TECHNICAL MEMORANDUM

502

Assimilation of cloud and precipitation affected microwave radiances

Peter Bauer, Philippe Lopez,
Deborah Salmond and Alan Geer

Research Department

Paper presented to the SAC, 35th Session

November 2006

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

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

The first implementation for assimilating cloud and precipitation-affected observations over oceans in the operational 4D-Var data assimilation system was successfully performed in June 2005 with model cycle 29R2. The methodology consists of a 1D-Var retrieval of total column water vapour (TCWV) inside clouds and precipitation using Special Sensor Microwave / Imager (SSM/I) passive microwave radiance observations over oceans. The 1D-Var retrieval is otherwise constrained with standard short-range model forecasts of temperature, moisture, and near-surface wind speed, their forecast errors, an empirical estimate of observation errors and an observation operator that consists of moist physics parameterizations and a multiple scattering radiative transfer model. The TCWV retrievals are assimilated in the 4D-Var system as single-variable pseudo-observations in the areas that were identified as cloud-affected. The method is therefore referred to as 1D+4D-Var assimilation of rain-affected microwave radiances.

Given the current data usage in the ECMWF system, the introduction of rain-affected observations promises to represent a valuable upgrade. This is because radiance measurements from satellites that are sensitive to clouds have been screened in previous systems leaving large parts of mainly oceanic areas and lower tropospheric atmospheric levels virtually unobserved. The generally less accurate moisture analysis in the lower troposphere over oceans was only supported by clear-sky SSM/I observations so that the modelled diabatic processes were only indirectly affected by observations.

However, the assimilation of cloud and precipitation-affected observations is associated with problems some of which can only be pragmatically solved in current systems: The *observation operator* that simulates microwave radiances using model control variables as input comprises cloud and convection parameterizations and radiative transfer models that account for scattering of radiation at hydrometeors. The complexity of this operator may introduce large modelling uncertainties and may exhibit non-linear sensitivity of operator output to changes in state vector input. The latter is in particular critical in incremental 4D-Var systems that rely on linear model behaviour in the vicinity of the first-guess state. The definition of *background and observation errors* in the presence of clouds and precipitation can be assumed less accurate than in clear skies. The observation error consists of instrument noise and the observation operator error. Apart from systematic errors that may be compensated by bias-correction schemes, the random error component may only be indirectly derived. In the current ECMWF system, model background errors exhibit only little geographic variability and are not precipitation specific.

The implemented configuration has been chosen for the following reasons:

- The 1D-Var retrieval of an intermediate pseudo-observation (TCWV) that is to be assimilated in 4D-Var permits additional quality control based on 1D-Var retrieval diagnostics and minimization performance and alleviates the problem of non-linear observation operator behaviour. The disadvantages are that (1) the model first-guesses of moisture and temperature related to this observation are used twice, namely in 1D-Var and 4D-Var assimilation; and that (2) the choice of TCWV as retrieval variable may reduce the potential impact of the observation in a 4D-Var context.
- The spaceborne passive microwave observations from existing operational instruments have a similar spatial resolution than global forecast models which alleviates spatial representativeness issues. Moreover, microwave radiances at spectral window frequencies over oceans mainly respond to integrated precipitation water/ice paths and thus show little sensitivity to hydrometeor profile details. The signal dynamic range is very large with little contributions from other parameters such as sea-surface temperature, atmospheric temperature profile or other atmospheric constituents.
- There is long-term experience with such data in operational data assimilation systems in terms of data accuracy and stability and its impact in clear skies. SSM/I data can be considered the main driver in lower tropospheric moisture analysis over oceans (Andersson et al. 2006).

Prior to implementation, a significant number of tests have been run to assess the performance of both the 1D-Var retrieval and the 4D-Var assimilation of the retrieved TCWV and to address the above mentioned obstacles. Despite the complexity of the modelled processes and the generally uncertain precipitation representation in NWP models, the observation operator performed very well. Radiance biases are of the same magnitude as in clear skies and radiance departure standard deviations are 2-4 times larger. The 1D-Var

retrieval of TCWV only uses SSM/I channels at 19.35 and 22.235 GHz (1.55 and 1.35 cm wavelength) because the radiance biases were found to be small, departure probability distributions are nearly Gaussian, the observation operator is nearly linear, and observation errors could be indirectly derived from the spatial co-variance of radiance departures.

The 1D-Var retrieval exhibited problems in cases of very cold and dry atmospheres, mainly over Southern hemispheric oceans. These situations were associated with significant amounts of snowfall and with the activation of the linearized convection scheme in the minimization. Due to the lack of sensitivity of the employed radiances to frozen precipitation, the observation operator may produce too large increments and therefore unrealistic TCWV retrievals. These situations are screened in the latest version of the 1D-Var retrieval scheme. Another improvement was introduced by adding near-surface wind-speed in the 1D-Var control vector to account for the sensitivity of microwave sea-surface emissivity to surface roughness variations.

After more than one year of operations and several upgrades to the system, the following scientific conclusions can be drawn:

- 1D-Var and 4D-Var analysis increments of total column water vapour are consistent. Comparably large areas of systematic drying by the analysis can be identified along the Western coast of the American and African continent that persist until day 3 of the forecast. Smaller areas of moistening, e.g. in the Southern Indian Ocean, do not maintain the moisture increment because of the moisture removal by precipitation. Globally, the drying signal dominates. This is mainly the effect of the chosen moisture control variable and the definition of the associated errors in situations near saturation that is in the presence of clouds and precipitation. The humidity background errors are defined such that only limited supersaturation is permitted. On average, this asymmetry amplifies negative and penalizes positive moisture increments. This effect can only be overcome by a change of control variable, for example, a total water variable that contains water vapour as well as condensed contributions. This work is part of the 4-year plan.
- The initial version of the rain assimilation framework (model cycle 29R2) produced improved forecast scores of relative humidity in the tropics at around 700 hPa. Elsewhere there was little obvious improvement or degradation. A region of degradation between latitudes of 30 and 60 degrees South at T+48 hours was found to be linked to problems of excessive amounts of frozen precipitation and very low TCWV amounts in the northward intrusion of cold airmasses from the Antarctic. This problem has been corrected in model cycle 31R1 by an improved screening and bias correction.
- The changes made to the rain assimilation system between cycles 29R2 and 31R1 lead, in isolation, to neutral or slight positive impacts on forecasts. The area South of 30 degrees which was noted to be affected by large amounts of snowfall was particularly improved as was the area of initial positive impact in the Tropics at 700 hPa. With cycle 31R1, a new scheme for the dynamical bias-correction of radiances has been introduced that was found to potentially interact with the assimilation of TCWV from rain-affected SSM/I observations. The interaction lies in the above mentioned weak but sytematic drying of large parts of the oceans by the TCWV observations and the change in the treatment of other moisture-sensitive radiance biases with the new model cycle. The underlying mechanism is currently under investigation.
- The analysis of tropical cyclone forecasts for the rather active Caribbean season in 2004 showed a reduced track forecast error standard deviation when rain-affected observations were assimilated. Also, the fit to temperature and wind observations from radiosondes and dropsondes was improved. For tropical cyclone Katrina, the rain assimilation proved to produce better cylcone forecast with respect to centre pressure. Also here, the fit to temperature and wind observations from conventional sondes was improved. Despite these new observations in the vicinity of cyclones, the general performance of the moisture analysis and the still insufficient model spatial resolution dominate the forecast quality. All results were obtained with the initial model cycle 29R2 and the initial version of the 1D-Var retrieval. Since then, model resolution has greatly improved and substantial modifications of model physics, satellite radiance treatment and 1D-Var TCWV retrieval scheme have been implemented. Ongoing efforts will be dedicated to tropical cyclone forecast performance evaluation and the assessment of the individual contribution of rain-affected observations.

The experience with the operational version suggests that the main efforts for future developments have to be towards the implementation of a direct 4D-Var radiance assimilation system that avoids the intermediate 1D-Var retrieval. As for clear-sky observations, a stronger and more direct impact on model dynamics can be expected and the treatment of observation (operator) biases as well as the experimentation with more microwave radiometer channels will be greatly facilitated.

At this stage, an experimental version of a direct 4D-Var radiance assimilation in clouds and precipitation has already been developed. This represents the first available version in a global and operational NWP model. The implementation proves to be computationally more efficient than the operational 1D+4D-Var system and does not deteriorate the 4D-Var system's convergence or the linearity of the 4D-Var model operator. The experimental implementation will be tested and is expected to replace the 1D+4D-Var system in the near future. This can be considered a major accomplishment because the 4D-Var assimilation of cloud and rain affected radiances was considered impossible until very recently.

Other important improvements can be expected from a change of humidity control variable for the reasons mentioned above. If water vapour and cloud water stay separate, in particular the treatment of positive moisture increments and their distribution into gaseous and condensed state will remain difficult. The impact of moisture-sensitive satellite observations can be further optimized if model first-guess errors can be defined with a better distinction of clear-sky vs. cloudy areas.

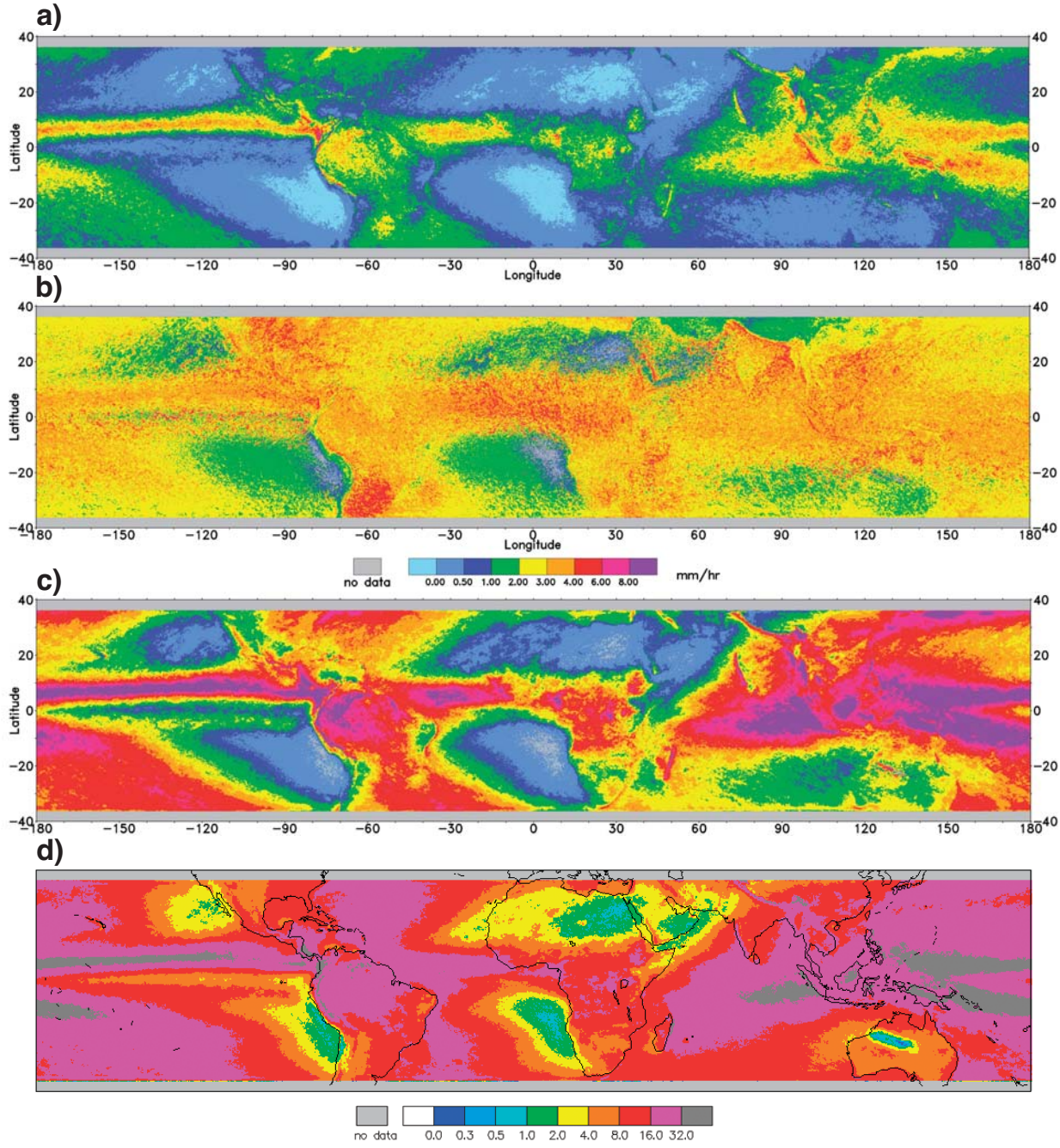


Figure 1: Mean tropical rainfall rate (a; in mm/h), conditional rainfall rate (b; = rainfall rate when raining in mm/h) and rain occurrence at 4 km (c; in % of valid observations) and at 26 km spatial resolution (d; similar to ECMWF model resolution; in % of valid observations) derived from TRMM PR data over period between January 1998 and December 2005 (Kidd, pers. communication).

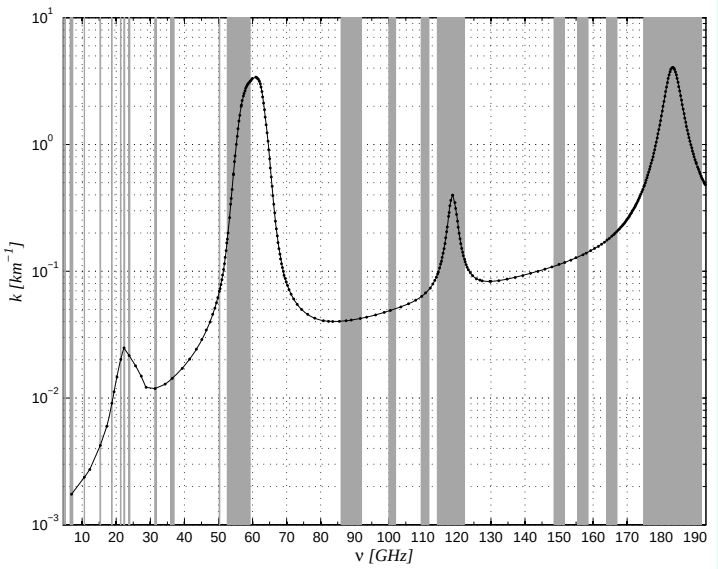
1 Introduction

Clouds and precipitation strongly affect the global hydrological and energy cycles, mainly through the interaction of solar and infrared radiation with cloud droplets and the release of latent heat in precipitation develop-

ment. The accurate modelling of clouds and precipitation in numerical weather prediction (NWP) systems is crucial for the forecasting of daily weather, in particular in case of extreme weather. Many catastrophic weather events are associated with extreme winds but also extreme precipitation. On climatological time scales, even droughts must be considered as weather extremes through the systematic and long-term anomalous lack of precipitation.

Microwave Radiances:

The microwave electro-magnetic spectrum comprises wavelengths, λ, between millimetres and centimetres. Microwave radiometer channel location is given in GHz with frequency $\nu = c/\lambda$ and the speed of light in vacuum c. The microwave spectrum that is currently used by operational meteorological instruments ($\nu \in [1 - 200\,GHz]$) is affected by oxygen (50-60, 118 GHz) and water vapour (22, 183 GHz) rotational line absorption as well as dry air (nitrogen, water vapour) continuum absorption. Millimetre-centimetre wavelengths interact with cloud droplets, rain drops and frozen particles due to absorption and scattering. Microwave radiances are expressed in black-body equivalent brightness temperatures, TBs, in units of degree Kelvin (K).



Clear-sky microwave absorption coefficient, k, for standard conditions (shaded areas indicated frequencies protected by International Telecommunication Union regulations).

About 3 million satellite observations[1] are currently assimilated per 12-hour analysis cycle at ECMWF (model cycle 30R1, June 2006). Those radiometer channel observations that may be affected by cloud and precipitation radiance contributions undergo screening checks to ensure clear-sky conditions. This is because the forward modelling of radiance signatures of clouds and precipitation is much less accurate than in clear skies. The screening checks are generally performed only based on observed information and neglect model information on clouds and precipitation. Most crucially affected are observations that are sensitive to the mid to lower troposphere and, in particular, to atmospheric moisture. Over sea, the most dominant observed humidity contributions in the analysis originate from Special Sensor Microwave / Imager (SSM/I) and Advanced Microwave Sounding Unit (AMSU) B observations. Over land, radiosondes, ground-based stations and AMSU-B observations dominate. In the upper troposphere, also infrared sounding channels from the High-resolution Infra-Red Sounder (HIRS), the Atmospheric Infra-Red Sounder (AIRS) and radiometers onboard geostationary satellites contribute significant information on humidity (Andersson et al. 2006).

Figure 1 highlights the distribution, intensity and occurrence of tropical rainfall. This climatology was obtained from Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (PR) observations between January 1998 and December 2005. The PR can be considered the most accurate spaceborne rainfall observation tool with an estimated rain detection threshold of 0.3 mm/h at 4 km spatial resolution and a calibration accuracy of 1 dBZ. The figure suggests that, in climatological terms, the main driver for regional differences is rainfall occurrence rather than rain intensity. In the tropical convergence zones, rainfall occurs at a rate of 8% and more while in areas with larger-scale precipitation systems

[1] multiple-channel observations are treated as independent samples

rainfall occurrence is of the order of 5%. These values apply only to the original spatial resolution of the PR. If they are averaged to the current approximate spatial resolution of the ECMWF model, rainfall occurence reaches about 25% and more in areas where convective precipitation regimes are dominant.

An example of data rejection due to cloud and precipitation occurrence is shown in Figure2 for SSM/I observations within a 12-hour data assimilation window used for the delayed cut-off analysis (DCDA) on June 1, 2006. Clearly, data coverage is greatly reduced over oceans. This means that, in particular over Southern Hemisphere oceans, large areas are not covered by observations that are sensitive to humidity so that the moisture analysis is mainly constrained by the model background (i.e. short-range forecast) in each analysis.

However, the interaction between humidity-related observations and the modelling system is not without problems. Several studies suggest that the systematic underestimation of atmospheric moisture by the model together with positive humidity increments from observations may lead to an imbalance in the hydrological budget that is accompanied by precipitation spin-up (e.g. Uppala et al. 2005). Recently, improvements in model physics and the data assimilation system helped to substantially reduce these problems (Andersson et al. 2005a). Apart from general predictability issues with regard to clouds and precipitation, the forecasting of processes involving the hydrological cycle is generally much less accurate than that of dynamical fields. This is mainly because of the less accurate moist physics parameterizations but also because moist processes strongly depend on less well predicted dynamical parameters like vertical wind.

Despite the underlying problems, both the continuous availability of accurate moisture observations and the progress in physical modelling and data assimilation system development pave the way for the use of cloud and precipitation affected observations in modern NWP systems. While these new observations represent greater technical challenges than existing ones, a substantial benefit for NWP skill improvement can be expected due to the lack of competing observations in cloudy areas.

The presented paper introduces a methodology for assimilating microwave radiometer observations in clouds and precipitation as it was operationally implemented at ECMWF in June 2005 (model cycle 29R2). The methodology will be referred to as one-dimensional plus four-dimensional variational (1D+4D-Var) system hereafter. The paper outlines the challenges imposed on the system and evaluates analysis and forecast impact of these new observations.

The paper begins with a short summary of past research on the topic (Section 2) and a description of the relevant issues of the technical implementation at ECMWF (Section 3). The impact evaluation is based on three systems, namely the first operational system (Section 4), its upgrade to be implemented with model cycle 31R1 (Section 4) and an experimental version of a full four-dimensional variational assimilation of cloud and rain affected microwave observations (Section 5). The paper is concluded by a summary including a discussion of outstanding issues and potential problems.

