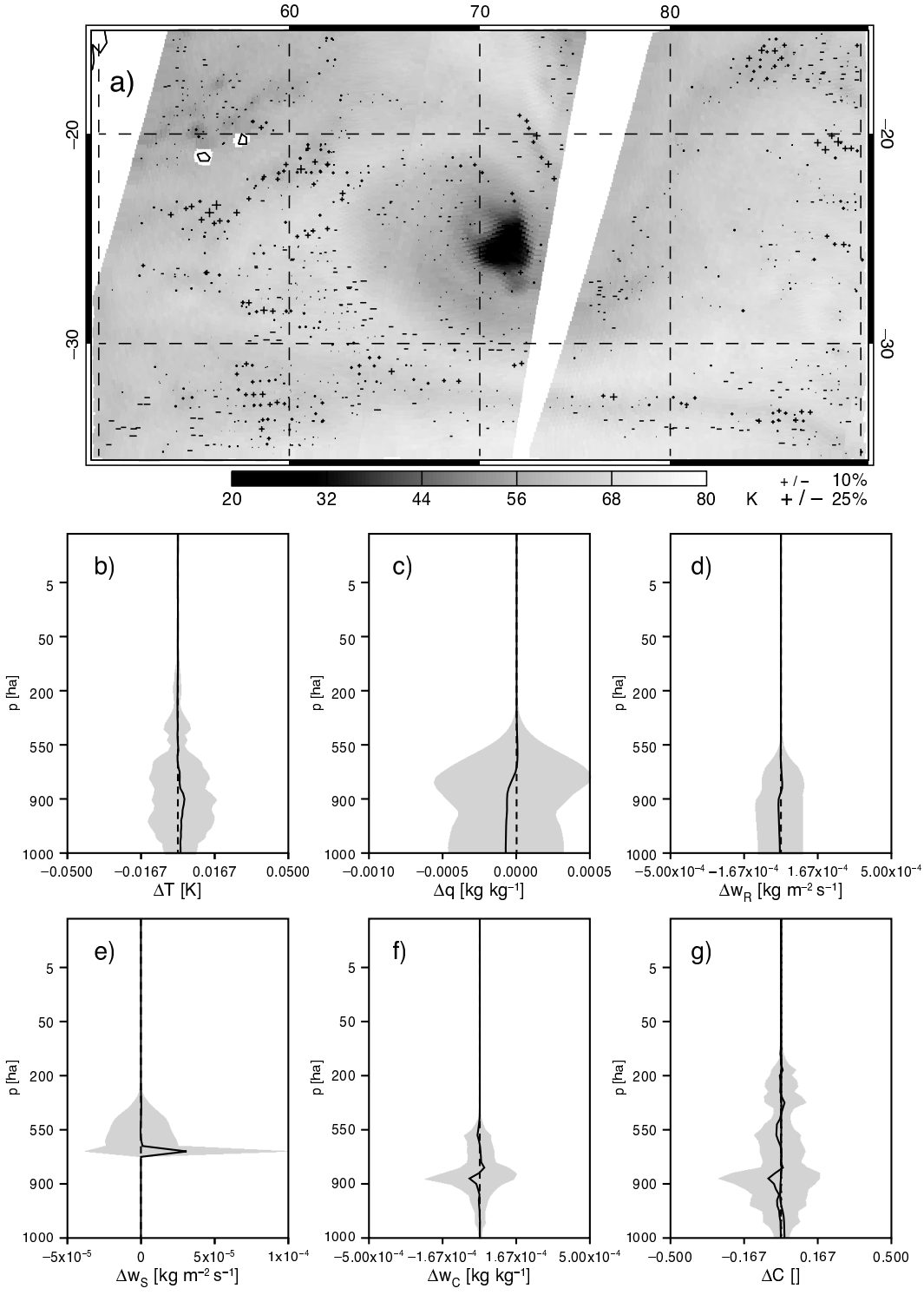

Figure 13: As Figure 12 for tropical disturbance No.6, 'Chambo', in the Southern Indian Ocean from 00 UTC analysis on December 31, 2004.

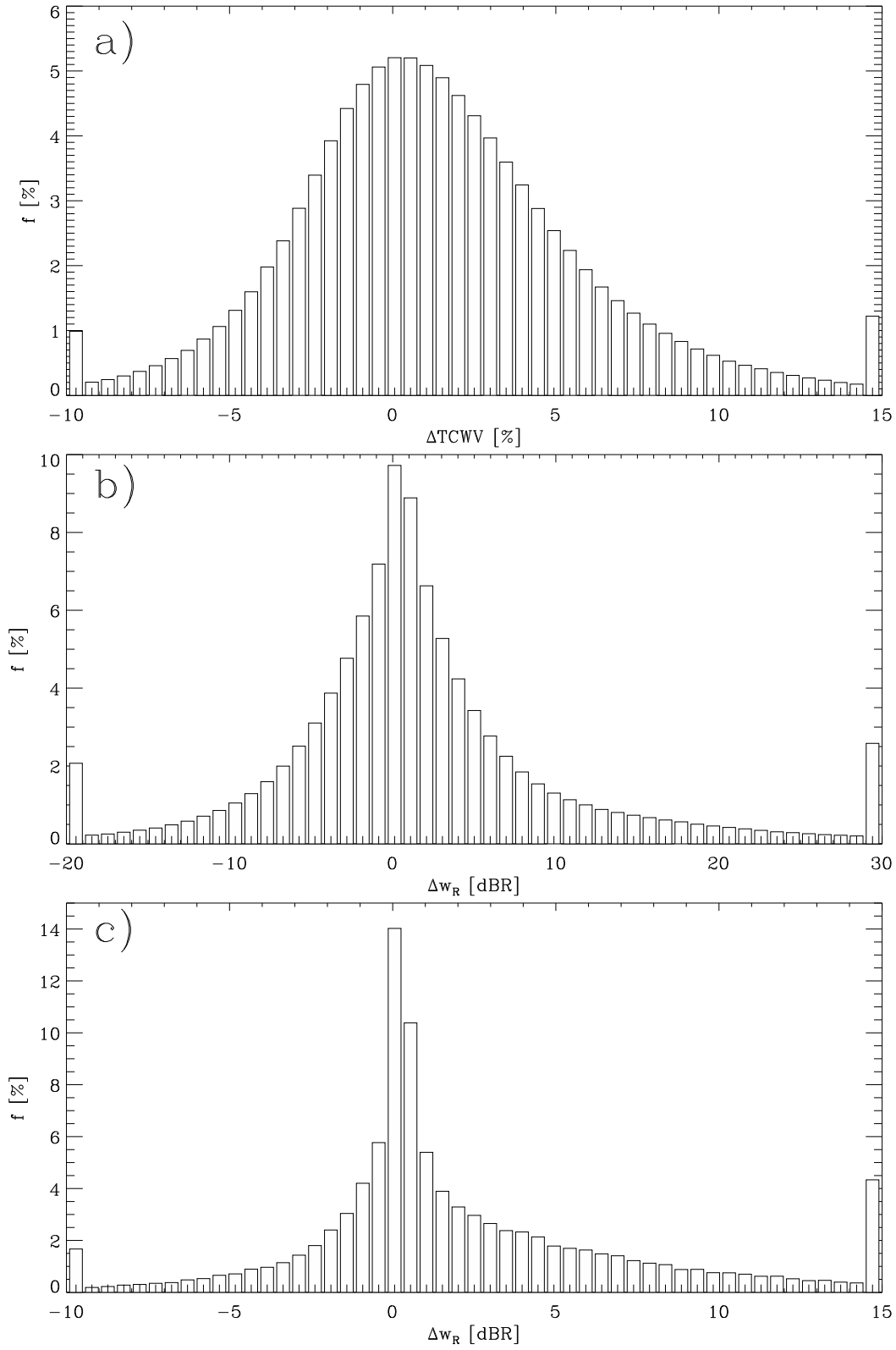

Figure 14: Frequency distribution of TCWV (a; in %), stratiform precipitation flux (b; in dBR) and convective precipitation flux (b; in dBR) increments from 1D-Var retrievals in September 2004.