Summary

- In large areas of the tropics, rainfall occurrence exceeds 25% at the current ECMWF model resolution. If only clear-sky data is assimilated, large areas are sparsely sampled by satellite observations in which a substantial part of the atmospheric latent heat budget is determined.
- Moisture is less well represented in the model than temperature and wind which can lead to an imbalance of the hydrological cycle. This may also lead to systematic model moisture biases and spin-up effects caused by the impact of moisture sensitive observations on the analysis.
- The assimilation of cloud and precipitation affected observations represents a new challenge for data assimilation but can be expected to overcome some of these deficiencies in the future.

2 Past research

Since weather system development in lower latitudes is more dominated by convection and less by large-scale dynamics, the first attempts to produce better moist initial conditions for model forecasts were made through adjustments of diabatic heating using infrared sounder data (e.g. Heckley et al. 1990) and moisture initialization and heating adjustments from outgoing longwave radiation observations (OLR; Puri and Miller 1990). The development of more complex physical initialization schemes included the use of OLR and retrieved surface rain-rates from satellite data to constrain surface fluxes, moisture flux profiles and their feedback with convection and radiation (Krishnamurti et al. 1984, Krishnamurti and Bedi 1996).

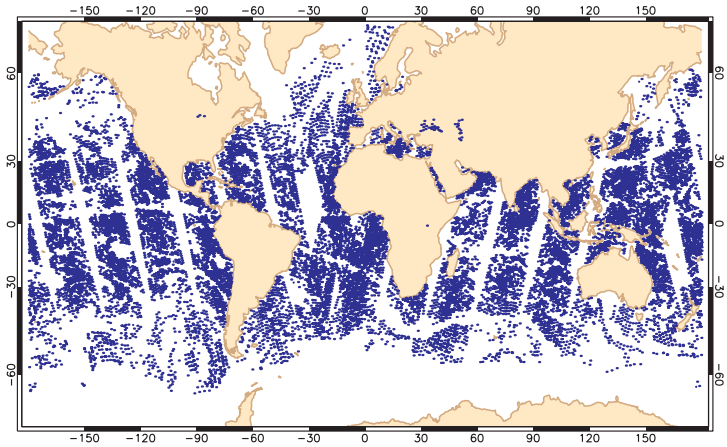


Figure 2: Active clear-sky SSM/I data distribution in operational ECMWF data assimilation on June 1, 2006 (DCDA analysis).

The advent of variational data assimilation schemes involving adjoint techniques initiated the employment of more complex and more consistent model physics in model initialization using cloud and precipitation observations. Apart from a number of experimental and case-study oriented methods (e.g. Zupanski and Mesinger 1995, Zou and Kuo 1996, Tsuyuki 1997, Kuo et al. 1997), only very few implementations in operational forecasting systems have succeeded so far. Currently, the National Center for Environmental Predition (NCEP) uses satellite derived rain-rates in their global 3D-Var system (Treadon et al. 2003) and the Japanese Meteorological Agency (JMA) uses ground-based observations of rain accumulation from radars and rain gauges in the operational meso-scale 4D-Var analysis (Tsuyuki et al. 2003). The ECMWF version of rainfall assimilation represents a unique step forward in that it is based on the variational assimilation of passive microwave radiance observations and that it performs the assimilation with an operational 4D-Var system on a global scale. The current operational version is a preparatory step towards the implementation of a direct 4D-Var assimilation of rain affected radiances. Both past and recent research highlight the following inter-connected issues to be addressed for cloud and rain data assimilation:

- Choice of observable: Derived vs. raw observation that can be instantaneous or accumulated near-surface precipitation, top-of-the-atmosphere infra-red/microwave radiance, or reflectivity profile. The choice may depend on the available observing system.

- Choice of data assimilation system: Depends on computational requirements and application. For global operational forecasts incremental 4D-Var systems are most prominent and most efficient in assimilating spatially/temporally inhomogeneously distributed data of different types. These systems' performance will be affected by non-linear observation operators and non-Gaussian error statistics.

- Choice of observation operator: Operators that provide the link between model fields and observation (e.g. radiances through moist physics parameterizations and radiative transfer models). May have to be linearized and regularized to perform in variational schemes.

- Definition of observation, plus observation operator errors: Quantification of biases and random error

components with Gaussian probability distributions. Error estimation may be more difficult for derived than raw observations.

- Choice of control variable: The control variable should also follow Gaussian error statistics and behave regularly near saturation. An explicit inclusion of condensed water (for example a total water variable) may be preferrable.
- Definition of model first-guess state errors: First-guess model field errors commonly do not differentiate between clear-sky and cloud/precipitation areas and lack geographical variation. Correlations between humidity and dynamic variables are not specified.
- Others: General degree of sensitivity of analysis to initial conditions and observations: Could be estimated by adjoint model studies and analysis sensitivity measures.

Some of these issues must be solved following pragmatic considerations. For example, observation error estimation may only be possible through indirect methods and non-linearity may only be avoided by applying strict screening. For details and a comprehensive discussion of these issues please refer to Errico et al. (2006).

Key Questions

- Which satellite observations are best suited for cloud and precipitation observation? Are they available from operational satellite series for continuous and real-time use?
- How accurately can satellite observations be simulated from current moist physics parameterizations and radiative transfer models?
- How relevant is the discrepancy between assimilating observations at model grid-points or observation locations?
- Is the current ECMWF model and data assimilation configuration suited for this purpose?
- Is the assimilation of cloud and precipitation affected observations computationally efficient?
- Does the impact of these observations overcome known model errors that are associated with the lack of observations in cloudy areas and the less accurate modelling of the atmospheric hydrological cycle?

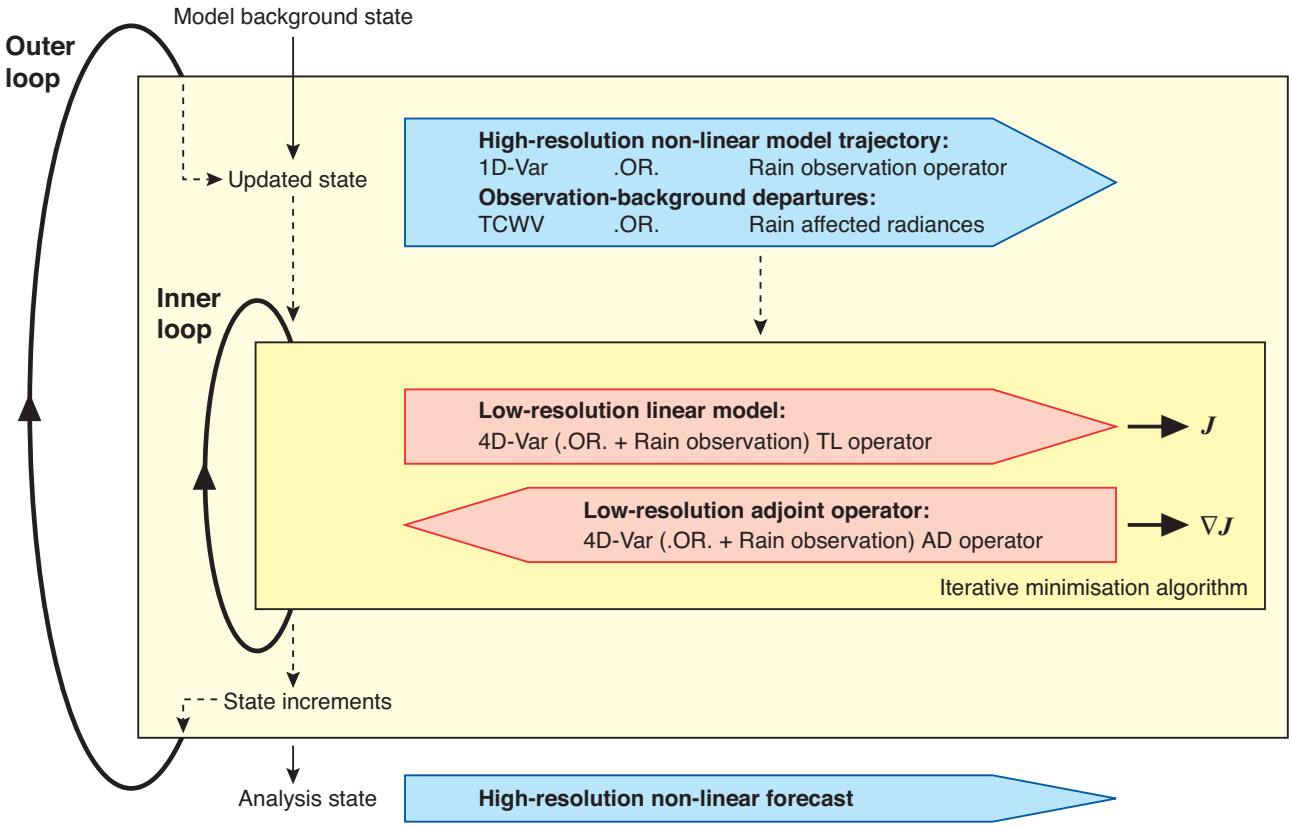


Figure 3: Basic incremental 4D-Var method implementation with optional 1D+4D-Var or 4D-Var assimilation of rain-affected radiances (after Trémolet 2004; '.OR.' refers to logical switch between 1D+4D-Var and direct 4D-Var assimilation option).

3 Operational implementation

3.1 1D+4D-Var system

In the past, several new satellite observations have been assimilated at ECMWF 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 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 cloud and precipitation 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 led to the implementation which became operational at ECMWF on June 28, 2005, were (1) the implementation of the 1D-Var retrieval using radiances in model grid-point space that is activated at each time step (currently 12 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.

Figure 3 illustrates the logical flow of information in the 1D+4D-Var methodology and compares it to the direct 4D-Var assimilation of cloud and precipitation affected radiances as presented by Trémolet (2004) for the 4D-Var algorithm in general. For enhanced computational efficiency, the incremental 4D-Var algorithm calculates departures between model state and observations at high spatial resolution and minimizes the cost-function (J) at lower resolution. These two steps are called 'outer loops' and 'inner loops', respectively.

Variational analysis:

In the variational analysis (e.g. Rodgers 2000), the optimum estimate of a state vector, $\mathbf{x}$ (here profiles of temperature, specific humidity, and 10-metre near-surface windspeed), is obtained using an observation vector (here TBs), $\mathbf{y}^o$, 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 Gaussian characteristics, the inversion problem can be formalized by the minimization of the well-known cost function, J:

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

with background error covariance matrix $\mathbf{B}$ and observation error covariance matrix $\mathbf{R}$. The latter includes the modeling error of operator H and the radiometer observation error.

In case of the 1D+4D-Var set-up, a 1D-Var retrieval of total column water vapour (TCWV) using microwave radiances as observations is performed and the high-resolution departures are calculated for TCWV's. In case of a direct 4D-Var assimilation of rain affected radiances, only the rain observation operator is applied at this stage, that produces model equivalent microwave radiances that are used for radiance departure calculations. In the minimization, the unchanged 4D-Var operators are applied for the 1D+4D-Var set-up that treat the derived TCWV values like direct moisture measurements. In the 4D-Var option, the tangent-linear and adjoint versions of the rain observation operator are applied together with the other components of the 4D-Var operator.

In both cases, the rain observation operator consists of a combined moist physics parameterization and 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 models 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 versions whose sensitivity to perturbations of the input parameters is more linear than that of the Tiedtke-parameterizations. This produces a more controlled behaviour 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. (2006c).

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 1D-Var retrieval algorithm uses temperature and moisture profiles as well as near-surface wind speed as the control vector and SSM/I radiance observations as observables. The control vector's first guess is taken from short-range forecasts, as in the operational 4D-Var system. The definition of model background errors and observation errors is briefly outlined in Section 3.3, more details can be found in Bauer et al. (2006a, b). The minimization itself is pre-conditioned and employs the M1QN3 algorithm developed by Gilbert and Lemaréchal (1989).

3.2 Observations

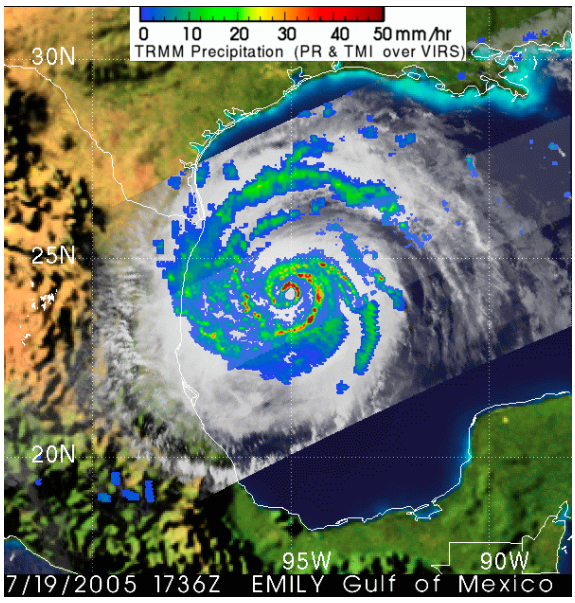


Figure 4: Hurricane Emily as seen by the VIRS, TMI and PR onboard the TRMM satellite on July 19, 2005 (courtesy NASA).

The currently available radiance observations cover visible (VIS) and infra-red (IR) wavelength ranges between 0.3-4 μm and 4-40 μm as well as the microwave part of the electro-magnetic spectrum between 6 and 200 GHz (i.e. between 5 cm and 1.5 mm). Of these, VIS and IR radiation does not penetrate water clouds and only provides bulk information on cloud top temperature and effective droplet size. Microwaves may penetrate clouds and precipitation depending on the choice of wavelength. For example, radiometer channels with frequencies between 10-20 GHz show the strongest sensitivity to rain water path while frequencies above 40 GHz become increasingly sensitive to cloud water and frozen precipitation. Except near strong atmospheric absorption lines (e.g. used for temperature sounding by AMSU-A radiometer) surface emission has to be taken into account due to the moderate atmospheric opacity at microwave frequencies.

In Figure 4, IR imagery from the TRMM Visible and Infrared Scanner (VIRS) has been overlayed on near-surface precipitation estimates from a combined PR (narrow swath) - TMI (wide swath) algorithm. The image shows the cloud vs. precipitation distribution and the lack of cloud penetration by VIS/IR radiation.

Figure 5 illustrates the microwave spectral sensitivity to perturbations of temperature, moisture, hydrometeor contents (cloud water, ice, liquid and frozen precipitation) as well as surface emissivity for an atmospheric profile over ocean. The figure suggests that there may be sensitivity to the clear atmosphere and the surface even in the presence of precipitation. Given the uncertainties in combined cloud-radiative transfer modelling,

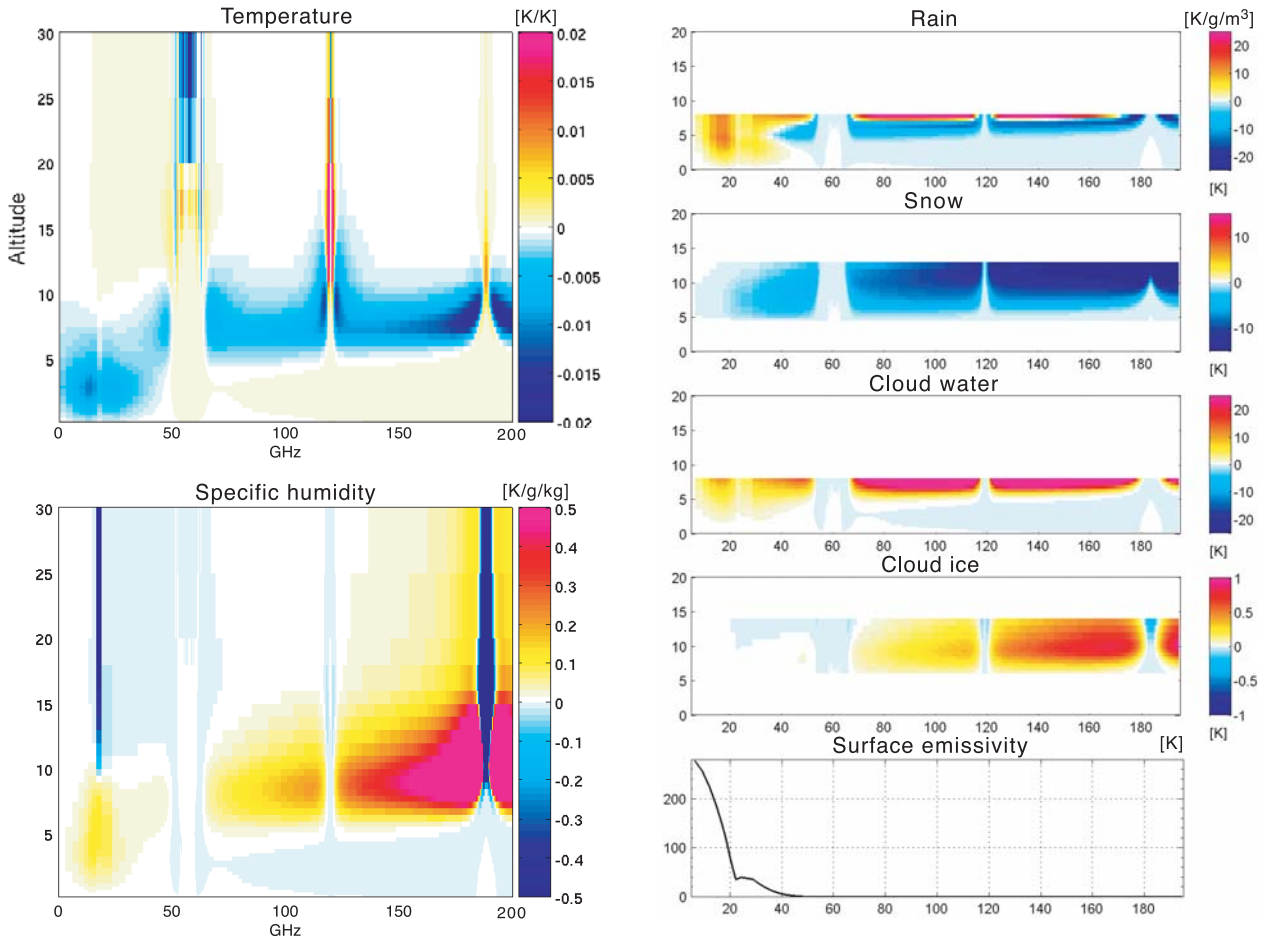


Figure 5: Sensitivity of microwave radiances between 10-200 GHz to perturbations in temperature, specific humidity, rain water content, snow content, cloud liquid water and cloud ice as well as surface emissivity for a tropical profile over ocean. Units are in degrees Kelvin (K) per unit change of the respective parameters.

the most appropriate frequency range is 10-20 GHz because scattering of radiation at frozen precipitation particles can be neglected and the main gaseous contribution is from integrated moisture only. At the same time, this choice requires a limitation to ocean surface applications because sea-water emissivity is modelled sufficiently accurate at most frequencies (Ellison et al. 2004).

Several satellite systems have provided observations in this frequency range for more than 20 years, the most prominent of which is the SSM/I that has been launched on several platforms of the Defense Meteorological Satellite Program (DMSP) since 1987. The SSM/I has seven channels at four microwave frequencies, namely window channels at 19.35, 37.0, and 85.5 GHz and a channel at the centre of the water vapour absorption line at 22.235 GHz. All window channels have two channels at vertical and horizontal polarization. For convenience, these channels will be denoted 19v, 19h, 22v, 37v, 37h, 85v and 85h hereafter.