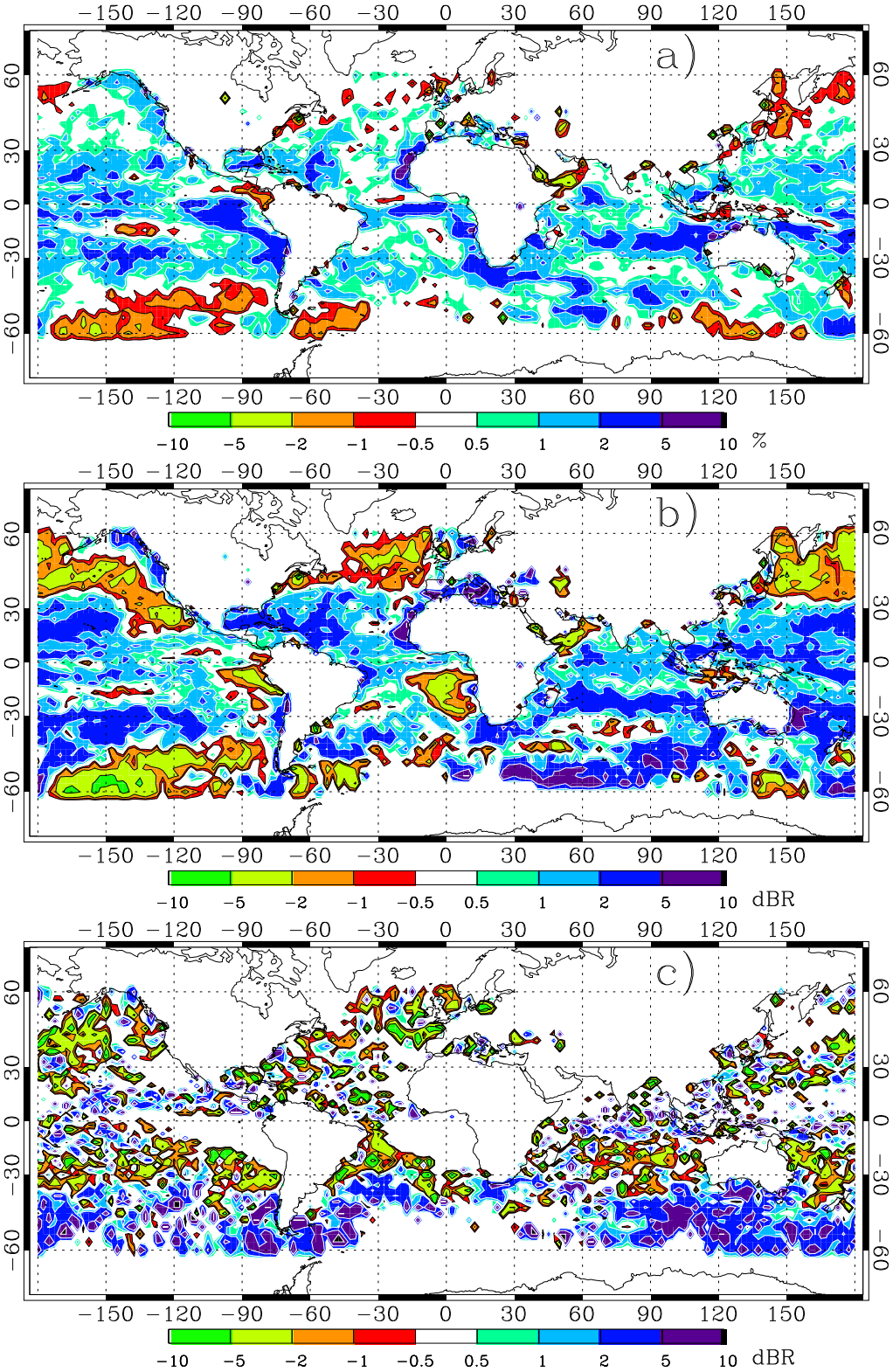

Figure 15: Global increment distribution of TCWV (a; in %), stratiform precipitation flux (b; in dBR) and convective precipitation flux (b; in dBR) from 1D-Var retrievals in September 2004 averaged to 2.5^o resolution.

increments correspond to only a few cases per month and to situations in which the first-guess convective rain rate is very small. Therefore, increments of, for example, 10 dB that is one order of magnitude still produce weak rain intensities.