Extended versions of this instrument have been available since the launch of the first Special Sensor Microwave Imager Sounder (SSMIS) in 2004 and will be continued after the replacement of DMSP with the National Polar-orbiting Operational Environmental Satellite System (NPOESS) foreseen for 2012. Additional experimental radiometers carry channels in this wavelength range such as the TRMM Microwave Imager (TMI) onboard the TRMM satellite and the Advanced Microwave Scanning Radiometer (AMSR-E) onboard the Aqua satellite.

The availability of long-term radiance observations at these frequencies from multiple satellites ensures data continuity and inter-calibration potential that is crucial for operational data assimilation and reanalysis.

3.3 1D-Var performance

Before radiance data can be assimilated, several fundamental issues have to be addressed, namely (1) the accuracy of the forward model operator, (2) the correction of systematic, potentially state dependent differences between observations and model predictions (= biases), (3) the definition of observation operator errors, and (4) the degree of non-linearity of the observation operator. These are described in more detail in Errico et al. (2006) and their technical solution in the ECMWF system is given in Bauer et al. (2006a).

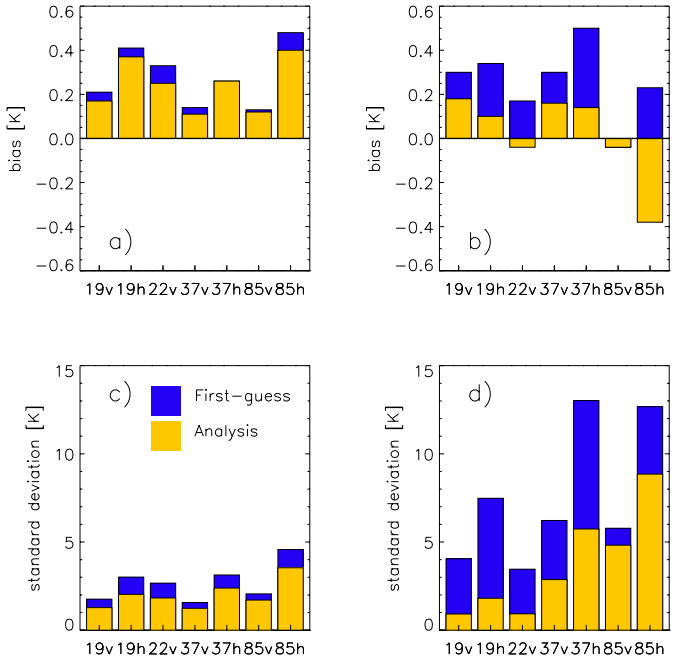


Figure 6: Bias-corrected first-guess and analysis departure statistics (in K) from clear-sky and rain affected SSM/I radiance assimilation. Mean observation-model differences for clear-sky (a) and rain affected (b) radiances for all SSM/I channels as well as standard deviations for clear-sky (c) and rain affected (d) radiances. First-guess departures in blue and analysis departures in yellow.

Biases were found to be surprisingly small and are corrected using a static parametric bias correction with TCWV, sea-surface temperature (SST), near-surface wind speed and precipitation water path as predictors. The originally implemented version (model cycle 29R2) only employed TCWV. The bias correction coefficients stabilize when data samples from 2-3 week periods are chosen. Figures 6 and 7 show examples of first-guess and 1D-Var analysis departure statistics. For comparison, similar statistics have been generated from the clear-sky SSM/I radiance assimilation in the ECMWF 4D-Var analysis.

The mean bias-corrected first-guess departures (Figure 6) show that the remaining radiance biases are very similar for both clear-sky and rain assimilation data. The first-guess departure standard deviations of rain affected radiances are between 1.5-4 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 departure standard deviations are smaller by 20% than the first-guess departures in clear conditions; this reduction inside clouds and precipitation amounts to 70% for the first three active channels.

Figure 7 shows, however, that the bias-correction works successfully only for the lower three SSM/I channels and cannot be significantly improved by a parametric bias correction method. The deviation of the departure distributions from a Gaussian shape also indicate that effects of non-linearity and both inaccurate first-guess model fields and observation operator inaccuracies are more dominant at higher microwave frequencies. This is supported by the linearity assessment carried out by Bauer et al. (2006a) that indicates a dependence of non-linearity on microwave frequency and convective activity. Therefore, only the lower three channels were

Table 1: 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. Background error distributions in radiance space from radiance transform calculations for cases with active large-scale condensation (LS) and convection (CV) schemes (bold figures indicate active 1D-Var channels).

SSM/I channel	**19v**	**19h**	**22v**	37v	37h	85v	85h
Observational method:							
Background error standard deviation	**2.2**	**4.1**	**2.0**	3.5	7.4	2.8	7.7
Observation error standard deviation	**2.8**	**5.2**	**2.5**	4.4	9.0	4.0	8.4
$\mathbf{HBH}^T$ *calculation:*							
Background error standard deviation, LS	**5**	**8**	**6**	5	10	5	12
Background error standard deviation, LS+CV	**20**	**37**	**7**	8	22	4	16

considered in the further analysis while the other four channels were maintained active in the 1D-Var retrieval for diagnostic purposes. With this set-up the minimization performance is satisfactory as is the reduction of cost function and cost function gradient (Bauer et al. 2006a).

The minimization is constrained by both background and observation-plus-modeling error covariance matrices. The background errors are 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 spatial resolution (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. Bauer et al. (2006b) performed 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 the production of 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. As an alternative solution, the background errors in radiance terms could be calculated from the application of the observation operator to the background error covariance matrix. However, this would also introduce modeling errors and would not produce a clean separation of error contributions.

Applying all different options leads to error estimates as displayed in Table 1. From the Hollingsworth-Lönnberg (= observational) method, 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.

The magnitude of background and observation errors obtained from the Hollingsworth-Lönnberg method are rather similar. Another method of background error calculation in radiance space is to apply the observation operator (moist physics and radiative transfer) to the temperature and moisture background error covariance matrix. This calculation produces error estimates that are 2-3 times larger than those obtained from the

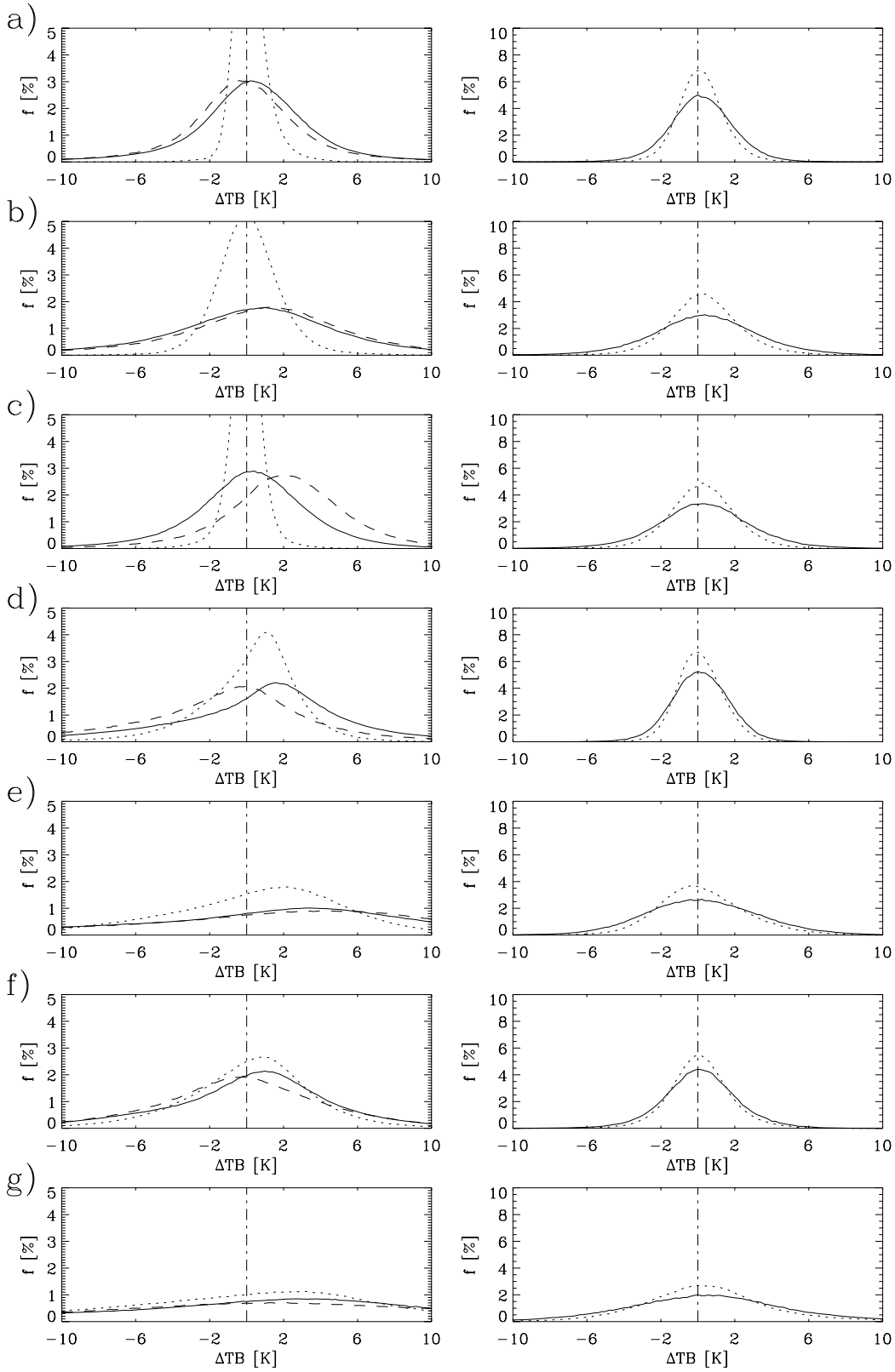


Figure 7: TB-departures between observations and bias-corrected first-guess model values (solid), analysis values (dotted) and uncorrected first-guess model values (dashed; for rain affected observations only). Left panels show distributions for cloud and rain simulations, right panels show clear-sky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

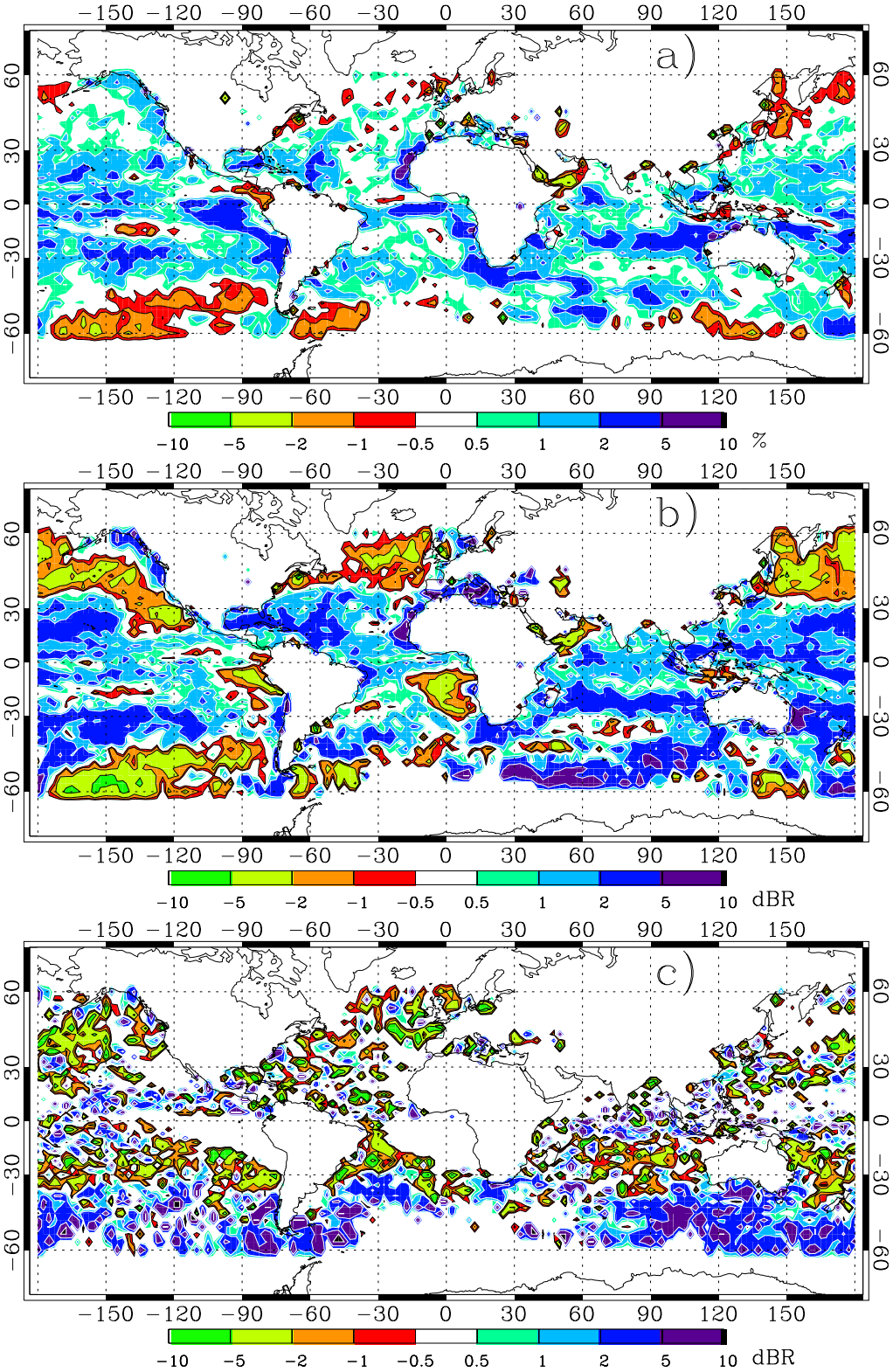


Figure 8: Global increment distribution of TCWV (a; in %), stratiform precipitation flux (b; in dBR) and convective precipitation flux (b; in dBR) from 1D-Var retrievals in September 2004 binned to 2.5 ^{o} resolution.

Hollingsworth-Lönnberg method. This may suggest that the B-matrix in clouds and precipitation contains too large covariances. As shown in Table 1, 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. This means that the combination of background errors and sensitivity of the observation operator to changes in model state do not reproduce the same error statistics that are obtained from comparison with observations.

3.4 Global 1D-Var results

Global increment statistics from the entire month of September 2004 are summarized in Figure 8 for TCWV (a), large-scale (b) and convective (c) precipitation, respectively. 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 shaped distribution.

There are regions with very pronounced positive and negative TCWV increments (Figure 8a), 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 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. The largest rain increments occur in areas with little rain so that the global impact of the rain assimilation is rather weak. However, 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.

3.5 1D-Var Case study: A good and a bad case

Two case studies of 1D-Var retrievals in the South Atlantic in the southern hemisphere winter were examined to demonstrate two different situations in which the 1D-Var retrieval succeeded or failed in reproducing realistic hydrometeor profiles and in matching the observed brightness temperatures. These case are illustrative of a local pattern limited to Southern hemispheric oceans in which the 1D-Var assimilation was responsible for a degradation in forecast scores. At high southern latitudes the 1D-Var retrievals showed areas of large and consistently positive TCWV increments. These areas are associated with the northward movement of very cold, dry, polar air. It was hypothesised that these areas of consistently large increments were responsible for the poor forecast scores.

Figure 9 shows the meteorological situation at 12 UTC on 14th August 2005. An area of high pressure lies in the subtropical South Atlantic in its usual position. There are low pressure systems near the Antarctic coast at 30W and 30E. Southwest of South Africa, an area of very low TCWV moves northwards, indicating an intrusion of cold polar air. The corresponding Meteosat visible channel image (Figure 10) shows a line of convective cloud along the cold front associated with this northward-moving air. Behind the front, the cold airmass shows

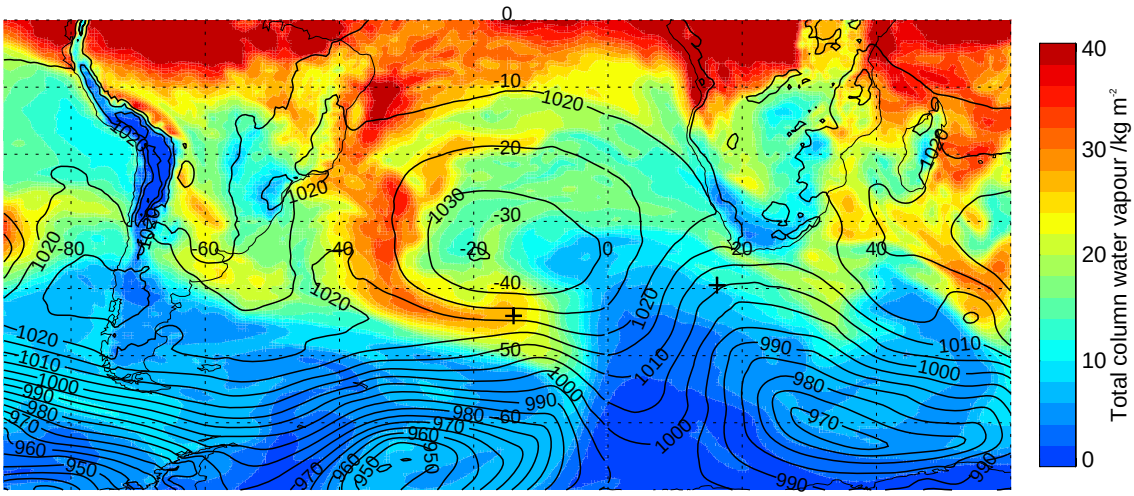


Figure 9: Total column water vapour (kg m^{-2}) and pressure at mean sea level (hPa) for 12 UTC, 14th August 2005, from the operational ECMWF analyses. Crosses at (A) 14W, 44S and (B) 16E, 40S indicate the locations of two selected examples of 1D-Var retrievals.

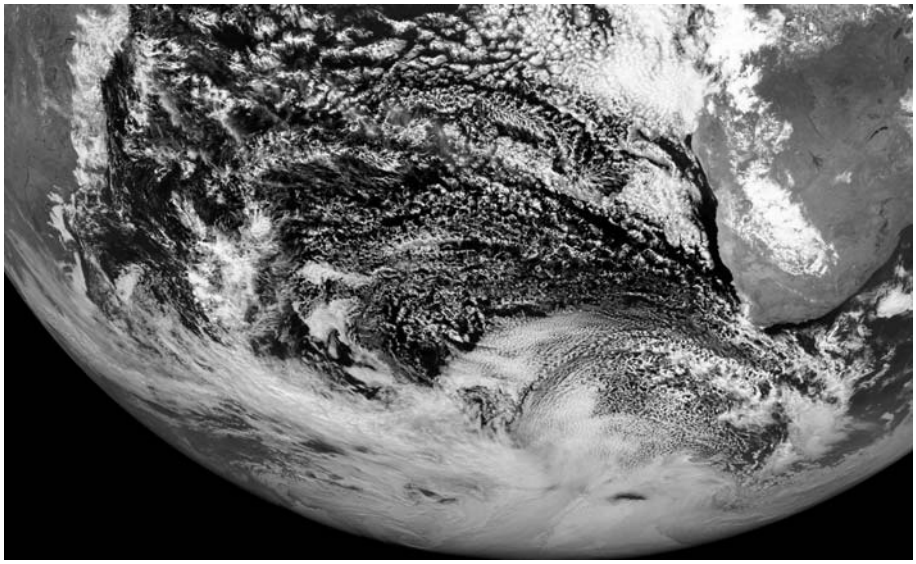


Figure 10: Visible (channel 1) image from Meteosat-8 showing the South Atlantic at 12UTC, 14th August 2005. Image copyright EUMETSAT 2005 and courtesy NERC Satellite Receiving Station, Dundee University, Scotland (http://www.sat.dundee.ac.uk/).

widespread scattered convective cloud. Further west, a band of high TCWV intrudes south-eastwards, moving around the subtropical high, associated with an area of heavy cloud cover (Figure10).