Generally, the large-scale condensation scheme increments are smoother and cover the entire globe. Most of the time, the areas of positive and negative TCWV increments directly translate into increments of large-scale precipitation of the corresponding sign. This is because of the higher sensitivity of the large-scale precipitation parameterization to moisture changes compared to the convection scheme because the air is already saturated in the presence of large-scale condensation.. The largest rain increments occur in areas with little rain so that the global impact of the rain assimilation is rather weak (see also Figure 14). Local increments can be large even in the presence of significant amounts of rain, for example in the Caribbean Sea and near mid-latitude frontal systems. How much of this impact remains in the 4D-Var analysis and how this will affect the medium-range forecasts will be shown in the subsequent paper by Bauer et al. (2006b).

4 Discussion

This paper is the first part of the description of methodology and results of the assimilation of rain affected SSM/I radiances at ECMWF. Based on previous studies, a 1D+4D-Var approach was chosen where the radiances are used as observations in a 1D-Var retrieval of total column water vapour that is then assimilated in the 4D-Var system. The main development in the first stage is the observation operator that consists of linearized convection and large-scale condensation schemes as well as a multiple scattering radiative transfer model.

Even though the 1D-Var retrieval employs a non-linear minimization algorithm, the linearity of the observation operator had to be evaluated because it determines the quality of the convergence. The linearity of the moist physics parameterizations and the radiative transfer model were investigated separately for realistic perturbations deduced from analysis increments that were obtained from the 1D-Var retrieval itself. Only for rather intense rain events, the linearity seemed deteriorated but even for most of the situations, the differences between finite-difference and tangent-linear calculations remained within the 10% level. The linearity performance of the entire operator was found to be better than for the microphysics parameterizations alone that showed the weakest statistics for frozen precipitation. The dependence on SSM/I channel was found to be rather weak.

A screening and bias correction procedure was introduced that reduces the analyses to those cases where the 1D-Var retrieval can be expected to perform well and an introduction of TCWV biases into the 4D-Var analysis is thereby avoided. Compared to the departure statistics of SSM/I radiances in clear skies, the rain observation operator performs extremely well for the lowest three SSM/I frequencies. Radiance departure standard deviations are less than a factor of 3 larger than those in clear skies. This is a very good result given the complexity of the observation operator.

The Hollingsworth-Lönnberg method for the separation of background and observation error standard deviations was implemented. This indirect algorithm provided realistic results and produced balanced error values for all SSM/I channels. A direct observation error estimation is not possible because the largest contribution can be attributed to the observation operator for which no independent validation exists.

The 1D-Var convergence performance was analyzed for a three-week period in January 2005 inside the current operational system. The number of minimization failures is below 5% and most data rejections are due to TCWV increments that were considered too large. The associated thresholds will have to be refined in the context of the 4D-Var analyses. Examples of 1D-Var increment statistics were shown in a mid-latitude frontal system and a weak tropical cyclone. Globally, the TCWV and precipitation increments were positive for the chosen analysis date. This moistening of the atmosphere was expressed as an increase in both large-scale and

convective precipitation. A slight dependence on latitude was identified with maxima around ± 40 degrees of latitude in rather dry areas. The geographical distribution of large-scale and convective rain increments showed several distinct features. The large-scale scheme has a stronger sensitivity to moisture changes and is activated almost everywhere. Outside the tropics, the convection scheme shows large increments in areas where little rain is present.

In summary, the implemented 1D-Var methodology performed very well with the chosen observation operator. Microwave radiances are mainly sensitive to integrated hydrometeor contents and always show sensitivity to TCWV and to a lesser degree to atmospheric temperature. Therefore, even in the absence of background precipitation a minimization can be performed during which clouds and/or precipitation are generated. The error distributions are very smooth and, at least for the lower frequencies, show near-Gaussian shapes with reasonable error values. Part II of this paper will investigate the impact of TCWV pseudo-observations produced by the 1D-Var retrieval on the 4D-Var analyses and forecasts. This is a new observation type in areas that are usually screened out in satellite data analyses. The assimilation of rain affected radiances therefore has a large potential for future forecast improvements but the impact will greatly depend on the assimilation system itself and the interaction between moisture-related observations and model dynamics prescribed by the assimilation system and by the formulation of background error covariance statistics.

Acknowledgements

The authors are grateful to Jean-Noël Thépaut for many valuable discussions on the subject and to Marine Bonazzola for technical support. The work was partly funded by the European Space Agency (ESA) under contract No. 17193/03/NL/GS.

References

Bauer, P., 2001: Over-ocean rainfall retrieval from multi-sensor data of the Tropical Rainfall Measuring Mission (TRMM), Part I: Development of inversion databases. *J. Atmos. Ocean. Tech.*,**18**, 1315–1330.