We examined two 1D-Var retrievals: (A) is located in the band of high TCWV and heavy cloud cover, where the 1D-Var appears to be performing well; (B) is located in the region of the cold front, where because of the large TCWV increments associated with 1D-Var, the retrievals are thought to be less reliable. See Figure9 for the exact locations.

Table 2 shows the first guess and analysed brightness temperature observation-minus-model departures from the 1D-Var retrieval at point (A). Only channels 19v, 19h and 22v are actively assimilated. However, the

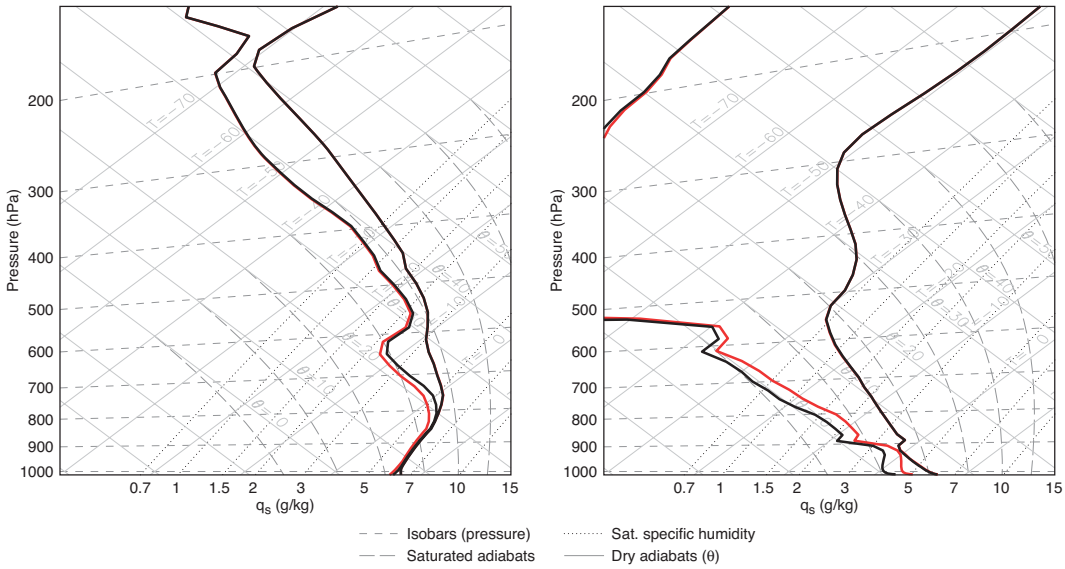


Figure 11: Tephigram showing the 1D-Var temperature (°C, right hand lines) and dewpoint temperature (°C, left hand lines) versus potential temperature at first guess (black line) and analysis (red line) from example (A) at 14W, 44S (a) and example (B) at 16E, 40S (b) on 14th August 2005. The analysed temperature is essentially no different to the first guess.

analysis departures are substantially smaller than those of the first guess in all seven available channels, giving confidence in the quality of the retrieval. The 1D-Var control vector contains the temperature and moisture profile, and the 10m windspeed.

Figure 11a shows the first guess and retrieved temperature and moisture profiles. The air temperature is close to dewpoint throughout the column, consistent with the heavy cloud cover seen in the visible satellite image (Figure 10). The temperature profile remains essentially unchanged between first guess and analysis, but water vapour amounts have been reduced in the lower and mid-troposphere. Table2 shows that this results in only a relatively small change in total column water vapour.

Table 2: Total column water vapour and brightness temperature departures from 1D-Var retrieval for examples (A) and (B). Bold columns indicate channels actively assimilated. The others are monitored but not assimilated.

	TCWV [kg m^{-2}]	Radiance departures [K]						
SSM/I channel		**19v**	**19h**	**22v**	37v	37h	85v	85h
Example (A):								
First guess	24.4	**-8.1**	**-13.6**	**-6.5**	-15.7	-30.0	-7.8	-20.6
Analysis	23.5	**-0.6**	**0.3**	**0.3**	-1.9	-2.0	-2.6	-5.7
Example (B):								
First guess	7.6	**4.5**	**7.9**	**2.0**	4.2	8.8	4.1	12.0
Analysis	8.8	**2.0**	**1.3**	**-1.1**	3.4	-1.8	13.3	5.5

Figure 12a-c shows precipitation and cloud parameters from the 1D-Var retrieval from example A. These quantities are generated by the moist physics part of the 1D-Var observation operator. Convection was not active in this profile; the cloud has been generated by the stratiform (large-scale) cloud scheme. This simulates three main cloud layers, with the upper two principally formed from ice. The lowest layer, between 800 hPa and the surface, is formed of liquid water and generates rain. The effect of reducing moisture in the 1D-Var retrieval is to reduce the cloud amount and its water content, and to reduce the precipitation amount. The upper part of Table 2 shows the effect on the simulated brightness temperatures. In general, SSM/I brightness temperatures

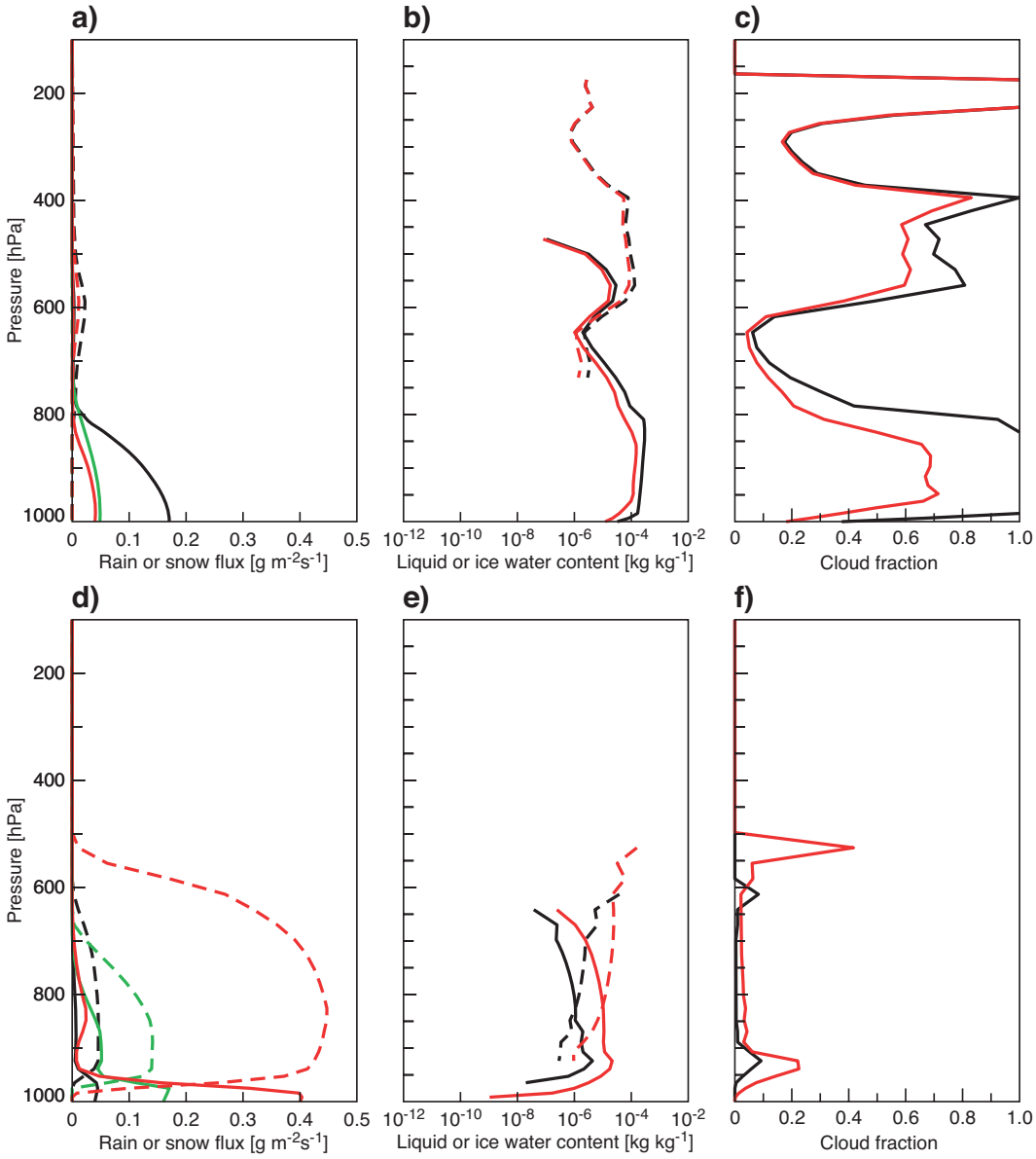


Figure 12: Rain (a,d; solid line) and snow (a,d; dashed) fluxes, liquid (b,e; solid line) and ice (b,e; dashed) cloud water content and cloud fraction (c,f) from a 1D-Var retrieval at 14W, 44S (a, b, c; example A) and at 16E, 40S (d, e, f; example B) on 14th August 2005. Black lines show the 1D-Var first guess and red lines the 1D-Var analysis. Green lines show the 3D-non-linear model's first guess (rain and snow only).

increase with increasing liquid cloud and rain amounts (see e.g. Chevallier and Bauer 2003). Here, first guess brightness temperatures were too high compared to observations; reducing the amount of cloud and rain has decreased brightness temperatures to better match the observations.

Despite the satisfactory performance of the 1D-Var retrieval for case A, Figure 12a indicates one area where 1D-Var retrievals are problematic. The green line shows the first guess rain profile calculated by the full non-linear moist physics scheme in the 3D-model. Given the same moisture and temperature profile, the simplified moist physics of 1D-Var simulates a substantially higher first guess rain amount. In fact, the 1D-Var analysis reduces rain amounts to those of the 3D-model. Ongoing research is devoted to reduce the different performance between non-linear and linearized moist physics parameterizations to minimize the potentially negative effect

of this problem.

We now examine retrieval (B), in the region of the cold front south of Africa. Here, the first guess total column moisture is just 7.6 kgm^{-2} (Table 2), compared to 24.4 kgm^{-2} in example (A). Surface air temperature is 7°C, similar to example (A), but the atmosphere above is much colder and drier (Figure 11b). The tephigram indicates a potential for convection between roughly 950 hPa and 550 hPa in the first guess profile. The 1D-Var analysis substantially increases water vapour amounts at levels below 550 hPa, making the column even more convectively unstable. Although column moisture has changed by a similar TCWV amount in both examples, in example (B) the increment is much larger relative to the background.

The lower part of Table 2 shows that first guess brightness temperatures were too low. Figure 12 suggests that by increasing the amount of moisture, the 1D-Var analysis has caused rain and cloud water amounts to increase, thus increasing brightness temperatures to better match SSM/I observations.

In this example, precipitation is mainly in the form of snow, melting to rain only near the sea surface. This precipitation has come mainly from the convection scheme. Significant cloud fractions are found only at 900 hPa, where the large scale cloud scheme is active, and at 500-600 hPa, where convective detrainment is simulated. In the 1D-Var analysis, the detrainment altitude and the cloud fraction at this level are substantially larger, indicating much stronger convection, as might be expected from the tephigram (Figure 11). Once again, the 1D-Var first guess rain profile is rather different to that from the 3D-model.

A clue as to what may be happening is provided by the analysis brightness temperature departures (lower part of Table 2). The 1D-Var analysis matches observations well at 19v/h, 22v and 37v/h. However, there are still large departures in the 85v and 85h channels. In contrast to the other SSM/I channels, 85v and 85h are sensitive to snow. In contrast to the influence of rain and cloud, snow tends to reduce 85v/h microwave brightness temperatures (see Chevallier and Bauer 2003). Hence these departures suggest that the 1D-Var analysis has too much snow. Why are channels 85v and 85h not actively assimilated? Microwave radiative transfer is thought to be less reliably simulated for snow than for rain and scattering of radiation at frozen hydrometeors produces larger deviations from linearity. Nevertheless, these large departures still provide an indication that there is a problem with the retrieved snow amounts (see also Figure 7 for non-Gaussian TB departure distributions). These situations are currently screened out by applying a threshold on integrated snow amounts and forecast scores in the Southern hemisphere were greatly improved as a consequence (Section 4.2).

Summary

- The combined moist physics - radiative transfer model observation operator performs very well. Radiance departure statistics reveal biases that are similar to SSM/I observations in clear skies and departure standard deviations that are 2-3 times larger.
- A stable radiance bias correction and an empirical formulation for observation-plus-modelling errors has been developed that produces an excellent performance of the 1D-Var minimization algorithm. The 1D-Var algorithm performance introduces another level of quality control that is used to prevent unwanted effects on the 4D-Var analysis.
- Since both the simplified moist physics and microwave radiative transfer may be less accurate for snow than for rain, problems of 1D-Var convergence may occur in very cold, dry regions associated with the northward intrusion of polar air in winter. TCWV increments can then be very large as a fraction of the background amount. However, these cases can be identified from the 1D-Var performance analysis and are removed from the assimilation.

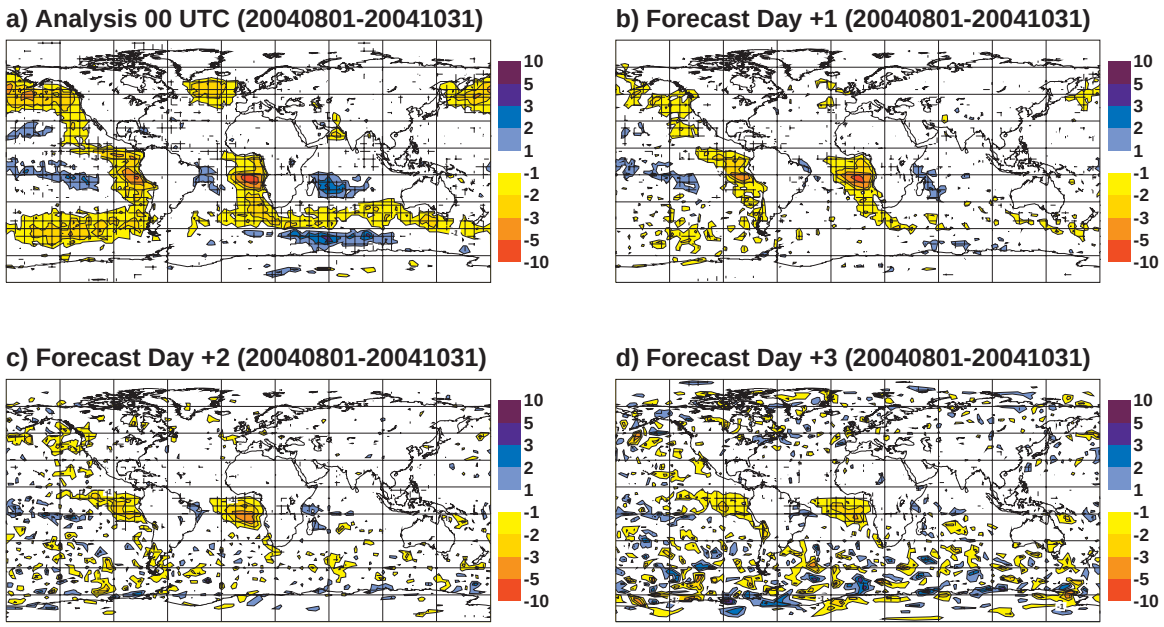


Figure 13: Mean normalized analysis (a) and 1-day (b), 2-day (c), 3-day (d) forecast difference of TCWV between RAIN and NORAIN experiments for period August-October 2004. Units are in % and crosses indicate statistical significance (95% significance interval with t-test on analysis differences against zero-difference).

4 4D-Var Impact evaluation

4.1 Analysis increments

The assimilation of SSM/I radiances in rain affected areas produces moisture increments that are expected to modify the hydrological cycle in the model and change dynamical patterns to accommodate the redistribution of moisture. This impact is not carried out through the background error covariance statistics because no cross-correlation exists between moisture and dynamics in 4D-Var. The feedback can only be produced by the 4D-Var that establishes the dependence of the analysis on the temporal evolution of the control vector during the minimization.

Experiment *RAIN* was carried out with model cycle 29R2 in which the rain assimilation was activated for the first time. In the control experiment (*NORAIN*), the observations were deactivated in the minimization. Both experiments were run at 40 km resolution (T511) and on 60 model levels. Figure 13 shows a three-month average of normalized TCWV differences between *RAIN* and *NORAIN* from the analyses (Figure 13a), the 24-hour (Figure 13b), 48-hour (Figure 13c), and 72-hour (Figure 13d) forecasts.

The TCWV-differences were normalized with the control TCWV analyses/forecasts to show the relative impact

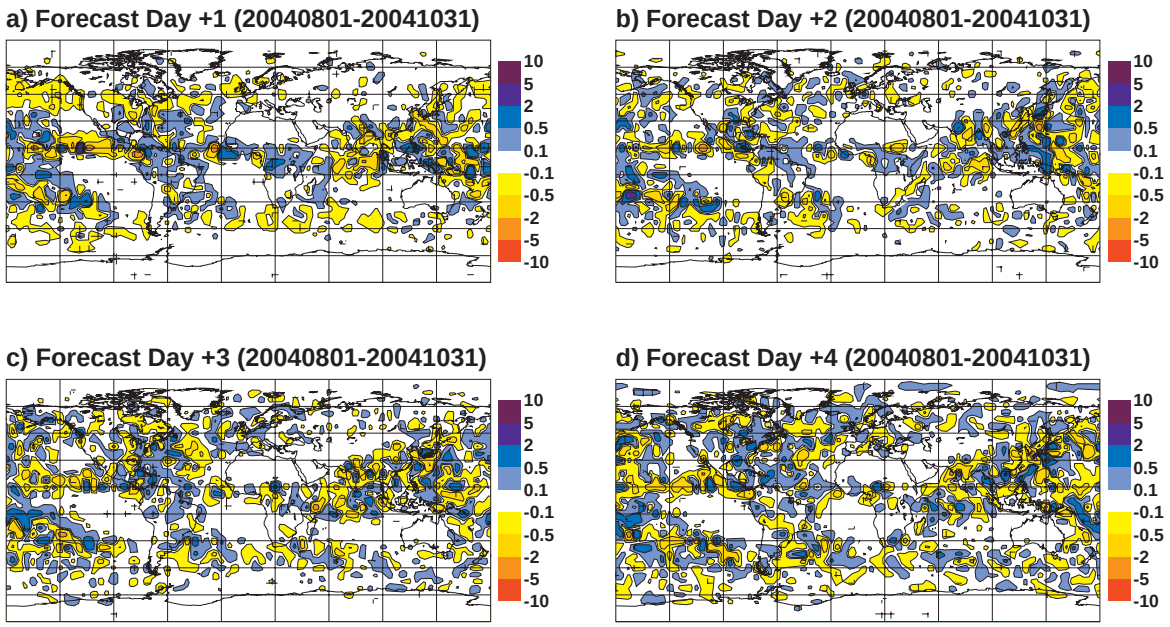


Figure 14: Mean 24-hour accumulated precipitation forecast difference for 1-day (a), 2-day (b), 3-day (c), and 4-day (d) forecasts between RAIN and NORAIN experiments for period August-October 2004. Units are in mm and crosses indicate statistical significance (95% significance interval with t-test on analysis differences against zero-difference).

and to avoid that the effect in mid-latitude areas is underestimated due to the lower moisture abundance. The hatched areas in Figure 13 indicate where the differences are statistically significant. The significance was calculated with a t-test applied to the TCWV-differences against the zero-difference and with a 5% significance level.

The mean analysis differences clearly identify areas of mean drying along the Western coasts of the American and African continents, the North Atlantic and Pacific and the latitude band between 30-50 degrees South. Positive differences are more localized just off the ITCZ and South of 50 degrees in the Indian Ocean. The amounts are of the order of a 1-3 %; larger amounts of -5 to -10 % are observed rather locally at 10 degrees South off the Western coasts of South America and Africa. The moistening generally remains below 5%.

The 4D-Var moisture differences are fairly similar to those produced by the 1D-Var (Figure8) even though the latter show more spatial detail because the 1D-Var retrievals are spatially independent. Some areas of local 1D-Var moistening in very dry subtropical regions are not reproduced in the 4D-Var analyses. The global 4D-Var picture generates a net drying of the atmosphere while the 1D-Var global statistics had revealed a small net moistening (Bauer et al. 2006a). This is explained by the difference in the control variables and their background error statistics. In the 1D-Var, specific humidity and temperature served as a control variable, while they are combined to a scaled relative humidity control variable in 4D-Var (Andersson et al. 2005a).

In areas of clouds and precipitation where the observations are mostly assimilated, the relative humidity is usually near saturation. The background error formulation in 4D-Var implies a non-linear behaviour in that moistening in nearly saturated areas is strongly penalized (small background errors) while drying is less penalized (larger background errors). Near saturation, a negative TCWV increment produced by the 1D-Var will therefore persist through the 4D-Var analysis while a positive TCWV increment may be suppressed.

Figures 13b-d show how the analyses are propagated into the forecast. Again, mean TCWV differences between *RAIN* and *NORAIN* are shown. From this illustration, the memory of the moisture increments from the rain assimilation seems to last longer for the dried areas (3+ days) and shorter for the moistened areas (1-2 days). The explanation for this is that the moist increments will accelerate condensation and produce more precipitation that removes the moisture from the atmosphere. However, only the larger areas with negative moisture increments can be identified in the forecasted precipitation differences (see Figure 14) between 24 and 96 hours. The patterns are rather noisy and suggest that most of the positive moisture increments are removed by precipitation very quickly.

The differences are generally more noisy in the Southern hemisphere because of the more active large-scale diabatic processes during winter. As already noted from the 1D-Var analyses, the global impact on the precipitation budget is small which means that the hydrological cycle in terms of global mean precipitation is not significantly changed.

Information content:

The individual impact of observations inside a data assimilation system that is constrained by a large number of diverse observations can be evaluated from the information content that is associated with an individual observation. Recently, this concept was formalized based on the ECMWF analysis system in terms of the so-called 'self-sensitivity' estimation that is based on the influence matrix (Cardinali et al. 2004). This matrix quantifies the sensitivity of the analysis value to individual observations. Moreover, the trace of the influence matrix provides the total information content from which the information content (here in terms of the degrees of freedom of the signal, *DFS*) of each observation type can be derived.

The table below summarizes *DFS* of each observation type from a single analysis experiment. Observation types with large data volumes therefore show comparably large contributions such as Atmospheric Infrared Sounder (AIRS) and Advanced Microwave Sounding Unit (AMSU-A) satellite data. However, the TCWV observations exhibit a significant contribution per single observation that is about 4 times larger than clear-sky SSM/I radiances, 1.5 times larger than AMSU-B and 2 times larger than HIRS data. It has to be kept in mind that *DFS* depends on the definition of observation errors thus observations with small errors have more weight in the analysis.

Observation type	*Number*	*DFS [%]*	*DFS per single observation*
AIRS	1,280,872	30.94	0.0863
AMSU-A	627,019	17.91	0.1020
Aircraft	151,745	6.66	0.1567
HIRS	94,886	5.84	0.2198
Temp	66,029	5.03	0.2720
Satob	110,704	4.95	0.1596
TCWV	**45,742**	**4.76**	**0.3715**
Meteosat	101,585	4.25	0.1494
AMSU-B	66,038	4.22	0.2280
Quikscat	113,702	3.64	0.1142
Pilot	53,004	3.31	0.2228
Synop	62,186	2.87	0.1648
GOES	59,162	2.84	0.1714
SSM/I	84,164	2.06	0.0876
Drift buoys	4,564	0.60	0.4730
Ozone	11,010	0.11	0.0362

Information content (DFS in %) per observation type and single observation in 12 UTC analysis on August 27, 2005 (GOES: Geostationary Operational Environmental Satellite, AMSU: Advanced Microwave Sounding Unit, HIRS: High-resolution Infrared Sounder, AIRS: Atmospheric Infrared Sounder, SSM/I: Special Sensor Microwave / Imager).

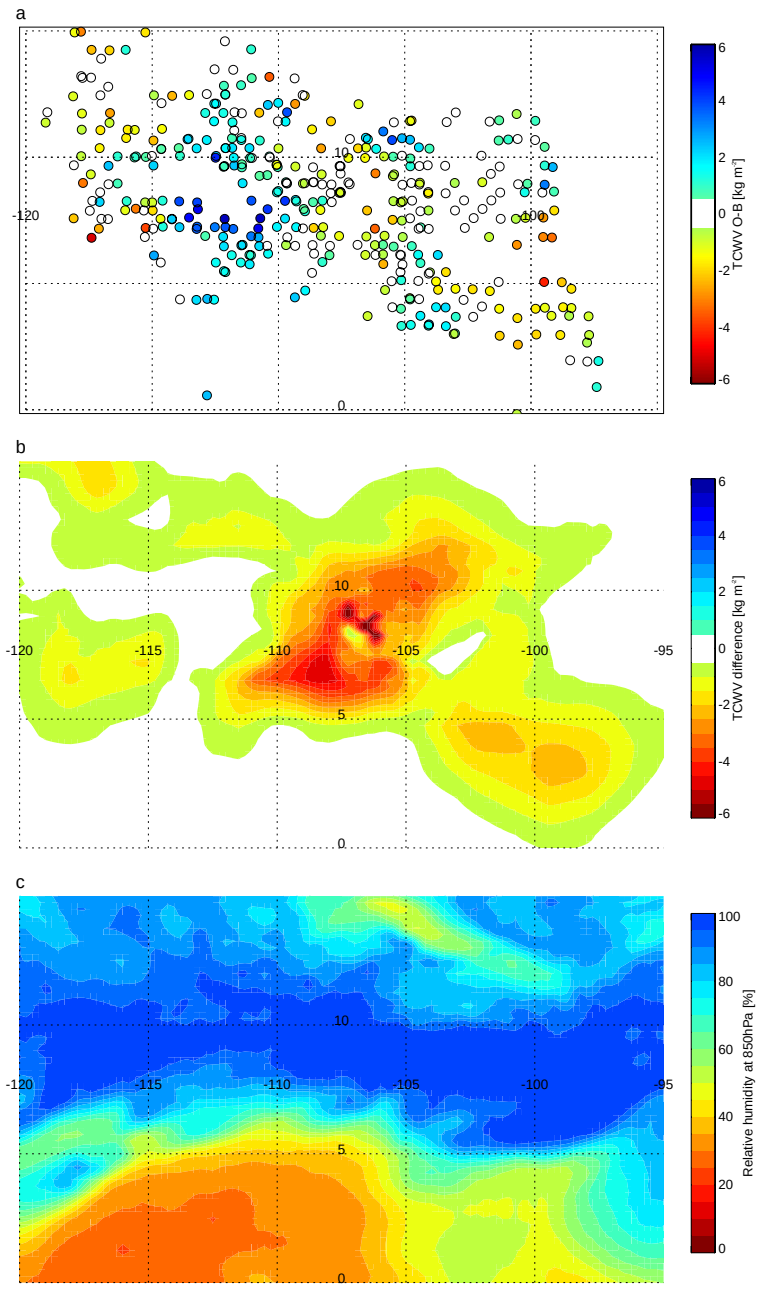


Figure 15: a) Observed-minus-background TCWV at the locations of 1D-Var retrievals used in the 00 UTC analysis of experiment RAIN, 1st August 2005; b) Difference in TCWV increments between RAIN and NORAIN experiments at 21 UTC, 31st July 2005 (both experiments use the same background state for this analysis); c) relative humidity at 850 hPa at 21 UTC, 31st July 2005.

As an example, TCWV increments in the ITCZ in the East Pacific are examined. Figure 15 shows first guess departures for 1D-Var retrievals of TCWV during one analysis cycle, and their resulting impact on TCWV in the 4D-Var analysis. There are regions of positive and negative departures (Figure 15a), however, the TCWV increments are biased towards negative values (Figure 15b). The relative humidity control variable is normalised such that positive increments are strongly penalised when the background is near saturation. Figure 15c shows that in much of this region the relative humidity is indeed near 100%. These high relative humidities are associated with the ITCZ and the high occurrence of clouds and precipitation from deep convection. In this moist region, positive departures have therefore been suppressed while negative departures result in drier 4D-Var analyses.

Figure 16a shows the mean sea-level pressure differences between *RAIN* and *NORAIN* for the same period as Figures 13. As expected, areas with systematic drying by the new observations exhibit enhanced surface pressure and vice versa. The dynamic response through the 4D-Var analysis system is illustrated in Figure 16b that shows the impact of the precipitation assimilation on the analyzed divergent wind component at 850 hPa. Again, the statistical significance is indicated by hatched areas. Regions of systematic convergence match well with those where the mean moisture increments are positive, that is in the southern and tropical Indian Ocean and the northern part of the SPCZ. The most intensely dried areas (off western coasts of South America and Africa) show strong local divergence that is statistically significant. As mentioned before, this connection between moisture and convergence is a product of 4D-Var and would be less well reproduced within, e.g., a 3D-Var system unless cross-correlations between these parameters would be defined for the

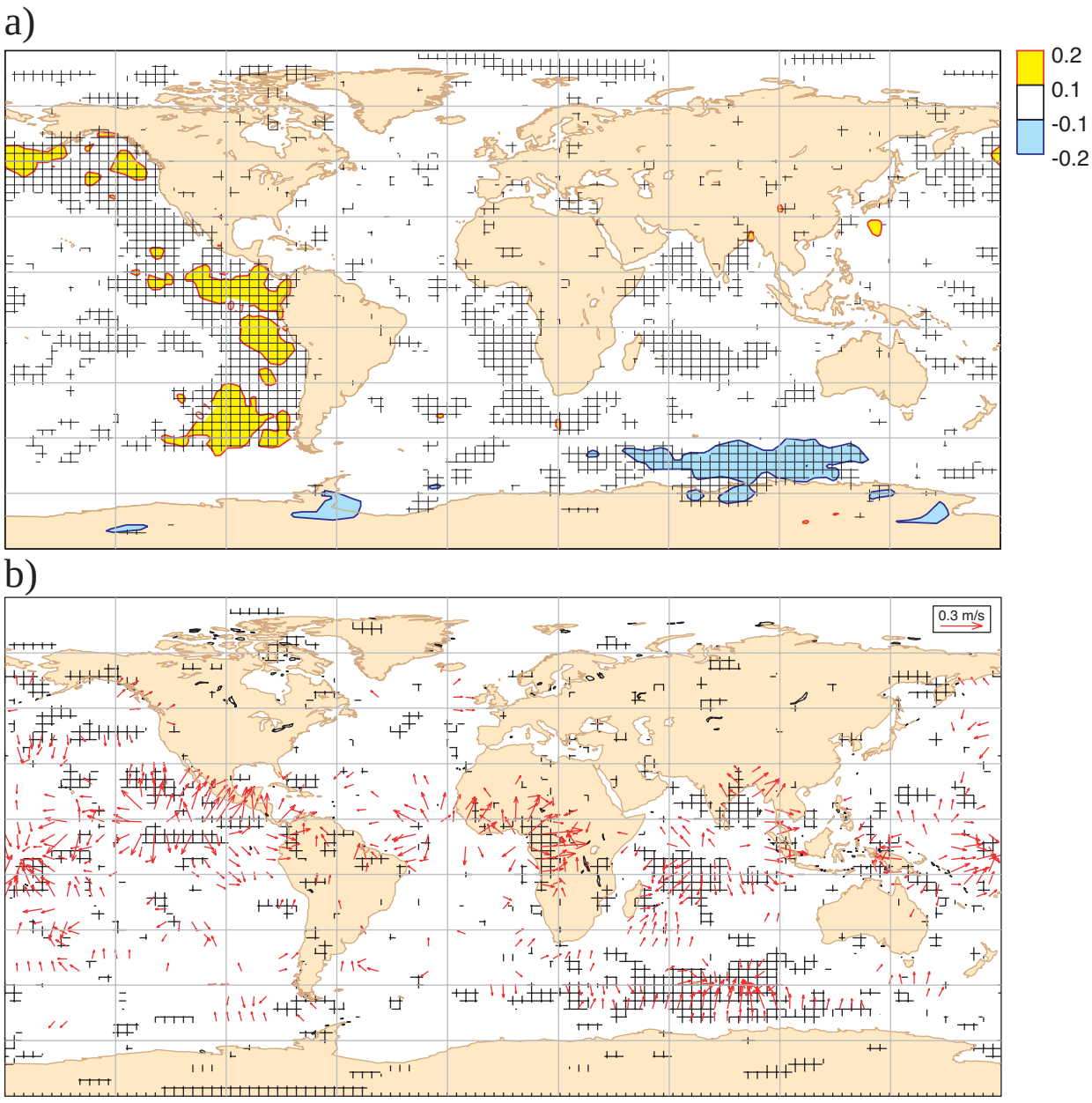


Figure 16: Mean surface pressure (a) and divergent wind component (b) difference (hPa at mean sea level and m/s at 850 hPa) from RAIN and NORAIN analyses. Calculated over period August-October 2004. Crosses indicate statistical significance (95% significance interval with t-test on analysis differences against zero-difference).

background error covariance statistics. However, the patterns are confined to limited areas which may suggest that the large-scale dynamical features are not very strongly affected. This would mean that an evaluation of forecasts that is mostly based on the assessment of key parameters associated with large-scale dynamics may show a small impact of the rain assimilation on forecast skill.

Summary

- The 4D-Var moisture analysis reproduces similar features to the 1D-Var analyses. For the period of August-October 2004, large-scale areas of systematic drying are exhibited along the West coast of the American continent, Africa and over 10-degree wide latitude bands in both Northern and Southern oceans. Areas of systematic moistening are smaller (Indian Ocean).
- During the forecasts, the drying signal is more stable because moistening increments tend to be converted to precipitation, removing the additional moisture within one day.
- The impact on the moisture analysis affects sea-surface pressure and divergence fields through the 4D-Var analysis. Those areas characterized by systematic drying produce increased surface pressure fields and divergent motions and vice versa.
- Information content analysis suggests that the rain affected observations have rather strong weight in the analysis if quantified per single observation. This is a consequence of the lower number of moisture related observations in general and the lack of other moisture related observations in rain areas in particular.

Table 3: Operational model changes between CY29R2 and CY31R1 (Haseler, pers. communication).

Model cycle	Implementation Date	Contents
CY29R2	28/06/2005	Assimilation of rain-affected SSM/I radiances, changes to AIRS assimilation, use of Meteosat-8 (MSG) winds, Jb statistics from latest ensemble data assimilation, refinement of surface-pressure bias correction, humidity analysis changes affecting spin-up and stratosphere, use of SMHI Baltic sea-ice analysis, convection changes, revised Gaussian.
CY30R1	01/02/2006	T799, L91. Model top level raised to 0.01 hPa. 4D-Var T95/255/799. Use grid-point humidity and ozone in 4D-Var. Revised ozone chemistry scheme, global ocean wave model resolution changed to 0.36 degress, with spectral resolution 24 directions, 30 frequencies. Jason altimeter wave height data and ENVISAT ASAR spectra used by wave model assimilation, ERS-2 SAR spectra no longer used. EPS resolution increased to T399, L62 with top model level at 5hPa.
CY31R1	planned 08/2006	Revisions to the cloud scheme including treatment of ice supersaturation and new numerics, implicit computation of convective transports, introduction of turbulent orographic form drag scheme and revision to subgrid scale orographic drag scheme, gust fix for orography and stochastic physics, air/sea interaction with qsat=98% over ocean, revised assimilation of rain-affected radiances, variational satellite bias correction, thinning of low level AMDAR data.

4.2 Forecast evaluation

Two pairs of experiments were run for the evaluation of the impact of rain assimilation on global model forecasts. The first pair is based on the model cycle with which the rain assimilation became operational in June 2005 that is cycle 29R2 (see 4D-Var analysis impact in Section 4.1). The second pair refers to an intermediate model cycle 30R2 which prepared model cycle 31R1 that is to become operational later in 2006. The first pair of experiments is referred to as *RAIN*29*R*2 and *NORAIN*29*R*2 as in Section 4.1 and was run for the months of August-October 2004. The second pair is referred to as *RAIN*31*R*1 and *RAIN*30*R*2 and was run for December 2005. Note that the 1D-Var scheme did not change between model cycles 29R2 and 30R2.

Table 3 summarizes the changes to the operational model configuration between cycles 29R2 and 31R1. The main improvements of the 1D-Var (see Section 3.5) scheme between cycles 30R2 and 31R1 are (1) a better radiance bias correction, (2) the screening of profiles that contain large amounts of frozen precipitation, (3) the inclusion of near-surface windspeed in the 1D-Var control vector and (4) the adjustment of the linearized moist physics parameterization to perform more similar to the non-linear parameterizations. For the sake of computational efficieny, all experiments were set up with 40 km spatial resolution despite the operational model upgrade to 25 km with cycle 30R1. However, the cycle 30R2 experiments used 91 model levels compared to 60 layers for cycle 29R2.

A number of other modifications between 29R2 and 31R1 can affect the moisture analysis and may interact with the impact of the rain assimilation. These are the increase in model resolution, the implementation of grid-point humidity, modifications to non-linear cloud and convection schemes and the introduction of the variational bias correction scheme (Dee 2004).

4.2.1 Global

Figures 17, 18 and 19 show the zonal mean change in forecast scores between *RAIN*29*R*2 and *NORAIN*29*R*2. Scores are based on root-mean-square (RMS) differences between the forecast and its originating analysis for the months of August-October 2004. The differences are normalised by the *NORAIN* experiment's zonal mean score at that level and latitude. Blue areas show regions where the rain assimilation improves forecast scores and red areas where the scores are degraded. An indication of statistical significance is given (cross-hatched areas) showing that small differences and differences at longer forecast times should be treated with caution.

In relative humidity (Figure 17), forecast scores are improved in the tropics at around 700 hPa over the first few days. Elsewhere there is little obvious improvement or degradation, though there is a region of degradation between latitudes of 30 and 60

Forecast scores:

Forecast evaluation is generally carried out with model analyses as a reference. The forecast root-mean-square errors (RMSE) are calculated from the forecasted fields, *FC*, and the respective analyzed fields, *AN*, with:

$$RMSE = \left[\frac{1}{N} \sum_{N} (FC - AN)^2 \right]^{1/2}$$

The errors may also be normalized to emphasize the relative difference of forecast errors between two experiments. The *score* is then calculated from the RMSE of the experiment, $RMSE_{EXP}$, and that of the control run, $RMSE_{CTRL}$:

$$score = \left(\frac{RMSE_{EXP} - RMSE_{CTRL}}{RMSE_{CTRL}} \right)$$

which produces the relative normalized difference. Positive (negative) numbers indicate deterioration (improvement) of forecast skill in the experiment.