Bauer, P., J.-F. Mahfouf, S. di Michele, F.S. Marzano, and W.S. Olson, 2002: Errors in TMI rainfall estimates over ocean for variational data assimilation. *Q. J. Roy. Meteor. Soc.*,**128**, 2129–2144.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, M. Bonazzola, and S. Saarinen, 2006a: Implementation of 1D+4D-Var Assimilation of Microwave Radiances in Precipitation at ECMWF, Part II: 4D-Var. *Q. J. Roy. Meteor. Soc.*, submitted.

Bauer, P., E. Moreau, F. Chevallier, and U. O'Keefe, 2006b: Multiple-scattering microwave radiative transfer for data assimilation. *Q. J. Roy. Meteor. Soc.*, submitted.

Chevallier, F. and P. Bauer, 2003: Model rain and clouds over oceans: Comparison with SSM/I observations. *Mon. Wea. Rev.*, **131**, 1240–1255.

Chevallier, F., P. Bauer, J.-F. Mahfouf, and J.-J. Morcrette, 2002: Variational retrieval of cloud profile from ATOVS observations. *Q. J. Roy. Meteor. Soc.*, **128**, 2511–2525.

Colton, M.C., and G.A. Poe, 1999: Intersensor calibration of DMSP SSMI's: F-8 to F-14, 1987-1997. *J. Geophys. Res.*, **37**, 418–439.

Deblonde, G., and S. English, 2003: One-dimensional variation retrievals from SSMIS-simulated observations. *J. Appl. Meteor.*, **42**, 1406–1420.

Errico, R., and K.D. Raeder, 1999: An examination of the linearization of a mesoscale model with moist physics. *Q. J. Roy. Meteor. Soc.*,**125**, 169–195.

Errico, R., L. Fillion, D. Nychka, and Z.-Q. Lu, 2000: Some statistical considerations associated with the data assimilation of precipitation observations. *Q. J. Roy. Meteor. Soc.*,**126**, 339–359.

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

Fillion, L., and R. Errico, 1997: Variational assimilation of precipitation data using moist convective parameterization schemes: A 1D-Var study. *Mon. Wea. Rev.*,**125**, 2917–2942.

Fillion, L., 2002: Variational assimilation of precipitation data and gravity wave excitation. *Mon. Wea. Rev.*,**130**, 357–371.

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

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*,**38A**, 111–136.

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

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

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

Marécal, V., J.-F. Mahfouf, and P. Bauer, 2002: Comparison of TMI rainfall estimates and their impact on 4D-Var assimilation. *Q. J. Roy. Meteor. Soc.*,**128**, 2737–2758.

Moreau, E., P. Bauer, and F. Chevallier, 2002: Variational retrieval of rain profiles from spaceborne passive microwave radiance observations. *J. Geophys. Res.*,**203**, D16, 4521, doi: 10.1029/2002JD003315.

Moreau, E., P. Lopez, P. Bauer, A.M. Tompkins, M. Janisková, and F. Chevallier, 2003: Rainfall vs. microwave brightness temperature assimilation: A comparison of 1D–Var results using TMI and SSM/I observations. *Q. J. Roy. Meteor. Soc.*,**130**, 827–852.

Phalippou, L., 1996: Variational retrieval of humidity profile, wind speed and cloud liquid-water path with the SSM/I: Potential for numerical weather prediction. *Q. J. Roy. Meteor. Soc.*,**122**, 327–355.

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

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

Saunders, R., P. Brunel, S.J. English, P. Bauer, U. O'Keeffe, P. Francis, and P. Rayer, 2005: RTTOV-8 Science and validation report. *NWP SAF Report, NWPSAF-MO-TV-007*, pp. 46.

Smith, E., P. Bauer, F.S. Marzano, C.D. Kummerow, D. McKague, A. Mugnai, and G. Panegrossi, 2002: Intercomparison of microwave radiative transfer models for precipitating clouds. *IEEE Trans. Geosci. Remote Sens.*, **40**, 541–549.

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

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

Tompkins, A.M.,and M. Janisková, 2004: A cloud scheme for data assimilation: Description and initial tests. *Q. J. Roy. Meteor. Soc.*,**130**, 2495–2518.