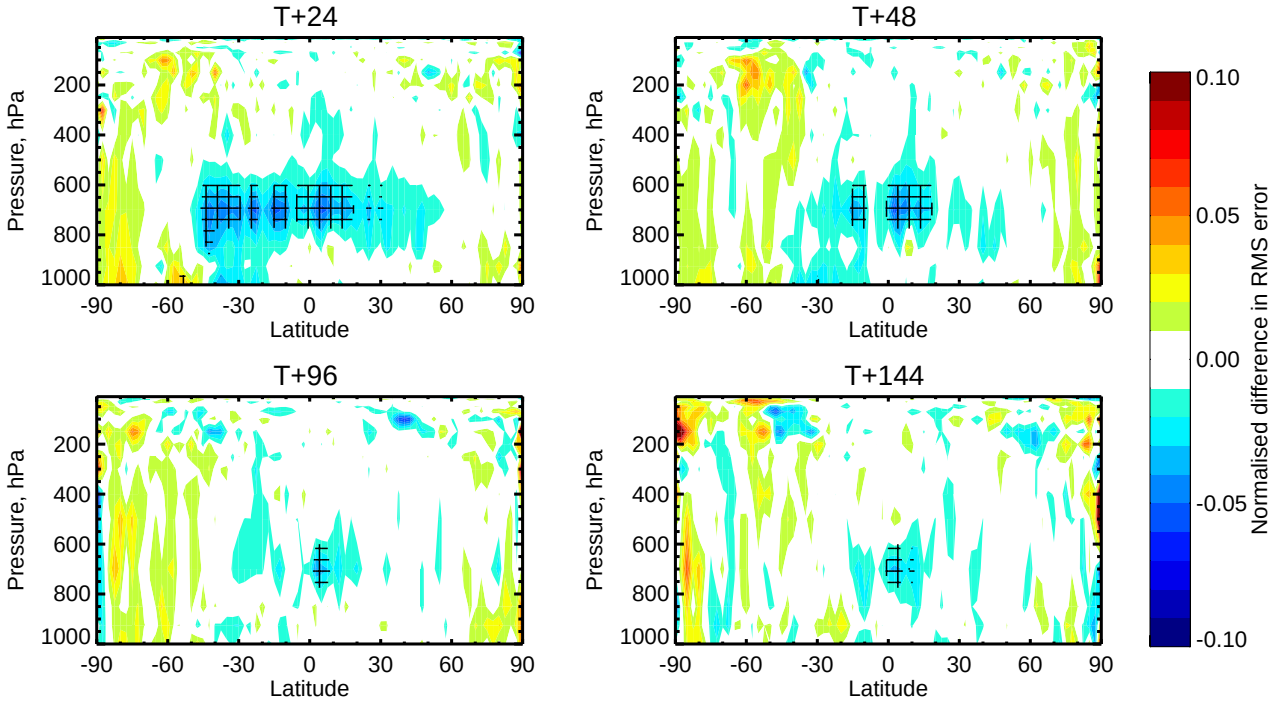


Figure 17: Normalised RMS forecast error differences in relative humidity between RAIN29R2 and NORAIN29R2, for August-October 2004. Forecasts errors are calculated with respect to own analysis. Negative (positive) contours in blue (red) indicate that RAIN29R2 performs better (worse); crosses indicate statistical significance at the 90% level.

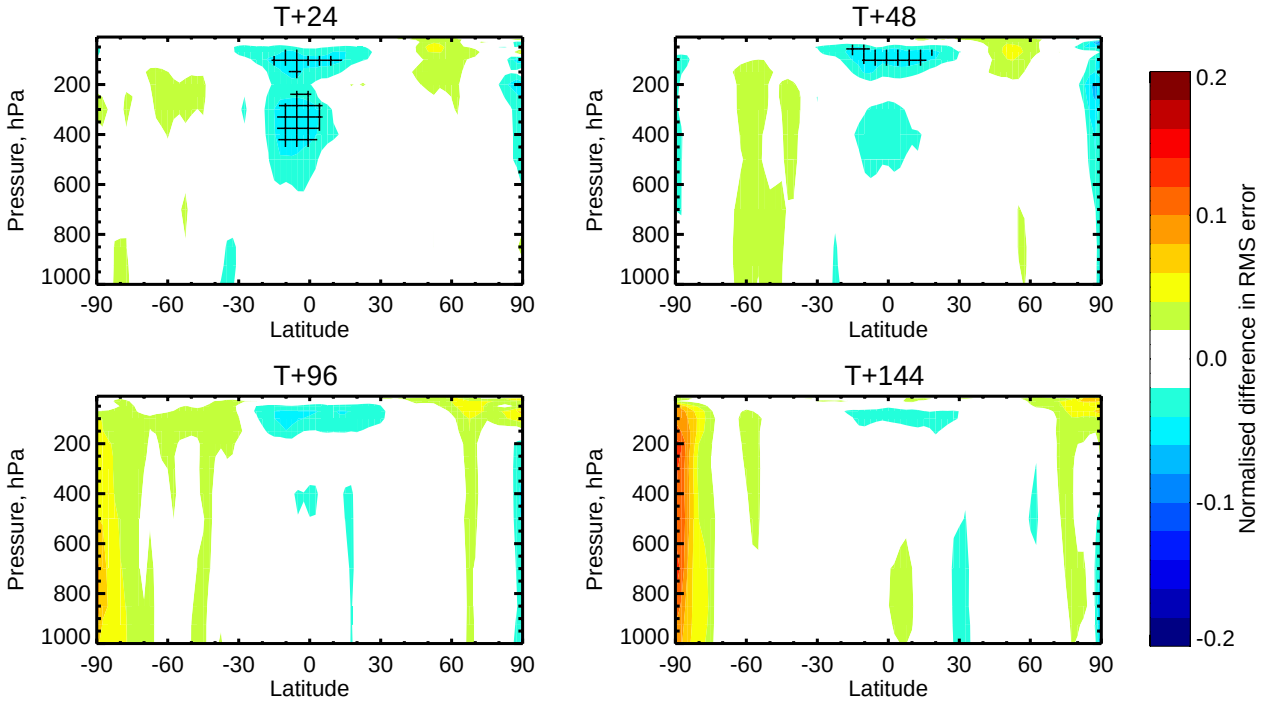


Figure 18: As Figure 17 for geopotential.

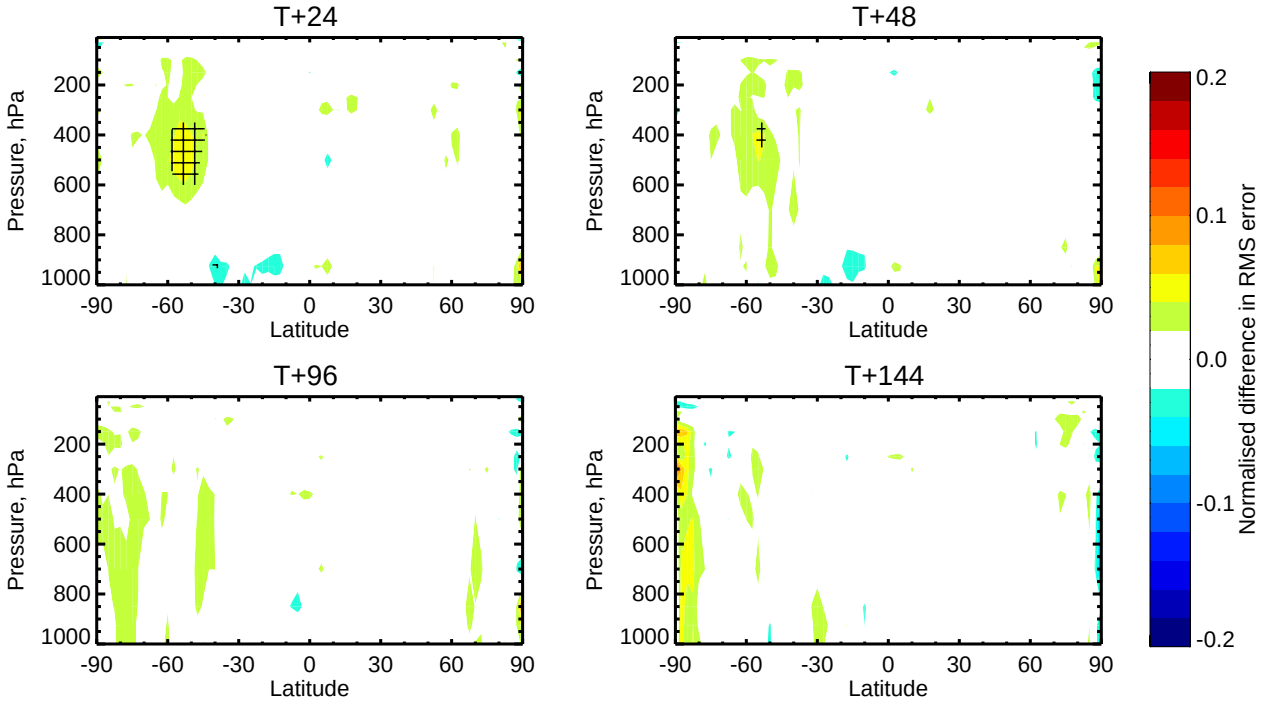


Figure 19: As Figure 17 for temperature.

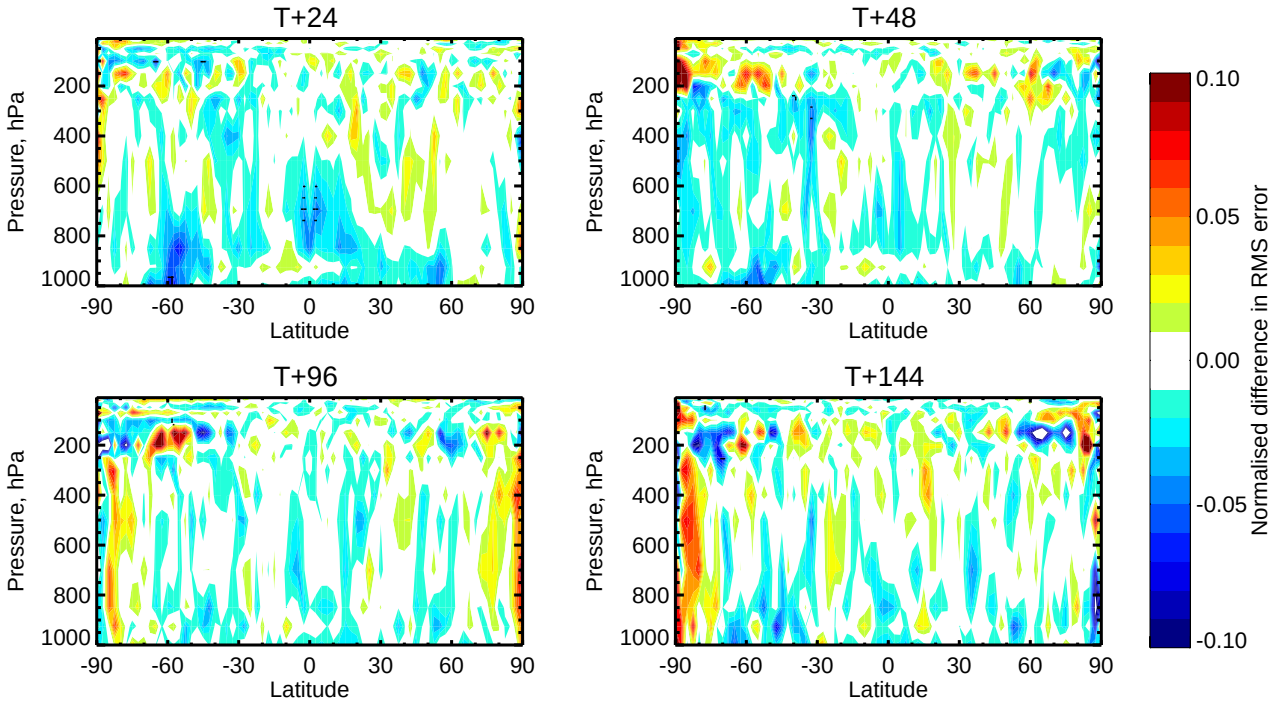


Figure 20: Normalised RMS forecast error differences in relative humidity between RAIN31R1 and RAIN30R2 for December 2005. Forecasts errors are calculated with respect to own analysis. Negative (positive) contours in blue (red) indicate that new 1D-Var scheme performs better (worse); crosses indicate statistical significance at the 90% level.

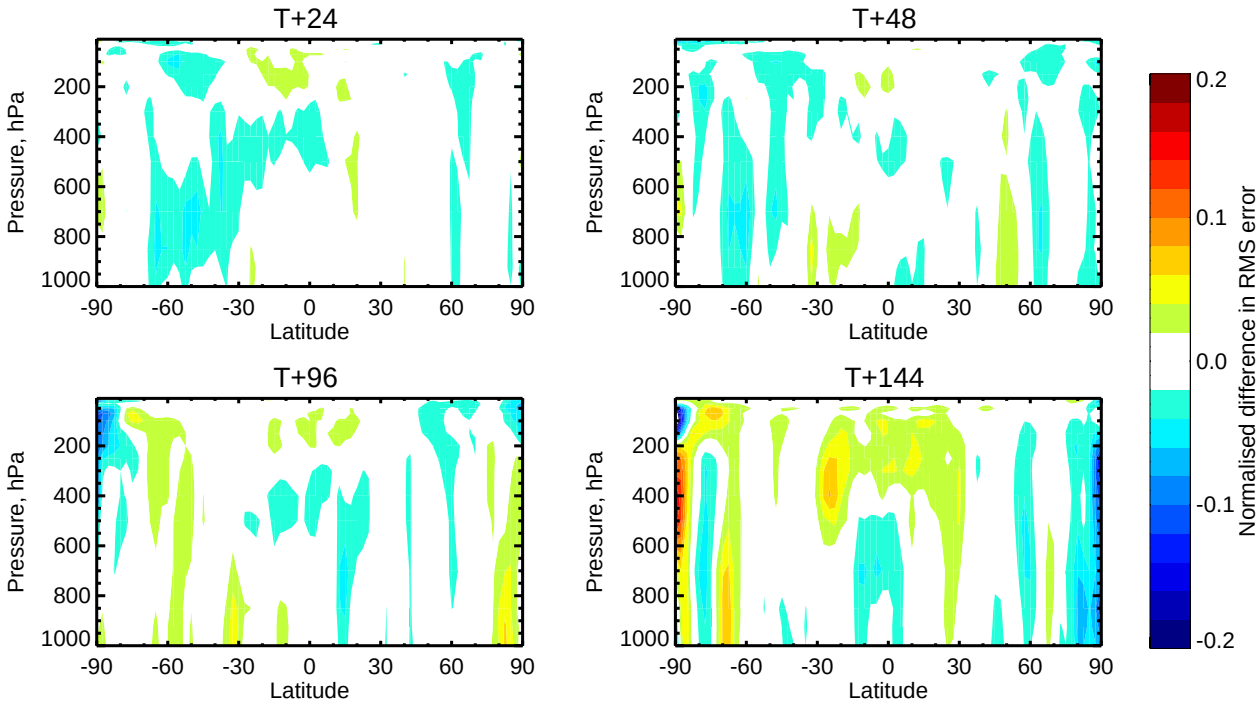


Figure 21: As Figure 20 for geopotential.

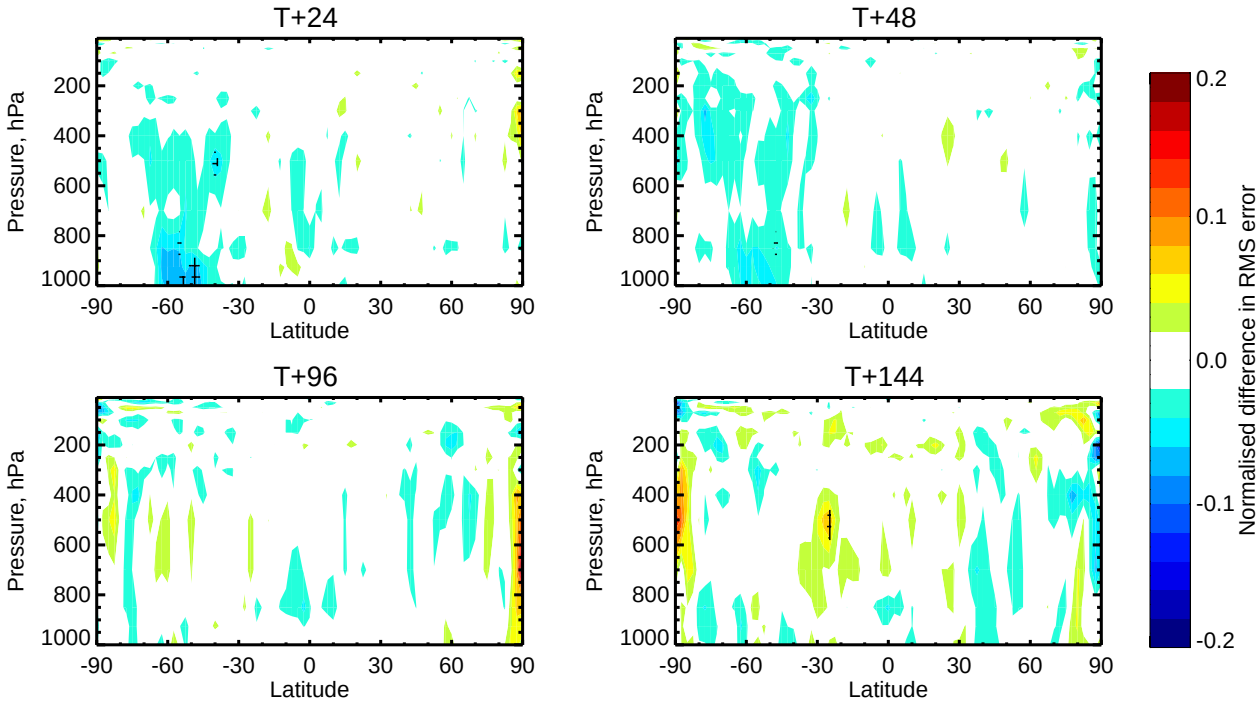


Figure 22: As Figure 20 for temperature.

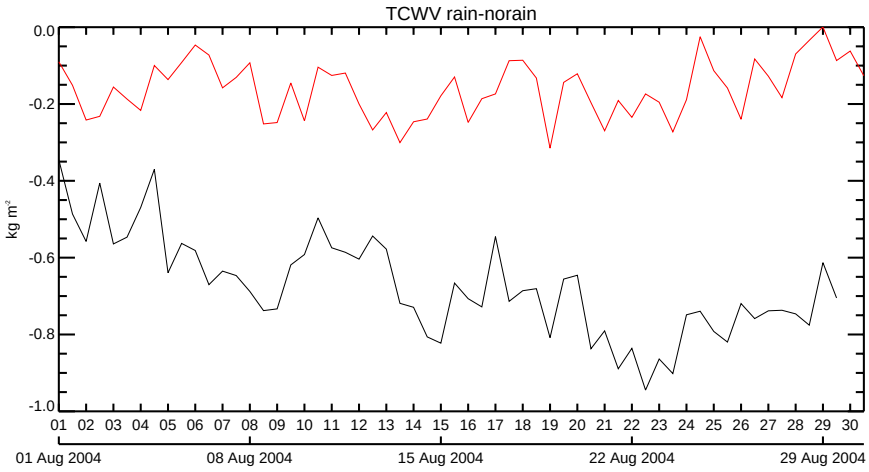


Figure 23: August 2004 mean RAIN - NORAIN difference in analysed TCWV in the region between the equator and 15 degrees North: cycle 31R1 (black); cycle 29R2 (red).

degrees South at T+48 hours. This is thought to be linked to problems of excess snow and very low TCWV amounts in the northward intrusion of cold airmasses from the Antarctic (see Section 3.5). This problem has been solved in CY31R1 by removing cases with excessive snow abundance. There are improved scores in geopotential (Figure 18) in the tropics, particularly in the upper troposphere and around the tropopause, persisting for at least the first two days. Elsewhere the changes are unlikely to be significant. There is little sign of any effect on temperature forecasts other than a small negative development near 500 hPa and at 50 degrees South that is associated with the previously mentioned area of large amounts of snowfall (Figure 19). The impact of rain assimilation on forecasts at cycle 29R2 is thus generally neutral in the extratropics and positive in the tropics.

Figures 20, 21 and 22 show the zonal mean change in forecast scores between *RAIN*31*R*1 and *RAIN*30*R*2. The revised 1D-Var scheme produces significant improvements near the surface at latitudes below 40 degrees South in relative humidity (Figure 20) and additional improvements in the tropics near 800 hPa. This confirms the effectiveness of the revised bias correction and data screening. This signal is carried forward into the geopotential forecasts (Figure 21) and to a large extent into the temperature forecasts (Figure 22). The previously noted negative temperature scores (Figure 19) at 500 hPa and at 50 degrees South were neutralized. After 48 hours, the signal dissipates for all parameters indicating that the impact of the different analyses is not memorized by the system longer than 2 days. The statistical significance is less pronounced due to the shorter experimentation period.

If the TCWV-observations are withdrawn from the analysis for model cycles 30R2 and 31R1, an important new feedback mechanism was noted which is illustrated in Figure 23. The mean difference of TCWV between the equator and 15 degrees North from *RAIN* minus *NORAIN* experiments appears to be significantly different for model cycle 29R2 and 31R1. While in model cycle 29R2, only a small systematic drying by the rain-affected observations is noted, the latest model cycles show a drying signal that is 4-5 times stronger. More detailed analysis (not shown here) reveals that this is due to the combined effect from different moisture-sensitive observations on the moisture analysis. This is mainly between SSM/I data used in the 1D-Var and the direct 4D-Var SSM/I radiance assimilation in clear skies. The latter radiances have been included in the variational bias correction scheme which dynamically updates bias corrections. This allows systematic increments from one observation type to be more efficiently absorbed in the model state by propagating this signal into the bias correction of another observation type. This mechanism is currently under investigation.

4.2.2 Tropical cyclones

The forecast skill of tropical cyclone tracks was investigated separately for the rather active Atlantic systems over the period of August-September 2004 with model cycle 29R2 and at 40 km spatial resolution. The means and standard deviations of the cyclone centre pressure differences, Δp and the centre locations, Δd, were calculated using a cyclone tracking software (van der Grijn 2002). In our analysis, between 40 and 80 track forecasts were evaluated for 1-5 day forecasts as well as the model analyses. A general feature of global model analyses and forecasts of these systems is that the analyses produce too high centre pressure due to the limited model spatial resolution thay is employed in the minimization (80-120 km).

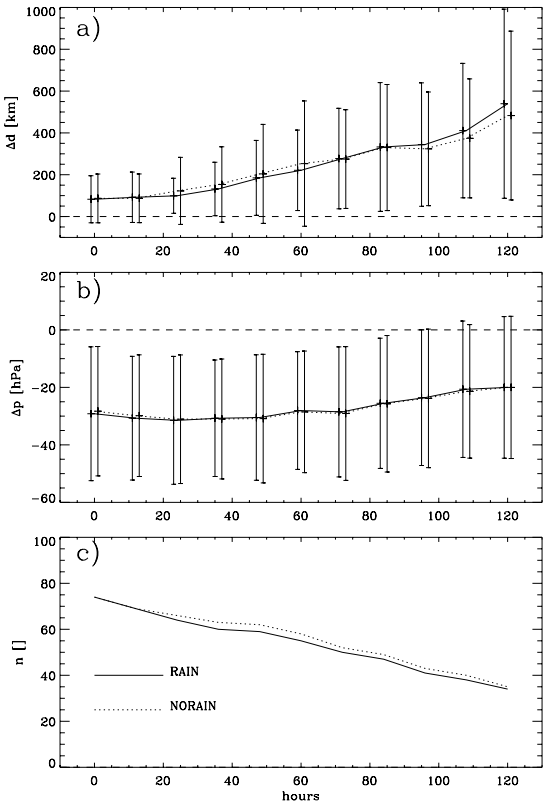


Figure 24: Mean and standard deviations of Tropical cyclone location distance error Δd (a; in km) and mean sea-level pressure error Δp (b; in hPa) and number of forecasts, respectively, for experiment RAIN and NORAIN. Error bars were offset by $\pm$ 1 hour with regard to valid forecast time to highlight differences.

Another source of error in central pressure analysis is that scatterometer data providing near-surface wind vectors is only assimilated in the vicinity of cyclones so that the drastic drop of pressure towards the cyclone's centre may be smoothened. This overestimation of centre pressure reduces during the forecast because the forecast is carried out at the highest model resolution (here 40 km).

Figure 24 summarizes Δd (Fig. 24a), Δp (Figure 24b) for the available number of cases (Figure 24c). Both mean and standard deviations for experiments *RAIN*29*R*2 (solid) and *NORAIN*29*R*2 (dotted) are shown. The sample size reduces with increasing forecast length because cyclones may dissipate and the tracking routine fails in identifying the decaying centres of cyclones.

Mean location errors increase from 100 km to 550 km. This error mainly consists of an along-track error because, apart from central pressure being systematically too high, the systems' speed is generally too low and the forecasted cyclones lag behind the observed paths. For both Δp and Δd no significant difference between the two experiments can be identified because the slight improvement in mean Δd (24-72 hours) in *RAIN* is well within the standard deviation. However, over this period the standard deviation in Δd is smaller by about 20-50% in *RAIN* which may be interpreted as an improvement.

The 4D-Var tropical cyclone analysis performance in comparison to dropsonde observations in the Caribbean Sea is displayed in the left panels of Figure 25 for the entire period between August and October 2004. Figure 25a-b shows the first-guess (thick) and analysis departure means and standard deviations of wind (Δ_w, σ_w), Figure 25c-d) the same for temperature (Δ_T, σ_T between *RAIN* (dotted) and *NORAIN*29*R*2 (solid). The sample size is 600-7,500 as a function of altitude and does not significantly change between experiments. While first-guess σ_w's are slightly smaller between 450 and 1000 hPa for *NORAIN*29*R*2, biases, Δ_w, are larger throughout the troposphere. Secondly, the analysis standard deviations are very similar while the biases of *NORAIN*29*R*2 remain larger. For temperature, a similar but weaker behaviour is observed. Again, the biases in *NORAIN*29*R*2 remain larger in the analysis.

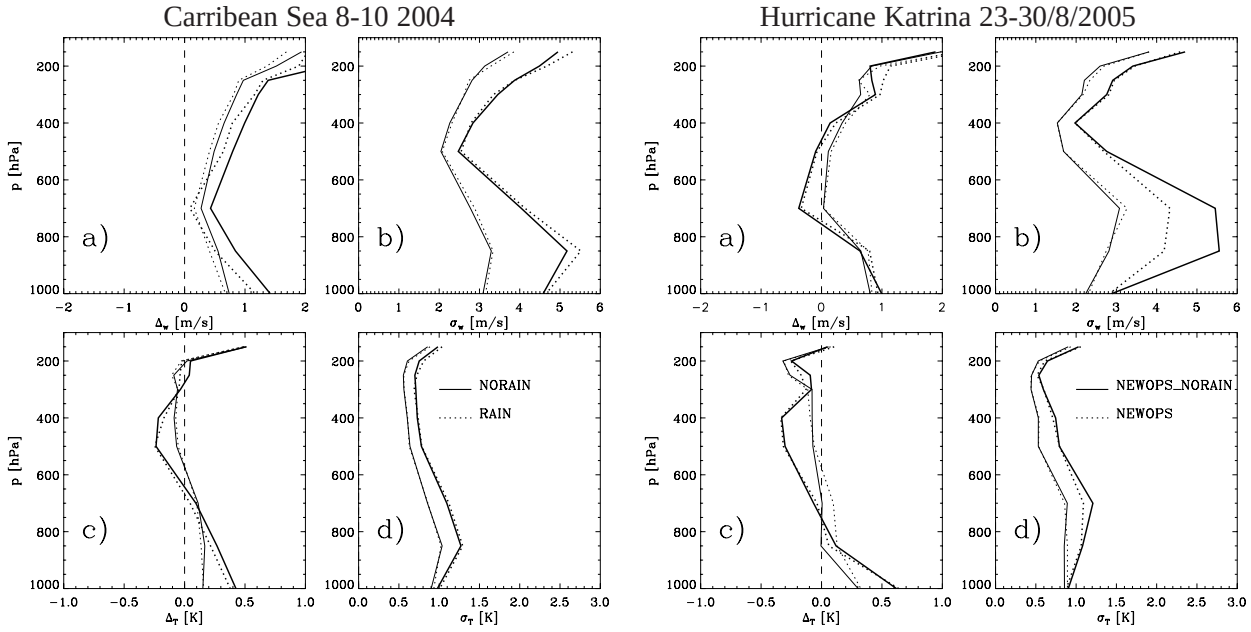


Figure 25: Biases, Δ, and standard deviations, σ, between dropsonde observations and model first-guesses (thick solid, thick dashed) or model analyses (thin solid, thin dashed) for experiment RAIN and NORAIN. Wind speed bias (a) and standard deviation (b), temperature bias (c) and standard deviation (d). Left: for period August-October 2004 in the Caribbean. Right: for period August 23-30, 2005, tropical cyclone Katrina for experiment NEWOPS (RAIN) and NEWOPS – NORAIN (NORAIN).

Similar statistics have been calculated for hurricane Katrina, which caused immense destruction in the south-east of the US in late August 2005 (Figure 25 right panels). Here, *RAIN*29*R*2 (*NEWOPS*) produces significantly smaller first-guess standard deviations in wind and slightly smaller ones in temperature. Analysis biases are slightly worse for *RAIN*29*R*2. In summary, *RAIN*29*R*2 produces better biases and, at least in the case of hurricane Katrina, *RAIN*29*R*2 produces much better standard deviations than *NORAIN*29*R*2 (*NEWOPS – NORAIN*).

Figure 26 shows the model forecasts from both experiments for hurricane Katrina in more detail. Both experiments have similar problems with the forecast of the cyclone's location in the 4.5-day forecast (Figure 26a, b). However, *RAIN*29*R*2 produces a deeper cyclone (992.1 hPa centre pressure) compared to *NORAIN*29*R*2 (996.7 hPa centre pressure) and a larger 12-hour rain accumulation (12.6 mm for *RAIN*29*R*2 vs. 10.1 mm for *NORAIN*29*R*2). Forecasts with the same target time but initialized one day later are shown in Figure 26c, d. The location is much closer to the observed one but about 50-100 km too far South in both experiments. Again, *RAIN*29*R*2 produces a slightly deeper cyclone (945.8 hPa) than *NORAIN*29*R*2 (946.5) and more precipitation (23.5 mm vs. 22.9 mm). The cyclone made landfall on August 29, 2005, at 11:21 UTC (see Figure 26f).

The operational analysis (Figure 26e) still does not reproduce the correct location and the correct centre pressure (971.1 hPa, observed: 910 hPa). The latter is a consequence of the generally limited model resolution (40 km) and the fact that the minimization is performed at very coarse resolutions (200 km). The discrepancy between spatial resolution in the minimization and forecast model also explains why forecasted cyclones usually deepen during the forecast.

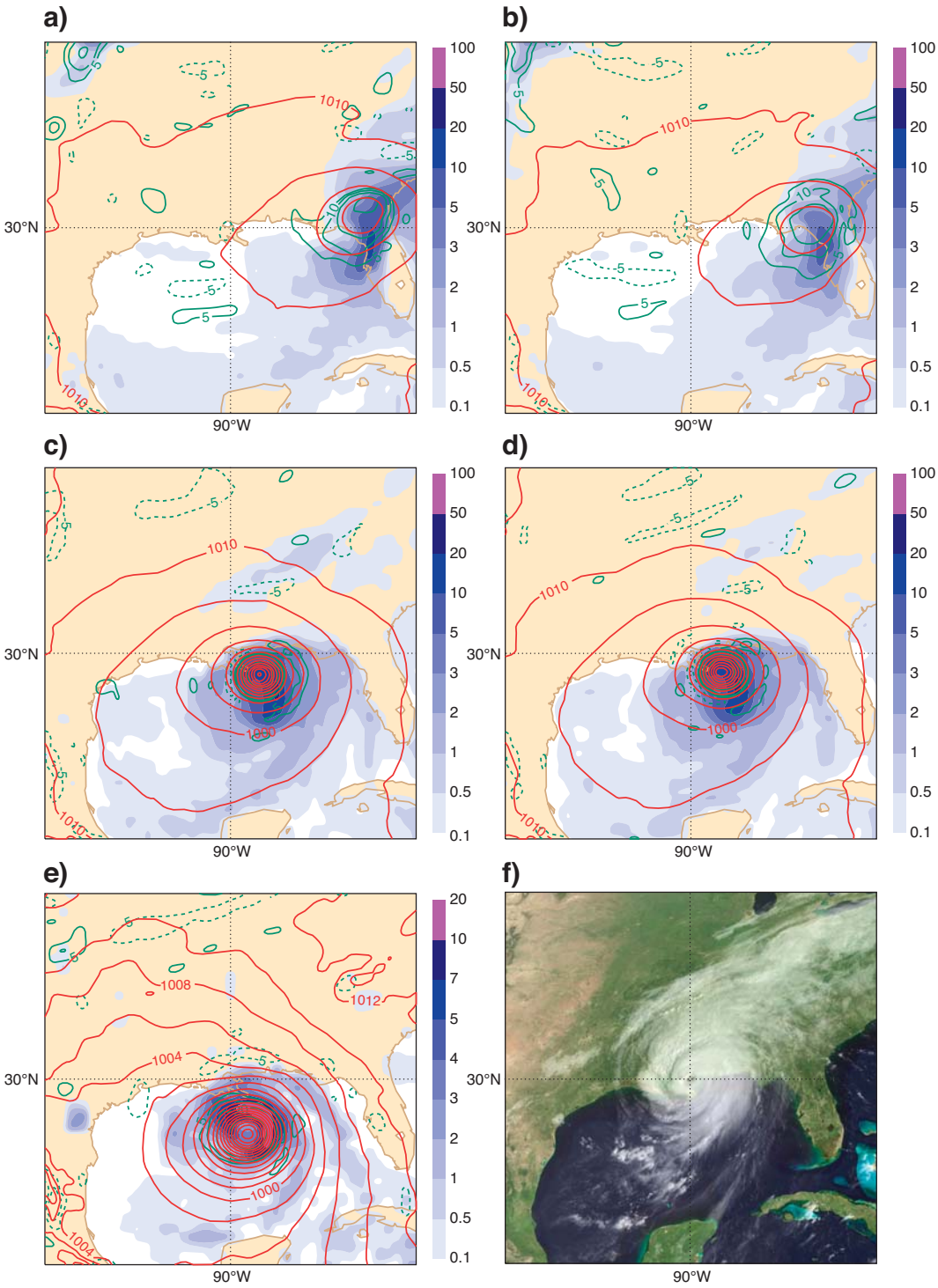


Figure 26: 4.5-day (108-hour) forecasts of mean-sea-level pressure (red isolines), 850 hPa potential vorticity (green isolines) and 12-hour accumulated rainfall (mm; blue shading) from experiment RAIN29R2 (a) and NORAIN29R2 (b) initialized on August 25, 2005, at 00 UTC. Same for 3.5-day (84-hour) forecasts from RAIN29R2 (c) and NORAIN29R2 (d) initialized on August 26, 2005, at 00 UTC. Verifying operational analysis (e) August 29, 2005, at 12 UTC and GOES-12 satellite imagery at 11:21 UTC.

Summary

- The initial version of the rain assimilation framework (model cycle 29R2) produces improved forecast scores of relative humidity in the tropics at around 700 hPa. Elsewhere there is little obvious improvement or degradation. A region of degradation between latitudes of 30 and 60 degrees South at T+48 hours is thought to be linked to problems of excess snow and very low TCWV amounts in the northward intrusion of cold airmasses from the Antarctic. This problem has been corrected in model cycle 31R1 by an improved screening and bias correction.

- The changes made to the 1D+4D-Var rain assimilation system between cycles 29R2 and 31R1 lead to neutral or slight positive impacts on forecasts. The improvements are observed in areas where large amounts of frozen precipitation reduce the sensitivity of the observation operator to moisture changes. This had been suspected to produce a slight forecast deterioration in model cycle 29R2.

- The interaction of moisture increments produced by the 1D+4D-Var system with the variational bias correction of moisture-sensitive radiance observations produces a stronger systematic drying of the northern Tropics than had been noted in model cycle 29R2. This effect is currently investigated.

- If all cyclones of the 2005 Carribean hurricane season are analyzed, no significant improvements of location and centre pressure predictions can be observed from the assimilation of rain affected observations at this stage. However, a slight improvement of location error standard deviation has been noted.

- Individual cases show a beneficial impact; a prominent example is hurricane Katrina for which central pressure and accumulated rainfall predictions were improved when rainfall data was assimilated. Both for Katrina and the remaining cyclones of the 2005 Carribean season, the fit to conventional observations in the analysis has been improved.

5 Experimental implementation of direct 4D-Var radiance assimilation

5.1 Demonstration

As noted in the introduction, the direct 4D-Var assimilation of rain-affected radiances is usually preferred because of the more direct impact on the entire model state vector and therefore the stronger effect of moisture related observations on model dynamics. Figure 3 illustrated the difference between 1D+4D and 4D-Var implementation. The main relevant differences are:

- The 1D-Var retrievals are only performed at the level of the first model trajectory that is at high resolution. The 4D-Var radiance assimilation requires forward observation operator calculations at the first and updated model trajectory runs and tangent-linear and adjoint calculations in the minimizations. The rain operator is therefore active at all stages of the analysis.

- The 1D-Var retrievals only produce TCWV pseudo-observations. Since there is no coupling of moisture and dynamics through cross-correlations in the background error covariance matrix this coupling can only be produced through the 4D-Var integration. The 4D-Var radiance assimilation allows a much more direct impact on all model state variables because the sensitivities of radiances to all relevant model state variables is represented in 4D-Var.

- The 1D+4D-Var system employs the first-guess twice because it constrains both 1D-Var and 4D-Var. This means that the TCWV observation that is used in the 4D-Var system has been constrained by the short-range model forecast and is therefore not model-independent. While this is formally inaccurate it avoids too strong discrepancies between TCWV pseudo-observations and model first guesses in 4D-Var.

- The 1D-Var retrieval step represents an opportunity for additional quality control given the 1D-Var minimization performance before TCWV pseudo-observations are assimilated in 4D-Var. The 4D-Var radiance assimilation must apply quality control beforehand and may be more vulnerable to erroneous or inappropriate observations. However, the lessons learned from the 1D-Var retrieval performance in ECMWF operations can help to refine 4D-Var radiance quality control.

- With the introduction of the variational bias correction (model cycle 31R1; Auligné and McNally 2006) most radiance observations have adaptive bias corrections while all other observations have static corrections. This will also apply to the 4D-Var rain-affected radiance assimilation system and produce a more consistent model vs. observation bias treatment among all radiance observations with sensitivity to atmospheric humidity.

Table 4 shows an example of 4D-Var timings to assess the computational efficiency of one scheme vs. the other and against a control experiment without rain assimilation. The 1D+4D-Var scheme always imposes large additional computing effort (+50%) on the first trajectory calculations but remains neutral for all other steps of the analysis. The 4D-Var scheme is much less expensive in the first trajectory (-30%) and only slightly more expensive (+5%) in the other steps of the analysis. This suggests that the net computing time for a future 4D-Var scheme will be reduced compared to the operational 1D+4D-Var scheme.

Figure 27 demonstrates the impact of the different systems on 5-day forecasts of cylone Matsa in the East China Sea for 5 August 2005 at 00 UTC. All forecasts were initialized with the same operational model first-guess state so that the different forecasts result from the use of rain-affected radiances in the 1D+4D-Var framework (Figure 27a), the 4D-Var framework (Figure 27c) compared to no rain-affected observations (control, Figure 27b). The verifying analysis (Figure 27d), surface radar (Figure 27e) and independent satellite radiometer

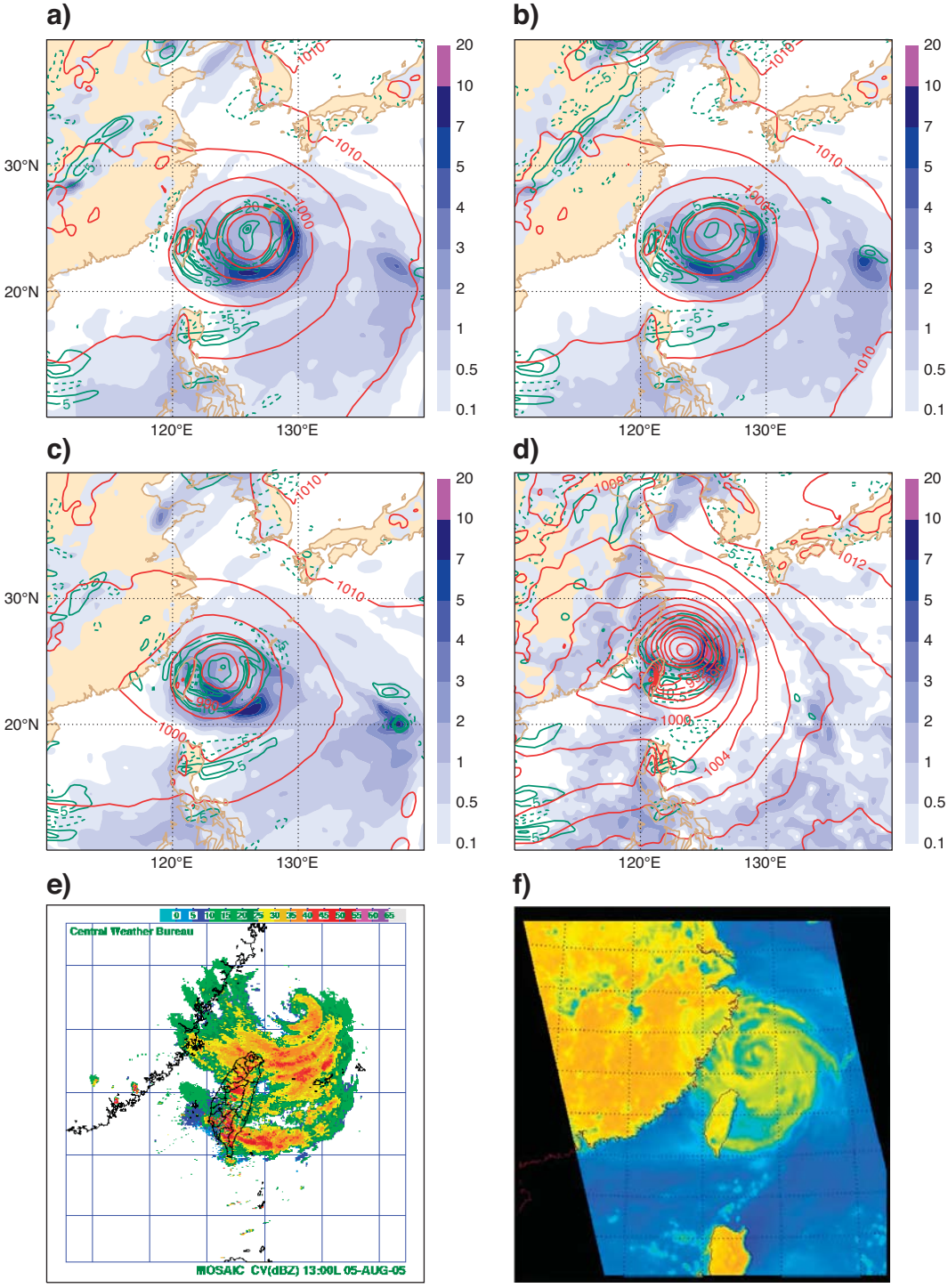


Figure 27: Tropical cyclone Matsa approaching China on 5 August 2005 at 00 UTC from 5-day forecasts of the 1D+4D-Var rain assimilation (a), 4D-Var without rain assimilation (b), 4D-Var rain assimilation (c), and the operational model analysis (d; rain amounts are from 3-hour forecast initialized at analysis time). Colour shading shows rain intensity in mm/h, red isolines are mean sea-level pressure and green isolines are 850 hPa potential vorticity. Taiwan Central Weather Bureau radar reflectivity at 13 LT (e) and the corresponding AMSR-E 10 GHz imagery (f).

Table 4: Total wall clock timing (in s) of analysis system using 384 processors (48 MPI x 8 OpenMP) for various rain assimilation configurations (Analysis/1st minimization/2nd minimization). Experiments were run with 91 model levels.*

	1D+4D-Var	4D-Var	4D-Var (no rain assimilation)
Model resolution: T799/T95/T255*			
1st trajectory	1092	-	721
1st minimization	456	-	456
2nd trajectory	563	-	563
2nd minimization	3146	-	3146
Model resolution: T799/T159/T255*			
1st trajectory	1223	766	-
1st minimization	1075	1138	-
2nd trajectory	549	540	-
2nd minimization	3390	3504	-

imagery (Figure 27f) are shown as well. Note that all experiments were run at 40 km spatial resolution while the verifying analysis from the operational model has 25 km spatial resolution.

The 1D+4D-Var method does not change position or shape of the rain distribution compared to the control experiment but increases precipitation intensity by a factor of 2. The forecasted magnitude of precipitation is confirmed by the analysis despite the inaccurate forecast of the cyclone centre position and centre presssure. The 4D-Var framework produces a narrower cyclone that is slightly displaced to the West with respect to the 1D+4D-Var method. The North-Eastern rainband is under-represented compared to the 1D+4D-Var system and the radar imagery. However, it has to be noted that all forecasts are based on first-guesses that originate from previous analyses containing 1D-Var retrievals of TCWV.

The current experimental implementation of the full 4D-Var system is technically sound and will be tested and evaluated over longer periods. It will be further upgraded by improved versions of the linearized moist physics parameterizations that are prepared for operational implementation in October 2006. The computational efficiency of the system will be increased compared to the 1D+4D-Var system and the impact of these observations on the analysis is expected to be stronger. Issues of observation and model background error definition are common between both implementations and require more research.

6 Summary and discussion

The first implementation for assimilating cloud and precipitation affected observations in the operational 4D-Var data assimilation system was successfully performed in June 2005 with model cycle 29R2. The methodology consists of a 1D-Var retrieval of total column water vapour inside clouds and precipitation using Special Sensor Microwave / Imager passive microwave radiance observations over oceans. The 1D-Var retrieval is otherwise constrained with standard short-range model forecasts of temperature, moisture, and near-surface wind speed, their forecast errors, an empirical estimate of observation errors and an observation operator that consists of moist physics parameterizations and a multiple scattering radiative transfer model. Previous tests had shown that, despite the complexity of the modelled processes and the generally uncertain precipitation representation in NWP models, the observation operator performed very well. 1D-Var and 4D-Var minimization do not indicate any problems with this new observation type or potential problems due to non-linear model behaviour.

After more than one year of operations and several upgrades to the system, the following scientific conclusions can be drawn:

- 1D-Var and 4D-Var analysis increments of total column water vapour are consistent. Comparably large areas of systematic drying by the analysis can be identified along the Western coast of the American and African continent that persist until day 3 of the forecast. Smaller areas of moistening, e.g. in the Southern Indian Ocean, do not maintain the moisture increment because of the moisture removal by precipitation. The drying signal is more pronounced because, near saturation, negative moisture increments are less penalized than positive increments through the moisture control variable and background error formulation.

- The initial version of the rain assimilation framework (model cycle 29R2) produces improved forecast scores of relative humidity in the tropics at around 700 hPa. Elsewhere there is little obvious improvement or degradation. A region of degradation between latitudes of 30 and 60 degrees South at T+48 hours is thought to be linked to problems of excess snow and very low TCWV amounts in the northward intrusion of cold airmasses from the Antarctic. This problem has been corrected in model cycle 31R1 by an improved screening and bias correction.

- The changes made to the 1D+4D-Var rain assimilation system between cycles 29R2 and 31R1 lead, in isolation, to neutral or slight positive impacts on forecasts. The improvements are observed in areas where large amounts of frozen precipitation reduce the sensitivity of the observation operator to moisture changes. This had been suspected to produce a slight forecast deterioration in model cycle 29R2. The interaction of moisture increments produced by the 1D+4D-Var system with the variational bias correction of moisture-sensitive radiance observations produces a stronger systematic drying of the northern Tropics than had been noted in model cycle 29R2. This effect is currently investigated.

- The analysis of tropical cyclone forecasts for the rather active Carribean season in 2004 revealed only very little systematic impact on track prediction. Only the track forecast error standard deviation was reduced. For tropical cyclone Katrina, the rain assimilation proved to produce better cylcone forecast with respect to centre pressure. Also, the fit to temperature and wind observations from radiosondes and dropsondes was improved. This indicates that despite new observations in the vicinity of cyclones, the general performance of the moisture analysis and the still insufficient model spatial resolution dominate the forecast quality.

An experimental version of a direct 4D-Var radiance assimilation in clouds and precipitation was developed. This represents the first available version in a global and operational NWP model. The implementation proves

to be computationally more efficient than the operational 1D+4D-Var system and does not deteriorate the 4D-Var system's convergence or the linearity of the 4D-Var model operator. The 4D-Var radiance assimilation can be expected to produce a significantly stronger interaction between cloud and rain affected observations and model dynamics. The experimental implementation will be tested and is expected to replace the 1D+4D-Var system in the near future. This can be considered a major accomplishment because the 4D-Var assimilation of cloud and rain affected radiances was considered impossible until very recently.

Answers to Key Questions

- *Which satellite observations are best suited for cloud and precipitation observation? Are they available from operational satellite series for continuous and real-time use?* The choice of observations is a compromise between operational availability, continuity, data reliability and quality, data coverage, regular sensitivity to cloud and rain water and near-Gaussian error statistics. Lower frequency microwave radiances (at frequencies between 10-50 GHz that is wavelengths of 3 cm to 6 mm) have the advantage of providing smooth sensitivity to cloud water and precipitation as well as water vapour. Measurements at these frequencies have been available for more than 20 years from well calibrated instruments that provide global coverage within 12 hours.

- *How accurately can satellite observations be simulated from current moist physics parameterizations and radiative transfer models?* The large-scale condensation and convection parameterizations as well as the microwave radiative transfer model operated at ECMWF, simulations of microwave radiance are very accurate and produce departure statistics that are comparable to clear-sky simulations for the same instrument.

- *How relevant is the discrepancy between assimilating observations at model grid-points or observation locations?* Assimilating observations at their proper locations requires model fields to be interpolated to observation locations. Assimilating observations at model grid-point locations requires to choose observations that are very near these grid-point locations or to average observations that can be associated with one model grid-box. When both model and observation sampling and spatial resolution are similar, as is the case for the current ECMWF model configuration and existing microwave instruments, the dicrepancy does not occur for high-resolution model trajectory calculations. However, a potential problem may occur in the minimization where spatial model resolution is significantly worse.

- *Is the current ECMWF model and data assimilation configuration suited for this purpose?* Both the 1D+4D-Var operational system and the 4D-Var direct radiance assimilation system do not deteriorate the minimization. This means that the employed observation operator and the current error formulation and bias correction are adequate. However, future research is required to optimize the meteorological impact such that background errors are more cloud/precipitation specific so that a more optimal impact on diabatic processes and their interaction with dynamics is achieved.

- *Is the assimilation of cloud and precipitation affected observations computationally efficient?* The operationally implemented 1D+4D-Var system increased the computational cost of the first model trajectory calculation by about 50% that corresponds to an increase of the entire 4D-Var analysis by 10%. The experimental 4D-Var radiance assimilation system proves to be less expensive than the operational 1D+4D-Var system.

- *Does the impact of these observations overcome known model errors that are associated with the lack of observations in cloudy areas and the less accurate modelling of the atmospheric hydrological cycle?* At present, the system is extremely well constrained by the combined effect of all observations sensitive to temperature, moisture and wind. The isolated impact of new observations is noticeable but can be optimized by improvements in model error definition and by the replacement of the 1D+4D-Var by the direct 4D-Var assimilation of radiances.

Acknowledgements

The authors are grateful to Carla Cardinali for running the information content analysis, to Gabor Radnóti for the evaluation of the 4D-Var tangent-linear approximation, to Chris Kidd (U. Birmingham) for generating and providing the TRMM PR rainfall climatology shown in Figure 1, to Thomas Jung for support with experiment diagnostics, to Angela Bendetti for running the original background error definition study, to Jean-Noël Thépaut and Adrian Simmons for their useful comments and dicussions and to Rob Hine for the great improvement of figures. There are many other people who helped along the way: Mats Hamrud, Sami Saarinen, Jan Haseler, Mike Fisher, Lars Isaksen.

References

Andersson, E., P. Bauer, A. Beljaars, F. Chevallier, E. Holm, M. Janiskovà, P. Kallberg, G. Kelly, P. Lopez, A. McNally, E. Moreau, A. Simmons and J.-N. Thépaut, 2005a: Assimilation and Modelling of the Hydrological Cycle. *Bull. Amer. Meteorol. Soc.*, **86**, 387-402.

Andersson, E., M. Fisher, E. Hólm, L. Isaksen, G. Radnóti, and Y. Trémolet, 2005b: Will the 4D-Var approach be defeated by nonlinearity? *ECMWF Technical Memorandum*,**479**, 26 pp.

Andersson, E., E. Hólm, P. Bauer, A. Beljaars, G. A. Kelly, A. P. McNally, A. J. Simmons, J N. Thépaut and A.M. Tompkins, 2006: Analysis and forecast impact of the main humidity observing systems *ECMWF Technical Memorandum*,**493**, 20 pp.

Auligné, T. and A. P. McNally, 2006: Adaptive bias correction of satellite data at ECMWF. *Proc. ECMWF/Eumetsat NWP-SAF Workshop on Bias estimation and correction in data assimilation*,8-11 November 2005, 127-142.

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, D. Salmond, A. Benedetti, and E. Moreau, 2006a: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF, Part I: 1D-Var. *Q. J. Roy. Meteor. Soc.*, in press.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, S. Saarinen and M. Bonazzola, 2006b: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF, Part II: 4D-Var. *Q. J. Roy. Meteor. Soc.*, in press.

Bauer, P., E. Moreau, F. Chevallier, and U. O'Keefe, 2006c: Multiple-scattering microwave radiative transfer for data assimilation. *Q. J. Roy. Meteor. Soc.*,**132**, 1259-1281.

Cardinali, C., S. Pezzulli, and E. Anderssson, 2004: Influence-matrix diagnostic of a data assimilation system. *Q. J. Roy. Meteor. Soc.*, **130**, 2767-2785.

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.

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

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

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

Dee, D.P., 2004: Variational bias correction of radiance data in the ECMWF system. In *Proc. ECMWF Workshop on assimilation of high spectral resolution sounders in NWP, ECMWF, 28 Jube - 1 July 2004*, 97-112.

Ellison, W.J., S.J. English, K. Lamkaouchi, A. Balana, E. Obligis, G. Deblonde, T.J. Hewison, P. Bauer, G. Kelly, and L. Eymard, 2004: A comparison of new permittivity data for sea water with AMSU, SSM/I and airborne radiometers observations. *J. Geophys. Res.*, **108**, D21, 4663, doi:10.1029/2002JD003213. ACL1-1-ACL1-14.

Errico, R., P. Bauer, and J.-F. Mahfouf, 2006: Issues regarding teh assimilation of cloud and precipitation data. *J. Atmos. Sci.*, accepted.

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.

Fisher, M., 2003: Generalized frames on the sphere, with application to the background error covariance modelling. In *Proc. Recent developments in numerical methods for atmospheric and ocean modelling, ECMWF, 6-10 September 2004.*, 87-102.

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

Heckley, W.A., G. Kelly, and M. Tiedtke, 1990: On the use of satellite-derived heating rates for data assimilation within the tropics. *Mon. Wea. Rev.*,**118**, 1743-1757.

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.

Krishnamurti, T.N., K. Ingles, S. Cooke, T. Kitade, and R. Pash, 1984: Details of low-latitude medium-range numerical weather prediction model using a global spectral model. Part 2: Effects of orography and physical initialization. *J. Meteor. Soc. Japan*,**62**, 613-648.

Krishnamurti, T.N., and H.S. Bedi, 1996: A brief review of physical initialization. *Meteorol. Atmos. Phys.*,**60**, 137-142.

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 vapour in rainy areas. *Mon. Wea. Rev.*,**130**, 43-58.

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.

Puri, K. and M.J. Miller, 1990: The use of satellite data in the specification of convective heating for diabatic initialization and moisture adjustment in numerical weather prediction models. *Mon. Wea. Rev.*,**118**, 67-93.

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.

Radnóti, G., Y. Trémolet, E. Andersson, L. Isaksen, E. Hólm, and M. Janisková, 2005: Diagostics of linear and incremental approximations in 4D-Var revisited for higher resolution analysis. *ECMWF Technical Memorandum*,**467**, 18 pp.

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.

Treadon, R.E., H.-L. Pan, W.-S. Wu, Y. Lin, W.S. Olson, and R.J. Kuliowski, 2003: Global and regional moisture analyses at NCEP. In *Proc. ECMWF/GEWEX Workshop on Humidity Analysis, 8-11 July 2002*, 33-47.

Trémolet, Y., 2004: Diagnostics of linear and incremental approximations in 4D-Var. *Q. J. Roy. Meteor. Soc.*,**130**, 2233-2251.

Tsuyuki, T., 1997: Variational data assimilation in the Tropics using precipitation data. Part III: Assimilation of SSM/I precipitation rates. *Mon. Wea. Rev.*, **125**, 1447-1464.

Tsuyuki, T., K. Koizumi, and Y. Ishikawa, 2003: The JMA mesoscale 4D-Var system and assimilation of precipitation and moisture data. In *Proc. ECMWF/GEWEX Workshop on Humidity Analysis, 8-11 July 2002*, 59-67.

Uppala, S.M., and co-authors, 2005: The ERA-40 re-analysis, *Q. J. Roy. Meteor. Soc.*,**131**, 2961-3012.

Zou, X., and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*,**124**, 2859-2882.

Zou, X., I.M. Navon, and J.G. Sela, 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus*,**45A**, 370-387.

Zupanski D., and F. Mesinger, 1995: Four-dimensional variational assimilation of precipitation data. *Mon. Wea. Rev.*,**123**, 1112-1127.

van der Grijn, G., 2002: Tropical cyclone forecasting at ECMWF: new products and validation. *ECMWF Technical Memorandum*, **No.386**, 13 pp.